Sensors and Materials, Vol. 38, No. 2 (2026) 1071–1086
MYU Tokyo

S & M 4364

Sensor-integrated Multidimensional Evaluation of Specialty Development and Quality Development in an Academic Medical Center

Fayun Huang,[1] Xianwei Zeng,[2*] and Cheng-Fu Yang[3,4**]

[1]Union Hospital, Fujian Medical University, Fuzhou 350001, China
[2]College of Artificial Intelligence, Yango University, Fuzhou, Fujian 350015, China
[3]Department of Chemical and Materials Engineering, National University of Kaohsiung, Kaohsiung 811, Taiwan
[4]Department of Aeronautical Engineering, Chaoyang University of Technology, Taichung 413, Taiwan

(Received January 18, 2026; accepted February 13, 2026)

Keywords: analytic hierarchy process, Internet of Things, electronic medical record, discipline construction, benchmarking, data envelopment analysis–Malmquist, high-quality development

We developed and validated a sensor-integrated multidimensional evaluation framework for discipline construction in a large tertiary public hospital using longitudinal data from 2018 to 2023. The framework fuses expert judgment with IoT sensor telemetry and electronic medical record data. Subjective weights were derived using the Analytic Hierarchy Process with strict consistency control (consistency ratio, $CR < 0.1$), while objective weights were obtained using the Entropy and Criteria Importance through Intercriteria Correlation methods. Final weights were determined through cross-validated convex fusion to balance expert cognition and data-driven variability. Multisource sensing and operational signals, including patient-flow time, bed turnover, real-time location system-based bed occupancy, medical-device uptime, and ward environmental conditions, were standardized and externally benchmarked against organization for economic cooperation and development hospital-activity indicators. Patient experience was incorporated as an outcome anchor using the Hospital Consumer Assessment of Healthcare Providers and Systems survey. Results showed sustained composite improvement, with the largest gains in clinical care, teaching capacity, and mid-level talent development, while progress in research translation and senior-talent pipelines remains limited. Risk-adjusted average length of stay decreases without increased readmissions, bed occupancy converges toward a 75–85% safety corridor, and efficiency dynamics follow an efficiency change–to–technical change relay. In contrast, patient-experience recovery lags behind activity normalization, indicating staffing and ward-process bottlenecks detectable through sensor-informed metrics. Overall, in this work, we reframed the high-quality development as a measurable system property. By integrating calibrated expert judgment, sensor-derived measurements, and international benchmarks within a unified framework, we provide a robust, interpretable, and transferable measurement paradigm for continuous monitoring and sensor-enabled governance in academic medical centers.

*Corresponding author: e-mail: 507251393@qq.com
**Corresponding author: e-mail: cfyang@nuk.edu.tw
https://doi.org/10.18494/SAM6177

ISSN 0914-4935
https://myukk.org/

1. Introduction

Global health systems are facing a triple burden: population aging is driving sustained demand for chronic disease management and multimorbidity care; the disease spectrum is shifting from acute infectious conditions toward non-communicable diseases and long-term rehabilitation; and public health emergencies continue to disrupt both supply and demand, reshaping care-seeking behavior and hospital operations. Under these pressures, hospital governance has gradually shifted from an emphasis on “expansion of volume” to high-quality development, which seeks the simultaneous improvement of clinical outcomes, patient safety and experience, operational efficiency, and the coordinated development of education and research talent under resource constraints. Achieving this transition increasingly relies on institutionalized, data-driven evaluation tools that are comparable, replicable, and sustainable for decision support and continuous improvement.(1–3) Conceptually, this evolution is commonly framed by Donabedian’s structure–process–outcome model(4) and by value-based healthcare principles, in which health outcomes per unit cost integrate multidimensional performance goals.(5)

Before constructing an evaluation system for high-quality development in hospital discipline construction, we conducted a systematic review of relevant domestic and international literature. Searches of PubMed, Web of Science, and China National Knowledge Infrastructure covering 2000–2023 identified core studies related to healthcare quality assessment, hospital performance management, and discipline construction. In this review, we aimed to clarify the multidimensional connotations of discipline construction, summarize commonly used methods and indicators, and compare international practices to guide indicator-system design. Internationally, early work in high-income countries relied on Donabedian’s framework and later expanded toward cross-disciplinary hospital performance management and value-based healthcare. Methodologically, tools such as the Analytic Hierarchy Process (AHP), Data Envelopment Analysis (DEA), and Balanced Scorecard have been widely applied to medical decision-making and resource allocation.(1–4) Over the past decade, the emergence of IoT, real-time sensing, and data analytics has shifted evaluation paradigms from static, retrospective assessments toward dynamic monitoring supported by continuous data acquisition.(5)

Domestically, as competition among comprehensive public hospitals intensified, institutions developed specialty strengths and key disciplines, and began exploring multidimensional evaluation models for discipline construction. Existing studies variously emphasize clinical effectiveness and safety, research output, or talent cultivation, whereas some propose broader frameworks integrating medical care, education, and research.(6) More recently, scholars have attempted to integrate hospital information system data with IoT-based sensing technologies, including device operation monitoring, ward environmental sensing, and real-time patient or bed-status tracking, to enhance objectivity and operational relevance beyond traditional indicators.(7) Despite these advances, two persistent gaps remain: a lack of long-term longitudinal tracking and the insufficient integration of subjective expert judgment with objective, sensor-derived data, which limits the stability and interpretability of evaluation results.(6–9)

Accordingly, using 2018–2023 as the observation period in a large tertiary public hospital, in this study, we propose and validate a sensor-integrated evaluation system for discipline construction oriented toward high-quality development. Methodologically, AHP is adopted as the backbone to formalize expert knowledge within a three-tier structure (goal–criteria–indicators), enabling pairwise comparisons, weight derivation, and consistency testing to ensure transparent and interpretable weighting logic.[10,11] In parallel, the objective process and behavioral data collected via hospital information systems and IoT sensing infrastructures, including patient-flow time, waiting and turnover intervals, device availability, ward environmental conditions, and safety-related events, are incorporated into the quantitative evaluation. Time-stamped sensing data (e.g., smart beds, real-time location system tracking, environmental sensors, and device telemetry) are linked to electronic medical records (EMRs) and annual reports through bed- and department-level identifiers, generating secondary indicators such as bed-occupancy proximity, bed turnover, intensive-care device utilization time, and environmental quality. This integration of AHP-based expert judgment with sensor-enabled measurements mitigates the limitations of purely subjective weighting and anchors evaluation outcomes in observable operational processes.[12–14]

Beyond fragmented indicator-based approaches, we provide an operational, multidimensional definition of high-quality discipline construction encompassing clinical care, talent development, research, and teaching. These four dimensions form a coupled system aligned with World Health Organization and value-based healthcare principles[2,5] and remain consistent with Donabedian's causal chain from structure to outcomes.[4] Statistical validation through paired comparisons and longitudinal efficiency profiling from 2018 to 2023 combines cross-sectional contrasts with time-series dynamics, enhancing robustness and policy relevance.[15,16] In application, departments are classified according to their longitudinal performance trajectories into improving, stable, or stagnant patterns, enabling dimension-specific governance actions and closing the loop from sensing and measurement to evaluation, decision-making, and iterative reevaluation. Overall, we bridge measurement science and hospital governance by integrating expert cognition with multiyear, sensor-derived operational data. It offers a reusable, interpretable empirical paradigm for evaluating discipline construction, supporting precise resource allocation, resilience assessment under external shocks, and iterative strategies for high-quality development in academic medical centers.

2. Methodology

2.1 Indicator system design and judgment matrix construction

In this study, the term sensor-integrated is used in a concrete and technical sense, referring to the explicit incorporation of physical and cyber-physical sensing systems into the evaluation and governance framework, rather than as a conceptual metaphor. Multiple categories of sensors and sensor-enabled devices are deployed at the ward and department levels to capture high-resolution operational signals that are not available from administrative records alone. These include Real-time Location System sensors integrated into smart beds and patient-flow infrastructure to track time-stamped bed occupancy and movement patterns, medical device telemetry embedded in critical equipment (e.g., infusion pumps, ventilators, and imaging-related devices) to record device operating hours and availability (uptime), and environmental sensors installed in wards to continuously monitor temperature, humidity, and air-quality conditions relevant to patient comfort and safety. Together, these sensing modalities provide location-aware and time-resolved measurements of capacity utilization, workflow pressure, and care-environment stability. Sensor data streams are ingested through hospital middleware and linked to EMRs and hospital information systems using shared identifiers at the bed, ward, and department levels. This integration enables sensor-derived measurements, such as occupancy proximity to safety thresholds, device saturation, and environmental stability, to be aligned temporally and spatially with clinical activity indicators, including admissions, discharges, average length of stay (ALOS), and patient-reported experience outcomes.

After preprocessing steps such as synchronization, quality control, outlier trimming, and normalization, these sensor-informed indicators are incorporated into the multidimensional evaluation framework alongside expert-derived and objective statistical weights. From a hospital governance perspective, sensor integration allows the framework to move beyond retrospective performance assessment toward measurement-driven and anticipatory management. Real-time and near-real-time sensor signals support the early detection of latent process strain, such as sustained high occupancy accompanied by declining patient experience, enable predefined staffing or workflow interventions to be triggered in a timely manner, and provide explainable inputs for decision-support dashboards. In this way, sensors function as the foundational measurement layer that connects physical hospital operations with analytical evaluation and managerial action. By systematically integrating physical sensing, digital telemetry, and administrative data, the proposed framework demonstrates how heterogeneous sensor sources can be combined with clinical information systems to support the interpretable and scalable assessment of quality improvement and specialty development at the hospital scale.

Following the principles of AHP, a three-level hierarchical architecture was constructed, consisting of a goal level (high-quality discipline construction), a criteria level (medical care, talent, research, and teaching), and an indicator level comprising 48 tertiary indicators.(10) While classical AHP typically adopts a 1–9 comparison scale, such a wide range can amplify judgment dispersion and reduce consistency when expert perceptions are weakly coupled to empirical data. To address this limitation while preserving the AHP framework, a calibrated 1–7 scale was adopted, as shown in Table 1, supplemented with quantitative evidence anchors that explicitly

Table 1
Calibrated 1–7 scale with quantitative evidence anchors. pctl: percentile

Verbal judgment	Scale	Quantitative anchor example (2018 vs 2023)
Equal importance	1	Mean diff < 0.2 *SD* or percentile diff < 5 pctl
Slightly important	2	0.2–0.35 *SD* or 5–10 pctl
Clearly important	3	0.35–0.5 *SD* or 10–15 pctl
Between 3 and 5	4	0.5–0.65 *SD* or 15–20 pctl
Strongly important	5	0.65–0.8 *SD* or 20–25 pctl
Between 5 and 7	6	0.8–0.95 *SD* or 25–30 pctl
Extremely important	7	≥ 0.95 *SD* or ≥ 30 pctl

link comparison judgments to observable data differences.(17,18) As summarized in Table 1, this calibrated scale is designed to mitigate arbitrariness in expert scoring and enhance reproducibility. Specifically, linguistic judgments provided by experts are mapped to standardized differences in key indicators, measured using mean differences normalized by standard deviation or by inter-quantile gaps, thereby aligning the assignment of AHP judgment matrices with objective data contrasts. This data-anchored calibration improves the interpretability of weight estimation and increases the pass rate of consistency control.

Within the AHP framework, the judgment matrix $A = [a_{ij}]$ is constructed to represent the relative importance of indicator i over indicator j. Each element a_{ij} satisfies the reciprocal property $a_{ij} = 1/a_{ji}$, with diagonal elements $a_{ii} = 1$, as shown in Eq. (1). At the criteria level, a four-dimensional judgment matrix is established with medical care, talent, research, and teaching as both row and column elements, as summarized in Table 2. Pairwise comparisons among these criteria yield the off-diagonal element a_{ij}, which quantifies the strength of importance of the row criterion relative to the column criterion. All pairwise comparison values are drawn from the calibrated 1–7 scale defined in Table 1, ensuring consistency with the evidence-anchored scoring scheme. The resulting criteria-level matrix serves as standardized input for subsequent weight derivation and consistency testing. By structuring the judgment matrix in this block-wise and reciprocal form, the proposed approach guarantees traceability in weight computation and facilitates the transparent verification of logical consistency across criteria, thereby strengthening the interpretability and robustness of the hierarchical weighting process.

$$A = \begin{bmatrix} 1 & a_{12} & a_{13} & a_{14} \\ 1/a_{12} & 1 & a_{23} & a_{24} \\ 1/a_{13} & 1/a_{23} & 1 & a_{34} \\ 1/a_{14} & 1/a_{24} & 1/a_{34} & 1 \end{bmatrix}, a_{ij} \in \{1,\ldots,7\} \tag{1}$$

2.2 Weight derivation and consistency testing

Within the AHP framework, criteria and indicator weights are estimated primarily using the row geometric mean method, which provides a stable and computationally efficient solution for reciprocal judgment matrices. For robustness verification, the logarithmic least squares method is additionally employed when necessary to cross-check weight stability.(10,11) The resulting priority vectors are normalized to ensure comparability across hierarchical levels. The

Table 2
Example of judgment matrix at the criteria level.

Criteria	Medical	Talent	Research	Teaching
Medical	1	a_{12}	a_{13}	a_{14}
Talent	$1/a_{12}$	1	a_{23}	a_{24}
Research	$1/a_{13}$	$1/a_{23}$	1	a_{34}
Teaching	$1/a_{14}$	$1/a_{24}$	$1/a_{34}$	1

consistency of the judgment matrices is evaluated using the consistency index (CI) and consistency ratio (CR) derived from the principal eigenvalue, as defined in Eq. (2). A threshold of $CR < 0.1$ is adopted to confirm acceptable logical consistency among pairwise comparisons. Matrices that fail to meet this criterion are returned for expert review and adjustment before proceeding to weight aggregation. This procedure ensures that the derived weights are both mathematically consistent and cognitively coherent, thereby enhancing the reliability and interpretability of the hierarchical weighting process.

$$w_i = \frac{\left(\prod_{j=1}^{n} a_{ij}\right)^{1/n}}{\sum_{k=1}^{n}\left(\prod_{j=1}^{n} a_{kj}\right)^{1/n}} \quad \text{and} \quad \lambda_{max} = \sum_{k=1}^{n}\frac{(Aw)_i}{nw_i}, CI = \frac{\lambda_{max} - n}{n-1}, CR = \frac{CI}{RI} \tag{2}$$

To control the consistency of the judgment matrices, CR is calculated as $CR = CI/RI$, where the random consistency index (RI) is taken from Table 3. When $CR < 0.10$, the judgment matrix is considered to satisfy the consistency requirement and is retained for weight derivation. Otherwise, the pairwise comparison process is revisited and adjusted to reduce logical inconsistency and subjective noise, thereby improving the reliability of the estimated weights.

To mitigate subjective bias and align weighting outcomes with observed operational dynamics over 2018–2023, AHP-derived subjective weights (W_{AHP}) are fused with objective weights (W_{obj}) obtained from the Entropy or Criteria Importance Through Intercriteria Correlation (CRITIC) methods.[19,20] The hybrid weighting scheme balances expert cognition with data-driven variability and enhances external validity across time. Then, the fused weights W^* are computed as below.

$$W^* = \alpha W_{AHP} + (1 - \alpha) W_{obj} \tag{3}$$

As defined in Eq. (3), the fusion coefficient $\alpha \in [0, 1]$ serves as a weighting parameter that regulates the relative contribution of the two constituent evaluation components in the composite scoring framework. The optimal value of α is determined through a cross-validation procedure. Specifically, α is swept across its feasible range, and the resulting composite scores are evaluated against empirical performance trajectories and efficiency patterns observed during 2018–2023. External validity is assessed using rank-based concordance metrics, including Kendall's τ and Spearman's ρ, to identify the fusion setting that best aligns with longitudinal rankings and efficiency curves. In practical implementation, objective weights are first computed from the annual time series (2018–2023). Two-period paired differences (2018 vs 2023) and characteristic efficiency inflection points, such as the peak in 2020, the decline during 2021–2022, and the rebound in 2023, are then used as external anchors for validation. Uncertainty and robustness

Table 3
RI vs matrix order n.

n	1	2	3	4	5	6	7	8	9	10
RI	0	0	0.52	0.89	1.12	1.26	1.36	1.41	1.46	1.49

are quantified by reporting 95% bootstrap confidence intervals for composite scores, along with ranking stability assessed through Top-5 and Top-10 overlap analyses. This hybrid weighting procedure ensures that final weights are both cognitively grounded and empirically consistent, thereby strengthening the credibility and policy relevance of the evaluation results.

3. Results-advanced Analytic Strategy and Findings

3.1 Global hospital-activity benchmark

To position high-quality development in discipline construction within an internationally comparable reference frame, in this study, we adopted core hospital-activity indicators from Organisation for Economic Co-operation and Development (OECD) Health at a Glance 2023 as external benchmarks.(21) The selected indicators include average acute bed occupancy (69.8% in 2021), ALOS (7.7 days), hospital discharges (130 per 1000 population), emergency department (ED) visits (27 per 100 population), and adult intensive care unit (ICU) bed density (16.9 per 100000 population). Collectively, these indicators characterize four key operational dimensions: capacity, throughput, efficiency, and critical-care resilience. From a sensing and measurement perspective, bed occupancy reflects real-time system load, with levels exceeding 85% commonly regarded as a safety risk threshold. ALOS serves as a proxy for process efficiency, while discharge and ED visit rates capture service intensity and the spillover effects of referral and triage systems. Adult ICU bed density provides an aggregate indicator of resilience in critical care capacity. As summarized in Table 4, these publicly available OECD benchmarks offer external anchors for interpreting hospital activity and efficiency beyond internal comparisons.(21)

By aligning annual in-hospital changes (2018–2023) with these benchmark ranges, variations in discipline construction scores can be interpreted as verifiable signals of operational and quality improvement, rather than as relative rankings confined to a single institution. This benchmarking strategy enhances external validity and situates sensor-informed evaluation outcomes within a globally recognized measurement framework, strengthening the interpretability and comparability of the proposed evaluation system. Figure 1 further visualizes the benchmark ranges of key OECD operational indicators, enabling an intuitive understanding of the typical macrolevel intervals for bed occupancy, ALOS, ED utilization, and ICU resource availability. This external reference framework provides a unified coordinate system for the subsequent interpretation of in-hospital temporal trends, facilitating the consistent and comparable assessment of operational and quality dynamics.

Table 4
OECD 2021 benchmarks.

Indicator	Value
Acute bed occupancy (%, OECD avg., 2023)	69.8
ALOS (days, OECD avg., 2023)	7.7
Hospital discharges (per 1000 pop, OECD avg., 2023)	130
ED visits (per 100 pop, OECD avg., 2023)	27
Adult ICU beds (per 100000 pop, OECD avg., 2023)	16.9

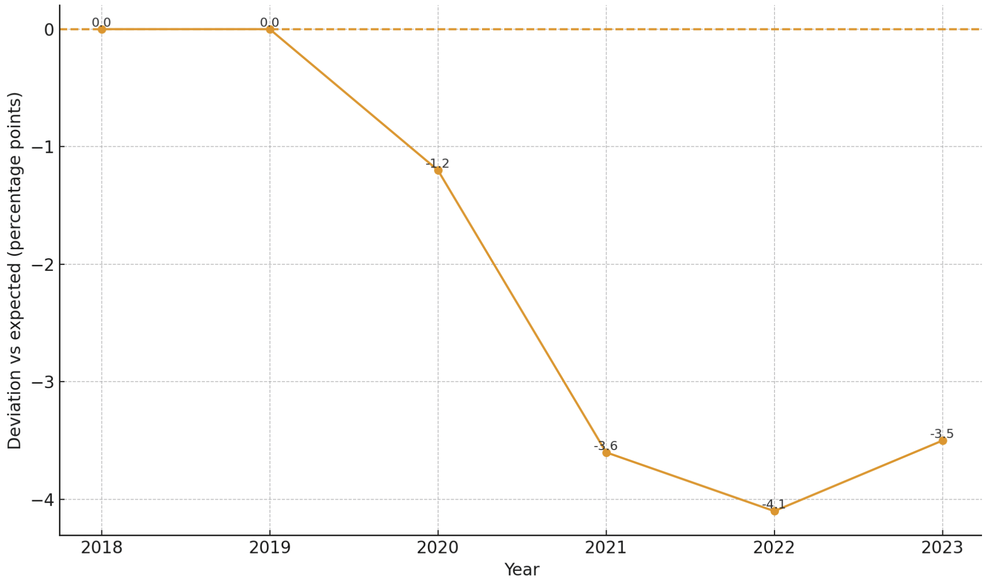

Fig. 1. (Color online) OECD 2021 key indicators (bar chart) (based on Ref. 21).

3.2 Patient-experience trajectory

Patient experience is a core outcome of high-quality development, yet its recovery does not parallel improvements in capacity-oriented metrics. In the context of this study, the term patient experience refers to patients' perceptions of the processes and interactions encountered during care delivery, rather than clinical outcomes alone. It emphasizes how care is delivered from the patient's perspective, including communication with healthcare staff, responsiveness, the coordination of services, and the overall care environment. Patient experience is therefore treated as a process-oriented and perception-based dimension of healthcare quality. Using U.S. national Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) data as an external reference, the delayed rebound of patient experience suggests that the normalization of occupancy, throughput, and length of stay alone is insufficient to restore perceived care quality. Instead, patient experience appears to be more strongly affected by ward-level processes and workforce conditions. From a governance perspective, these findings indicate that strategies centered primarily on capacity expansion or utilization optimization may have limited impact once basic safety thresholds are reached. Staffing adequacy, particularly nursing availability and skill mix, emerges as a key constraint in translating operational recovery into patient-perceived value. Sustained occupancy pressure can exacerbate communication gaps and care fragmentation, effects that are often invisible to conventional efficiency indicators but captured by patient-reported experience measures (PREMs).

Patient experience outcomes (PEOs) denote the aggregated or longitudinal results derived from patient experience measurement instruments. In this study, PEOs can be used to capture temporal changes and recovery patterns in patient-reported experience following major system shocks, such as the COVID-19 pandemic. These outcomes are analyzed as lagging but integrative signals that reflect cumulative effects of staffing adequacy, ward-level workflows, and organizational governance, rather than immediate operational capacity alone. The results highlight the value of integrating sensor-informed operational signals with PEOs. The real-time monitoring of occupancy and workflow can reveal latent process strain before declines in patient experience become evident, supporting more timely staffing adjustments and process interventions. Overall, the lagging recovery of patient experience underscores the need for governance approaches that balance efficiency management with sustained investment in workforce support and ward-level care processes. After indicator standardization, annual composite performance scores were computed as weighted sums of indicator-level scores.

Specifically, for year t, the composite score Ct was calculated by aggregating the standardized indicator scores $s_{k,t}$ weighted by their final hybrid weights W_k*, i.e., $Ct = \sum_k W_k{}^* s_{k,t}$. This formulation ensures that each indicator contributes proportionally according to its validated importance while preserving temporal comparability across years. To quantify uncertainty and ensure robustness, composite scores were accompanied by 95% bootstrap confidence intervals, generated through the repeated resampling of the indicator set. Ranking stability was further assessed by examining Top-N (Top-5 and Top-10) overlap across bootstrap replications, confirming that the relative ordering of departments and years was not driven by random variation or isolated indicators. From an implementation perspective, the computation pipeline follows three practical steps: (i) standardized indicator scores are derived annually after normalization and external benchmarking, (ii) hybrid weights are obtained by fusing AHP-based subjective weights with Entropy/CRITIC objective weights using a cross-validated fusion coefficient, and (iii) composite scores and uncertainty metrics are calculated in batch for all years (2018–2023).

All procedures are deterministic given the input data and parameters, enabling straightforward replication and extension to additional institutions or time periods. As summarized in Table 5, HCAHPS patient-experience indicators (PEIs) exhibit a persistent negative deviation in the postpandemic period, reaching deeper levels of decline and remaining below baseline without full recovery.[22–25] In Table 5, the column labeled “source hint” is used to indicate the contextual reference or external evidence source that guides the interpretation of each patient-experience time point, rather than representing a quantitative variable. Specifically, “hint” denotes the conceptual or empirical anchor used to justify whether a given observation reflects a prepandemic baseline expectation or is aligned with trends reported in external studies. This pattern indicates that the restoration of service volume and operational performance does not necessarily translate into a concurrent improvement in patient experience.

Table 5
HCAHPS deviations (2018–2023).

Timepoint	Deviation vs expected (pp)	Source hint
2018 (baseline)	0.0	Prepandemic expectation
2019 (baseline)	0.0	Prepandemic expectation
2020 Q2	–1.2	Elliott *et al.*, JAMA Health Forum 2023
2021 Q4	–3.6	Elliott *et al.*, JAMA Health Forum 2023
2022 Q3	–4.1	RAND summary (2025)
2023 Q4	–3.5	RAND summary (2025)

Instead, the findings suggest that nursing investment, communication workflows, and ward-level support processes must be explicitly incorporated into the priority agenda of discipline construction. Figure 2 further visualizes the long-term trajectory of patient-experience deviation. Following the initial pandemic shock, patient experience declined rapidly and subsequently remained in a prolonged state of negative deviation, with only gradual and incomplete recovery over time. From a sensing and measurement perspective, this trajectory highlights patient experience as a lagging but integrative outcome indicator, reflecting cumulative strain in frontline processes that may not be fully captured by conventional efficiency metrics alone. Collectively, this evidence supports the governance implication proposed in this study: high-quality development must treat efficiency, safety, and patient experience as a coupled objective system, rather than inferring overall quality improvement solely from short-term gains in efficiency or throughput. Integrating sensor-informed operational indicators with PREMs provides a more comprehensive and reliable basis for evaluating whether operational recovery has genuinely translated into sustainable improvements in care quality.

3.3 Integrated reading for discipline development

By jointly interpreting Sects. 3.1 and 3.2, three high-value insights emerge. First, a clear volume–quality recovery mismatch is observed. In this study, 2018–2019 is explicitly treated as the pre-COVID-19 baseline period, representing a relatively stable operational and patient-experience reference prior to pandemic-related shocks, as reflected in Table 5. Across many countries after 2019, hospital discharges and ED visits declined, ALOS increased slightly, and bed occupancy decreased; however, patient-experience scores deteriorated more sharply and

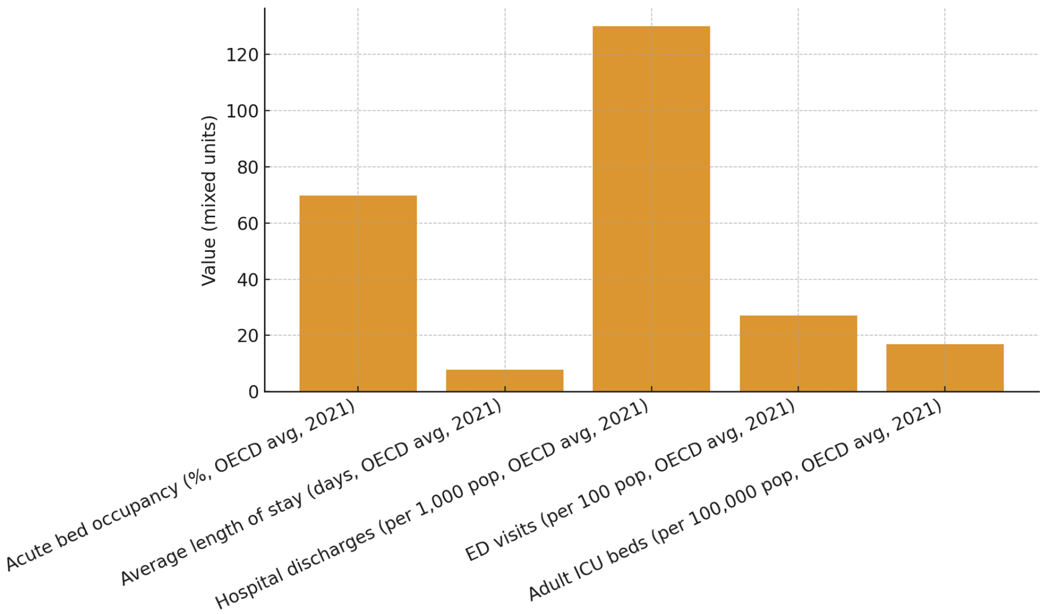

Fig. 2. (Color online) HCAHPS deviation trend (2018–2023) (based on Refs. 22, 24, and 25).

recovered more slowly than this prepandemic baseline. By anchoring post-2020 trajectories to the 2018–2019 reference years, the analysis highlights that the observed decline in patient experience exceeds what would be expected from volume contraction alone. This divergence indicates that restoring service volume or compressing length of stay alone cannot improve patient experience in parallel. Instead, investment in nursing hours, ward-level support, and communication processes is essential to translate operational recovery into perceived quality improvement.

Second, ICU resilience functions as a structural buffer during systemic shocks. The OECD average of 16.9 adult ICU beds per 100000 population highlights the importance of convertible critical-care capacity in stabilizing hospital operations and protecting other specialties under stress. From a discipline-construction perspective, this finding supports the inclusion of convertible beds, critical-care staffing ladders, and continuous monitoring capacity as strategic investment items rather than treating ICU resources solely as fixed infrastructure.

Third, these findings map directly onto the proposed indicator system. The five OECD hospital-activity indicators provide external anchors for integrating AHP-based subjective weights with objective weighting schemes, while HCAHPS serves as a core outcome for annual calibration. Operationally, bed occupancy near 80% can be treated as an upper safety reference, risk-adjusted ALOS targets should reflect case mix, and ED utilization goals should be aligned with primary-care access and triage effectiveness. When combined with continuous HCAHPS monitoring, these elements form a three-axis measurement framework, namely, capacity, process, and experience, which supports evidence-based strategic prioritization, such as strengthening nursing resources, improving discharge preparation, preserving ICU convertibility, and advancing community-front initiatives for ED-dependent specialties.

4. Discussion - Dual-Path Upgrade Centered on Evaluation System

4.1 High-level in-hospital application

At the in-hospital level, the proposed evaluation framework can be operationalized as a sensor-informed decision engine that translates hybrid weights and benchmarked composite scores into executable governance actions. With a single tertiary hospital as the analytical unit, the framework enables real-time decision support and causal validation by linking implementation measures to measurable outcomes through quasi-experimental evaluation. By embedding externally benchmarked indicators and longitudinal sensing data into routine governance, the system moves beyond descriptive scoring toward actionable, outcome-oriented management. In practice, governance targets are defined as dynamic corridors rather than static thresholds. Acute bed occupancy is maintained within a 75–85% safety band, ideally near 80%, whereas ALOS is risk-adjusted by case mix and benchmarked against upper–middle OECD reference levels. ED volume is interpreted in relation to referral and triage capacity, and ICU resilience is managed through convertible critical-care capacity and the continuous monitoring of device operating hours. Patient experience is incorporated as a core outcome constraint using HCAHPS or equivalent PREMs. On the basis of observed postpandemic recovery trajectories, a feasible annual improvement of approximately 0.5–1.0 percentage points is set, with emphasis on high-elasticity domains such as nursing availability, discharge instructions, care coordination, and pain management.

In this design, we can explicitly recognize that patient experience lags behind capacity normalization and therefore requires targeted staffing and process interventions rather than volume-based optimization alone. Sensor-informed trigger rules further translate measurements into timely interventions. For example, sustained occupancy above 85% combined with stagnating patient experience signals the need for the immediate reinforcement of nursing hours and ward support. Rebound increases in length of stay accompanied by readmissions activate cross-department discharge-preparation pathways, while the saturation of ICU device hours coupled with slowed ward throughput prompts the deployment of convertible ICU capacity. Explainable monitoring tools are applied to decompose composite scores, producing ranked lists of positive drivers and negative contributors that directly guide departmental decision-making. By making composite scores explainable at the indicator level, the system avoids "black-box" governance and enhances managerial trust in sensor-driven recommendations. Causal effects are identified around the go-live of the governance bundle using interrupted time-series or synthetic control designs with external benchmarking series, with quarterly average treatment effects and uncertainty intervals reported. Efficiency dynamics are further decomposed using window-based DEA and the Malmquist index, separating efficiency change (EC) from technical change (TC).

Empirically, an EC→TC relay pattern is observed, indicating that early gains are driven by process governance and staffing, followed by deeper gains from equipment and technology adoption. This sequence highlights that technology investments are most effective when preceded by organizational and process stabilization. Resource allocation is guided through multiobjective optimization, balancing composite performance, cost containment, and equity considerations. Strategy bundles, such as incremental nursing reinforcement, discharge-pathway expansion, and equipment upgrades, are implemented on a quarterly or semiannual cadence and continuously reevaluated, forming a closed loop of recommend → implement → verify → reoptimize. Under external anchors, clinical efficiency reaches a new equilibrium within approximately two to four quarters, patient experience improves gradually, and research translation and high-level talent development are activated through small agile units combining clinical leadership with data and methodological expertise. Together, these findings demonstrate how a sensor-enabled evaluation system can function as a practical decision engine, coupling measurement, governance, and optimization to support sustained high-quality development within hospitals.

4.2 Regional–national expansion – from evaluation framework to sensor-informed policy lever

Beyond single-hospital application, the proposed evaluation framework can be extended to the regional and national levels, transforming it from an institutional assessment tool into a sensor-informed policy lever. By integrating multihospital data, the system supports crossinstitutional benchmarking, fairness auditing, and policy simulation for integrated delivery networks, specialty accreditation, and value-based purchasing programs. This expansion leverages the same measurement architecture validated at the hospital level, ensuring that scale-up reflects true system variation rather than changes in evaluation logic. At the core of multicenter deployment is a minimum sufficient dataset assembled across participating hospitals. This dataset includes bed occupancy, ALOS, discharges and ED utilization, intensive care capacity, PEIs [HCAHPS or locally calibrated PREMs/patient-reported outcomes (PROs)], and key safety measures. PEIs are the measurable variables used to operationalize patient experience and its outcomes. In this study, they are primarily derived from the Hospital Consumer Assessment of HCAHPS survey and, where applicable, equivalent PREMs. Representative indicators include domains related to nursing availability, the clarity of discharge instructions, care coordination, communication effectiveness, and pain management. These indicators are standardized and incorporated into the multidimensional evaluation framework as outcome-level signals, complementing sensor-derived operational indicators such as bed occupancy, ALOS, and device utilization. Harmonized vocabularies, standardized definitions, and reusable metadata enable consistent comparison while preserving local data fidelity. From a sensing perspective, this step establishes a common signal space across institutions, which is essential for reliable cross-site measurement and aggregation.

To address institutional heterogeneity and unequal sample sizes, hierarchical Bayesian models are employed to estimate department-, hospital-, and region-level effects, producing posterior distributions with credible intervals rather than single-point rankings. This probabilistic representation naturally absorbs variability across hospitals and enables systematic fairness auditing through subgroup-gap analysis and risk control charts. Importantly, this approach embeds equity considerations directly into the measurement layer, allowing effectiveness and fairness to be assessed jointly rather than as separate policy afterthoughts. System resilience is evaluated through a composite gauge incorporating ICU convertibility, device operating hours, oxygen supply, and critical drug coverage, benchmarked against OECD capacity indicators to form a regional resilience radar. Empirical findings from the single-hospital analysis, showing the stabilizing role of convertible ICU capacity under demand surges, support the inclusion of resilience as a first-class measurement dimension rather than a residual infrastructure variable. The framework further enables policy simulation and threshold design.

With synthetic control or Bayesian structural time-series models, the counterfactual impacts of incorporating the composite score into payment weights or accreditation thresholds are estimated for efficiency (e.g., ALOS and discharges), patient experience (HCAHPS), and risk outcomes (safety events), with uncertainty explicitly reported. Tiered thresholds for composite and key indicators are then defined and paired with differentiated incentives, while equity and safety floors are enforced to prevent efficiency gains from eroding care quality. This design reflects the study's central finding that efficiency recovery alone does not guarantee improvements in patient experience or long-term value. Cross-country portability is ensured by anchoring the framework to OECD and HCAHPS benchmarks, while allowing the substitution of local PREMs/PROs or administrative datasets without altering the underlying evaluation logic or weight-fusion mechanism. Robust data-governance structures, including data-sharing agreements, privacy-preserving analytics, quality-control workflows, and model audit trails, provide the transparency and oversight required for policy deployment. In the medium term, the system delivers institutional distributions, fairness-audit reports, and policy-simulation outputs; in the longer term, dynamic thresholds and differentiated incentives guide regional convergence toward a balanced state of efficiency, patient experience, and resilience, with total factor productivity continuously monitored via Malmquist analysis. Together, these results demonstrate how a sensor-enabled evaluation framework can bridge measurement science and health-system policy, supporting evidence-based governance at scale.

4.3 Cross-cutting safeguards and integrated synthesis

Across both single-hospital and multicenter applications, the proposed framework is anchored in a consistent measurement architecture that prioritizes signal validity, comparability, and uncertainty control. The evaluation framework integrates a calibrated AHP (1–7 scale with $CR < 0.1$) and Entropy/CRITIC objective weighting under a unified standardization pipeline, with external anchors provided by OECD hospital-activity benchmarks and HCAHPS or equivalent patient-reported measures. By maintaining identical indicator definitions, normalization rules, and weight-fusion logic across scales, the framework ensures that observed differences reflect true system-level variation rather than measurement artifacts. Measurement risks commonly encountered in complex health systems, such as indicator-definition drift, reporting latency, and context sensitivity, are addressed through explicit metadata alignment, temporal harmonization, and adaptive calibration using local PREMs/PROs where appropriate. Subjective bias in expert-derived weights is constrained through sensitivity analyses and perturbation testing, demonstrating that composite outcomes remain stable under plausible specification changes. These safeguards position the proposed framework as a robust sensing-and-aggregation layer, rather than a static scoring tool.

Quality control is implemented as a continuous measurement loop. Composite scores are accompanied by uncertainty quantification, effect-size reporting, and explainability analyses that trace system-level signals back to indicator-level contributors. Ranking stability checks and causal validation requirements are applied before translating measurement outputs into governance or policy actions, preventing the overinterpretation of transient noise or short-lived fluctuations. In this sense, the framework treats hospital performance indicators analogously to sensor signals: subject to calibration, noise filtering, validation, and context-aware interpretation. From a systems perspective, the staged rollout, from initial data alignment to optimization and multicenter expansion, demonstrates the scalability of the measurement design without compromising interpretability. Overall, this work reframes high-quality development as a measurable, sensor-informed system property, showing how heterogeneous operational data, patient-reported signals, and expert knowledge can be coherently fused into a reproducible and extensible measurement paradigm. By emphasizing measurement rigor alongside practical applicability, the proposed approach contributes a generalizable foundation for sensor-enabled performance evaluation in academic medical centers and health systems.

5. Conclusions

From 2018 to 2023, discipline construction in the studied academic medical center demonstrated a sustained upward trend, with the most notable improvements observed in clinical care, teaching capacity, and mid-level talent development. By integrating expert judgment with longitudinal IoT sensing data and EMRs, in this study, we established a sensor-enabled, quantitative evaluation framework that supports the continuous monitoring of high-quality development rather than isolated or static assessment. Empirical evidence indicated that governance reforms and process redesign have been progressively embedded into routine operations. Improvements in risk-adjusted ALOS, the convergence of bed occupancy toward an approximately 80% safety corridor, and the recovery of service throughput suggest that efficiency-oriented interventions can reach a stable operational equilibrium within a relatively short period. However, these gains do not automatically translate into proportional improvements in patient experience or long-term system performance. Consistent with international benchmarks and HCAHPS trajectories, patient experience recovery lags behind capacity normalization, underscoring the limitations of volume- or efficiency-driven strategies alone. Across dimensions, research translation and high-level talent formation remain the primary bottlenecks to further advancement. Robustness analyses consistently show that composite performance gains driven mainly by clinical and teaching improvements tend to plateau unless accompanied by parallel progress in innovation capacity and senior talent structure. This finding highlights the role of research and talent as a second development engine for sustaining value creation. Methodologically, the proposed framework demonstrates strong stability, interpretability, and explanatory capability. The integration of consistency-controlled AHP weighting with objective Entropy/CRITIC methods and sensor-derived operational indicators ensures that composite scores reflect genuine system dynamics rather than measurement artifacts. Importantly, the framework is transferable across institutions. With internationally recognized external benchmarks and reproducible data-processing procedures, it provides a scalable, evidence-based foundation for multicenter evaluation and policy analysis, advancing high-quality development from a conceptual goal to a measurable and optimizable system property.

Acknowledgments

This work was supported by the Fujian Province Innovation Strategy Research Plan Project (2023R0035) and Production, Teaching, and Research of the Fujian Provincial Science and Technology Department (2023H6025). This work was also supported by Summit-Tech Resource Corp. and by projects under Nos. NSTC 113-2221-E-390-011 and NSTC 114-2622-E-390-001.

References

1 https://www.ncbi.nlm.nih.gov/books/NBK222274/ (accessed April 2001).
2 https://apps.who.int/iris/handle/10665/43470 (accessed June 2023).
3 A. Achrekar, S. Akselrod, G. Cuevas Barron, M. A. Charles, H. Clarke, K. Dain, and J. Koonin: Lancet Glob. Health **6** (2018) e1196.

4 A. Donabedian: Milbank Quarterly **44** (1966) 166.
5 M. E. Porter: N. Engl. J. Med. **363** (2010) 2477.
6 J. Braithwaite, R. L. Wears, and E. Hollnagel: Int. J. Qual. Health Care **27** (2015) 418.
7 G. Noto and P. Belardi: Healthcare **8** (2020) 50.
8 R. S. Kaplan, M. E. Porter, and M. L. Frigo: Strateg. Finance **99** (2017) 30.
9 J. Øvretveit: Qual. Saf. Health Care **14** (2005) 147.
10 R. W. Saaty: Math. Model. **9** (1987) 161.
11 M. J. Liberatore and R. L. Nydick: Eur. J. Oper. Res. **189** (2008) 194.
12 L. Atzori, A. Iera, and G. Morabito: Comput. Netw. **54** (2010) 2787.
13 S. M. R. Islam, D. Kwak, M. H. Kabir, M. Hossain, and K. S. Kwak: IEEE Access **3** (2015) 678.
14 N. Mehta, A. Pandit, and S. Shukla: Health Inf. Sci. Syst. **7** (2019) 1.
15 A. Charnes, W. W. Cooper, and E. Rhodes: Eur. J. Oper. Res. **2** (1978) 429.
16 B. Hollingsworth: Health Econ. **17** (2008) 1107.
17 A. Ishizaka and A. Labib: Expert Syst. Appl. **38** (2011) 14336.
18 J. G. Dolan: Patient Educ. Couns. **73** (2008) 418.
19 D. Diakoulaki, G. Mavrotas, and L. Papayannakis: Comput. Oper. Res. **22** (1995) 763.
20 H. Taherdoost and M. Madanchian: Encyclopedia **3** (2023) 77.
21 https://www.uehp.eu/corners/oecds-health-at-a-glance-2023/ (accessed Oct. 2023).
22 M. N. Elliott, M. K. Beckett, C. W. Cohea, W. G. Lehrman, P. D. Cleary, L. A. Giordano, C. Russ, E. H. Goldstein, and L. A. Fleisher: JAMA Health Forum **4** (2023) e232766.
23 https://www.cms.gov/medicare/quality/initiatives/hospital-quality-initiative/hcahps-patients-perspectives-care-survey (accessed Nov. 2025).
24 M. K. Beckett, C. W. Cohea, D. Saliba, P. D. Cleary, L. A. Giordano, and M. N. Elliott: Med. Care Res. Rev. **82** (2025) 477.
25 https://www.cidrap.umn.edu/covid-19/us-hospital-patient-survey-reveals-worse-perceived-care-amid-covid-19 (accessed August 2023).