ISCTE IUL
REPOSITÓRIO
INSTITUTO UNIVERSITÁRIO DE LISBOA

Repositório ISCTE-IUL

Deposited in *Repositório ISCTE-IUL*:
2025-09-05

Deposited version:
Accepted Version

Peer-review status of attached file:
Peer-reviewed

Citation for published item:
Yao, S., Ma, S., Shi, L., Wu, F. & Reis, E. (2025). Balancing stakeholder interests: A Balanced Scorecard perspective on performance appraisal implementation in China's public hospitals. International Journal of Health Planning and Management. 40 (5), 1122-1139

Further information on publisher's website:
10.1002/hpm.3948

Publisher's copyright statement:
This is the peer reviewed version of the following article: Yao, S., Ma, S., Shi, L., Wu, F. & Reis, E. (2025). Balancing stakeholder interests: A Balanced Scorecard perspective on performance appraisal implementation in China's public hospitals. International Journal of Health Planning and Management. 40 (5), 1122-1139, which has been published in final form at https://dx.doi.org/10.1002/hpm.3948. This article may be used for non-commercial purposes in accordance with the Publisher's Terms and Conditions for self-archiving.

Use policy

Creative Commons CC BY 4.0
The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a link is made to the metadata record in the Repository
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Serviços de Informação e Documentação, Instituto Universitário de Lisboa (ISCTE-IUL)
Av. das Forças Armadas, Edifício II, 1649-026 Lisboa Portugal
Phone: +(351) 217 903 024 | e-mail: administrador.repositorio@iscte-iul.pt
https://repositorio.iscte-iul.pt

Balancing Stakeholder Interests: A Balanced Scorecard Perspective on Performance Appraisal Implementation in China's Public Hospitals

Shufang Yao[1], Shaozhuang Ma[2*], Lishuo Shi[3], Fang Wu[2], Elizabeth Reis[2*],

1. Institute of Asclepius Hospital Management, Guangzhou, China; iren.yao@ailibi.com
2. ISCTE Business School, BRU-IUL, University Institute of Lisbon, Lisbon, Portugal;
3. The Sixth Affiliated Hospital of Sun Yat-sen University, Guangzhou, China;

Corresponding Author:

Shaozhuang Ma: ma.shaozhuang@iscte-iul.pt, ORCID-0000-0002-2199-1400;

Balancing Stakeholder Interests: A Balanced Scorecard Perspective on Performance Appraisal Implementation in China's Public Hospitals

Abstract

In 2019, the Chinese government initiated a National Performance Appraisal for Tertiary Public Hospitals (the "National Appraisal"), but limited study has been conducted on its appraisal indicators and its implementation effects. This study aims to assess how well this system balances the interests of key stakeholders in Chinese public hospitals and examines the positive changes and concurrent challenges it has brought to hospital operations.

Utilizing stakeholder theory and the balanced scorecard, we conducted two rounds of Delphi consultations with experts (N_1 = 46, N_2 = 29). We also analyzed archival data from three hospitals' "National Appraisal" records covering 2018 to 2020. Additionally, we conducted semi-structured interviews (N = 41) with key stakeholders from these hospitals.

The results show that: (1) The key stakeholders of Chinese public hospitals include patients, hospital executives, health care workers, health authorities, and the public, but the "National Appraisal" indicators only cover the first three stakeholders; (2) The "National Appraisal" system adopted a balanced scorecard approach with emphases on internal process and patient-center; (3) The "National Appraisal" had positive impacts on hospitals, clarifying hospital strategy and prioritizing the public roles, enhancing health care quality, safety, and efficiency, as well as patient and health care worker satisfaction. However, implementation challenges arise from resource constraints, discrepancy between appraisal indicators and patient needs, misalignment between appraisal indicators and doctor's patient care practices, and the tension between nationwide standardized appraisal and local context.

This study significantly contributes to the literature by empirically examining performance appraisal implementation in an under-researched country's public hospitals, offering practical implications and policy recommendations for practitioners, managers, and policymakers.

Keywords: Stakeholder theory; balanced scorecard; hospital performance appraisal; public hospitals in China

Highlights:

- China’s National Performance Appraisal for Tertiary Public Hospitals improved hospital strategy, quality, and efficiency.
- Key stakeholders lack full representation in the appraisal, excluding the public and authorities.
- Challenges of the performance appraisal implementation include resource constraints, misaligned indicators, and local adaptions.
- Findings inform global debates on public hospital performance appraisals.

1. Introduction

In an international context, performance management in public hospitals is a complex and multifaceted process, particularly when considering the diverse stakeholders involved and their varying interests and demands. These challenges underscore the necessity for robust performance management frameworks to offer a holistic view of hospital performance and align it with stakeholders' needs and interests.[1] One such framework that has gained significant attention is the Balanced Scorecard (BSC) approach.[2,3] The BSC is a management tool that focuses on measuring performance in multiple areas, including financial, customer, internal processes, and learning and growth perspectives.[4] The BSC approach helps in translating the hospital's strategic objectives into concrete measures and targets, ensuring that stakeholders' interests are taken into consideration.[1]

In the context of Chinese healthcare, performance management challenges in public hospital are particularly pronounced. The healthcare system in this emerging country operates under a less developed healthcare infrastructure and faces distinct socioeconomic factors, including a huge and increasingly aging population. Public hospitals in China play a pivotal role in the country’s healthcare system. As of 2022, China boasts a total of 36,976 hospitals, with 11,804

falling under the public sector, thereby providing nearly 60% of healthcare services.[5] Public hospitals are stratified into primary, secondary, and tertiary levels, each fulfilling distinct roles. Tertiary hospitals prioritize advanced care and research endeavors, as specified by China's Ministry of Health. Their mandates encompass providing safe health care, managing critical illnesses, advancing medical technologies, and developing talent.

However, fiscal constraints stemming from public funding pose challenges for these hospitals, prompting a shift towards profit-oriented behaviors.[6] China's healthcare expenditure, amounting to 5.59% of Gross Domestic Product (GDP) in 2020,[7] trails behind that of developed nations, thereby compelling hospitals to generate revenue through patient health care services.[8] For instance, in 2022, government subsidies constituted only 18.2% of the total revenue of healthcare institutions in Guangdong Province, an economically advanced region in southern China, with the remaining 72.2% derived from medical income.[9] Within this context, public hospitals tend to tie health professionals' performance to financial gains,[8] resulting in unnecessary procedures, inflated medical costs, and strained relations between health care workers (HCWs) and patients. [6,10]

To effectively tackle these challenges, a shift in performance management within public hospitals was necessary. This shift aims to balance financial and non-financial indicators, meet diverse stakeholder needs, and prioritize public interests, thereby reaffirming the crucial role of public hospitals in serving the community. With such aim, the General Office of China's State Council issued the "Opinions on Strengthening the Performance Appraisal of Tertiary Public Hospitals" (referred to as the "National Appraisal" hereafter) to introduce standardized assessment criteria for tertiary public hospitals in 2019.[11]

The implementation was not an easy endeavor, but empirical studies examining the intricacies and obstacles associated with such implementations in hospitals remain scarce.[1,12] In this context, the "National Appraisal" introduced in 2019, warrants closer scrutiny to examine the

indicators and comprehend its impact on diverse stakeholders fully. Building upon existing literature and theoretical frameworks, this study first aims to understand the key stakeholders of Chinese public hospitals and the extent the "National Appraisal" covers and balances the interests of those stakeholders; this study also aims to elucidate the implementation effects of the "National Appraisal" on key hospital stakeholders and the concurrent challenges. The first aim is achieved through the Delphi method while the second aim is through quantitative archive data analysis and semi-structured interviews conducted with key stakeholders from three hospitals.

By so doing, we present a holistic account of a practical "case study" of performance appraisal implementation in public hospitals from China. Additionally, we highlight both the benefits and challenges of addressing and balancing the varied interests of key stakeholders, using the lens of stakeholder theory and the BSC. Such understanding is essential for evidence-based decision-making, policy refinement, and continuous enhancement of public service delivery. The study will contribute to the performance management literature with a long-term implementation study from a complex health care sector in a under researched country, and more importantly provide managerial implications and policy recommendations for practitioners, managers, and policymakers with real application of performance appraisal system.

2. Literature Review

2.1 Hospital Performance Management: A Multidimensional Perspective

Hospital performance, distinct from enterprise performance, involves broad economic and social outcomes that affect a wide range of stakeholders.[1,13] Recognizing the importance of hospital performance assessment, the World Health Organization (WHO) introduced "health system performance assessment" (HSPA) in the early 2000s as a critical managerial tool to achieve health system goals (health, responsiveness and financial fairness).[14,15] This

framework evolved into a system with six building blocks: service delivery, health workforce, information, medical products, vaccines & technologies, financing, and leadership/governance. Additionally, in 2007, the WHO Regional Office for Europe developed the "Performance Assessment Tool for Quality Improvement in Hospitals" (PATH) framework, which comprises four domains (clinical effectiveness, efficiency, staff orientation, and responsive governance) and two transversal perspectives (safety and patient centeredness).[16]

Developed countries boast more advanced performance management systems for healthcare organizations. However, the notable disparities in health care organization and financing among developed nations result in distinct approaches to performance assessment, as exemplified by the UK, the US, Australia, and the EU.

The UK's healthcare system, the National Health Service (NHS), operates under a nationalized, centrally driven model. The NHS Performance Assessment Framework, based on the balanced scorecard approach, includes six dimensions: health improvement, fair access/equality, effective delivery, efficiency, patient satisfaction, and health outcomes.[2] Recently, the NHS has granted health care providers greater autonomy over financial and management matters, with national indicators focusing on finance and service quality (e.g., user experience).[17,18]

In contrast, the US healthcare system is pluralistic and market-based, driven by the private sector.[19] Consequently, various hospital performance assessments have been developed by government, private, non-profit organizations, and academic institutions, using outcomes and process-driven indicators.[20] Safety and effectiveness are primary priorities in hospital performance indicators, followed by timeliness, efficiency, patient-centeredness, and equity.[20]

Australia adopts a middle ground between the UK and US healthcare approaches. Its system is pluralistic and complex, featuring universal access supplemented by market-oriented financing incentives and private sector stimulation.[19] In August 1999, the National Health Performance Committee (NHPC) was established to develop and maintain a “national health performance

framework". This framework comprises three tiers: health status and outcomes, determinants of health, and health system performance. The latter is measured by effectiveness, appropriateness, efficiency, responsiveness, access, safety, continuity, capacity, and sustainability.[5]

The European Union (EU) has developed the European Core/Community Health Indicators (ECHI), comprising 88 health indicators classified by policy area, to monitor health at the EU level.[21] Within the "health services and healthcare" policy area, dimensions such as sustainable health care systems, health system performance, quality of care, efficiency of care, and patient safety are measured with specific indicators.

In summary, different health care systems employ varying approaches to hospital performance assessment and the performance assessment approach are evolutionary and dynamic. Overall, these approaches emphasize measuring hospital performance on health care safety, effectiveness, access, efficiency, health workforce and patient-centeredness to need the diverse interests of hospital stakeholders.[20,22,23]

2.2 Performance Management of Public Hospitals in China and the "National Appraisal"

Since the founding of the People's Republic of China, the evolution of performance management in public hospitals in the country has progressed through distinct stages.[24,25]

Government top-down planned management stage (1949–1985): Initially, the government had centralized control over public hospitals, focusing on tasks and resources without a formal performance appraisal system.

Hospital Autonomy Stage (1985–1991): China's economic reforms led to decentralization, granting hospitals more autonomy in operations and budgeting, shifting towards economic interests over public roles.[6]

Marketization Stage (1992-2009): Market-oriented reforms transformed public hospitals into profit-driven entities, relying on user charges and revenue generation, deteriorating patient relationships.[6]

Structured Appraisal Stage (2009–2018): There was a notable shift towards health service quality, patient interests, and sustainability, transitioning from government-led to third-party evaluations, with a focus on patient satisfaction and comprehensive appraisals.

Balanced Approach Stage (2019–Present): Introduced a nationwide unified performance appraisal system in 2019, the "National Appraisal," standardized assessment criteria for tertiary public hospitals, prioritizing patient-centeredness, health care quality and efficiency, aiming to address diverse stakeholder.

In short, this evolution highlights the intricate dynamics among various stakeholders in Chinese public hospitals (Fig. 1) and reflects a rational shift towards a more structured and balanced approach in performance appraisal in public hospitals in China.

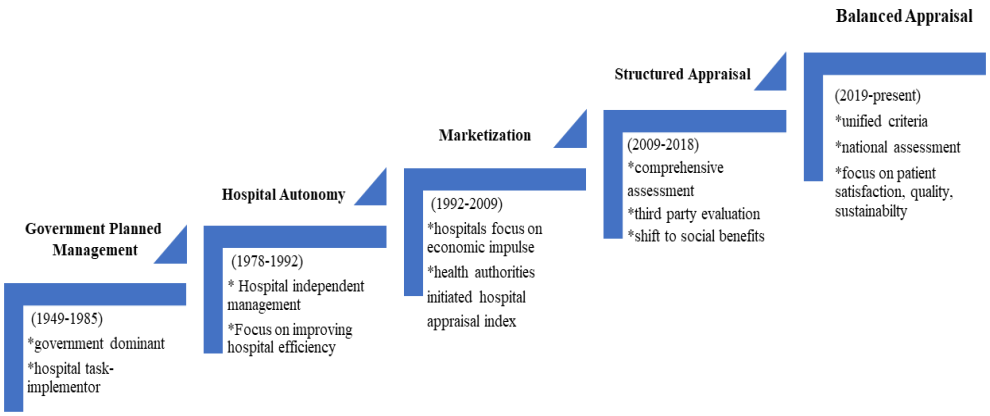


Fig. 1 – Performance Management Approaches of Public Hospital in China

The "National Appraisal" framework encompasses three tiers of performance evaluation indicators. The first tier comprises four indicators that pertain to health care quality, operational

efficiency, sustainable development, and satisfaction. The second tier includes 14 indicators, and the third tier includes 55 indicators (see Fig. 2 and Appendix Table).

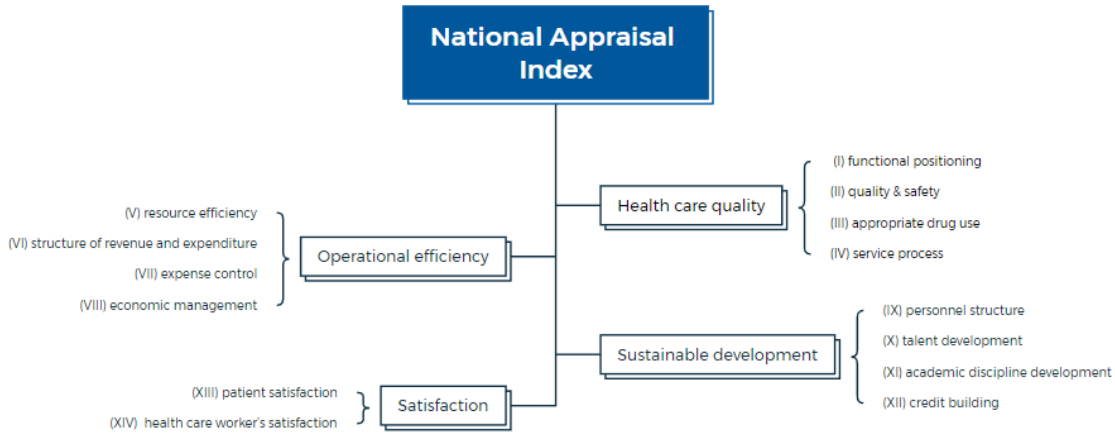


Fig. 2. Structure of National Appraisal Index (part)

As illustrated in Figure 2, under the four dimensions at 1st tier, there are respective indicators at 2nd tier and 3rd tier for assessment.

Health care quality: measured with 24 indicators at 3rd tier from four perspectives (2nd tier), including 1) functional positioning (7 indicators at 3rd tier); 2) quality & safety (8 indicator); 3) appropriate drug use (6 indicators); 4) service process (3 indicators).

Operational efficiency: measured with 19 indictors at 3rd tier also from four perspectives (2nd tier): 1) resource efficiency (2 indicators at 3rd tier); 2) structure of revenue and expenditure (10 indicators); 3) expense control (5 indicators); 4) economic management (2 indicators).

Sustainable development: measured with 9 indicators at 3rd tier from four perspectives (2nd tier): 1) personnel structure (3 indicators at 3rd tier); 2) talent development (3); 3) academic discipline development (2) and 4) credit building (1).

Satisfaction: measured with 3 indicators at 2nd tier, including 2 indicator for patient satisfaction and 1 for health care workers (HCWs)' satisfaction.

Among the 55 indicators, 26 are the key indicators annually monitored and assessed by China's national health care authority, the National Health Commission (NHC). The process is centralized through a dedicated online Performance Assessment Management Platform established and overseen by the NHC. This platform features a uniform interface for data submission to ensure consistency across all submissions. Each year, the government specifies a timeframe during which the system is open for hospitals upload their performance assessment data.

Among the 26 key indicators, particular emphasis is placed on ***health care quality*** as a pivotal dimension under scrutiny by the government. This dimension is evaluated through 10 key indicators, accounting for 38% of the total 26 indicators. ***Operational efficiency***, reflecting the degree of streamlined management within public hospitals, is measured by 9 key indicators, comprising 35% of the total. ***Sustainable development*** gauges the establishment of skilled teams and educational research capabilities within public hospitals, incorporating 4 key indicators, which account for 15% of the total indicators. ***Satisfaction*** is bifurcated into patient satisfaction and HCWs' satisfaction through 3 indicators, contributes to 12% of the overall indicators.

Given the large number of indicators of the "National Appraisal", in this study, we focus on 26 key indicators which are the national monitoring indicators that are assessed and published by heath care authorities.

2.3 Stakeholder Theory

Stakeholders are broadly defined as individuals or groups affecting or affected by an organization's operation.[26] Stakeholder theory proves vital in navigating complex stakeholder demands in hospitals, and a valuable lens to examine the performance management of public hospitals.[1] Balancing diverse interests is crucial for comprehensive hospital performance.[22]

Scholars apply stakeholder theory to public hospital management and performance appraisal, highlighting the intricate interplay between stakeholder behavior and hospital outcomes.[27] In the context of health care institutions, it is essential to identify and prioritize various stakeholders, based on their interests, significance and impact on the hospital operation.[1,28] The application of stakeholder theory to performance management in public hospitals can help identify the key stakeholders, their needs and expectations, and their influence on hospital performance.[29]

Stakeholders of public hospitals vary, including health authorities, hospital executives, patients, HCWs, the public, competitors, suppliers, industry associations, and the media.[30] These stakeholders have diverse and often competing interests.[18] Health professionals are concerned with providing high-quality care to patients and maintaining a positive work environment, while hospital administrators may focus on financial sustainability and operational efficiency. Government entities prioritize healthcare accessibility and cost-effectiveness, and patients seek timely and effective treatment, as well as a positive overall experience. Understanding and balancing these varied stakeholder interests are essential for effective performance management in public sectors like hospitals.[1] Considering the complexity of stakeholders' demands and their influence on hospital performance, this study first aims to identify the key stakeholders and their relationship with the "National Appraisal" indicators.

2.4 The Balanced Scorecard

The Balanced Scorecard (BSC) as proposed by Kaplan and Norton (2005), incorporates four perspectives: financial, customer, internal processes, and learning and growth. It is a comprehensive performance evaluation tool that has been applied in various sectors, including healthcare. The BSC framework has garnered substantial attention in hospital management for its efficacy in aligning strategic objectives with operational measures, ultimately bolstering performance and elevating patient care quality.[3,22,31] Studies emphasize that while financial

metrics are vital, non-financial indicators, such as patient satisfaction, clinical outcomes, and operational efficiency, are equally crucial in assessing hospital performance.[32] The BSC is pivotal in healthcare, as it prioritizes non-financial measures, significantly influencing patient care quality. For example, its implementation optimizes processes, reduces waste, and enhances resource utilization and operational efficiency, leading to cost savings and improved clinical outcomes.[33] Recent reviews, such as that by Amer et al. (2022), affirm the positive impact of BSC on patient satisfaction and financial performance, though its influence on HCWs' satisfaction remains marginal.

However, despite its comprehensive approach, implementing the BSC in hospitals presents challenges. First, the list of indicators used in the BSC for healthcare settings is heterogeneous, and each BSC should be developed in a unique way. Furthermore, the implementation of the BSC in hospitals involves multiple stakeholders, adding complexity to the process.[24] Taken together, the BSC is considered an effective framework for measuring performance in public hospitals, but customization is necessary to address the specific requirements of each organization.

We apply the BSC framework to understand how the national appraisal indicators are balanced for the diverse interests of various stakeholders and financial and non-financial outcomes of hospitals?

2.5 Integrating Stakeholder Theory and the Balanced Scorecard

Integrating stakeholder theory and the BSC in public hospitals' performance management is a key concern in the healthcare sector.[22] The application of the BSC in conjunction with stakeholder theory provides a comprehensive framework for performance management in public sectors like hospitals.[1] For example, the BSC approach allows for the integration of financial and non-financial measures, aligning with the diverse interests of stakeholders.[2] By

including metrics related to patient satisfaction, HCWs's learning and growth, and operational efficiency, public hospitals can address the multifaceted needs of their stakeholders.[32] Moreover, the BSC's focus on strategy implementation and continuous improvement complements the stakeholder theory's emphasis on long-term value creation and sustainability.

In order to effectively apply and implement BSC framework for performance management in public hospitals, it is essential to engage with various stakeholders, communicate openly about performance objectives, and solicit feedback to inform performance assessments.[2] In other words, stakeholder theory is essential for the successful implementation of BSC in public hospitals.

In conclusion, a combined application of stakeholder theory and the BSC perspective not only offers a robust framework for enhancing performance management in public hospitals, but also makes performance management more viable and effective (Cheng et al., 2019).[1] By recognizing the diverse interests of stakeholders and integrating a range of performance measures, public hospitals can optimize their operations and better serve their communities, ultimately improving healthcare outcomes.[32]

Specifically, this study embarks on an exploratory inquiry into the following research questions around key stakeholders, BSC and implementation outcomes of the "National Appraisal":

RQ1: Who are the key stakeholders of tertiary public hospitals in China? To what extent are the "National Appraisal" indicators related to these key stakeholders? Are the interests of the key stakeholders balanced and taken into due account?

RQ2: How should the key indicators of the "National Appraisal" be categorized into the four dimensions of the balanced scorecard? And to what extent, the national appraisal indicators are balanced?

RQ3: What positive outcomes have arisen from the implementation of the "National Appraisal" for key stakeholders? What challenges are confronted by tertiary public hospitals at the prefecture level in executing the "National Appraisal"? What can be improved in the "National Appraisal"?

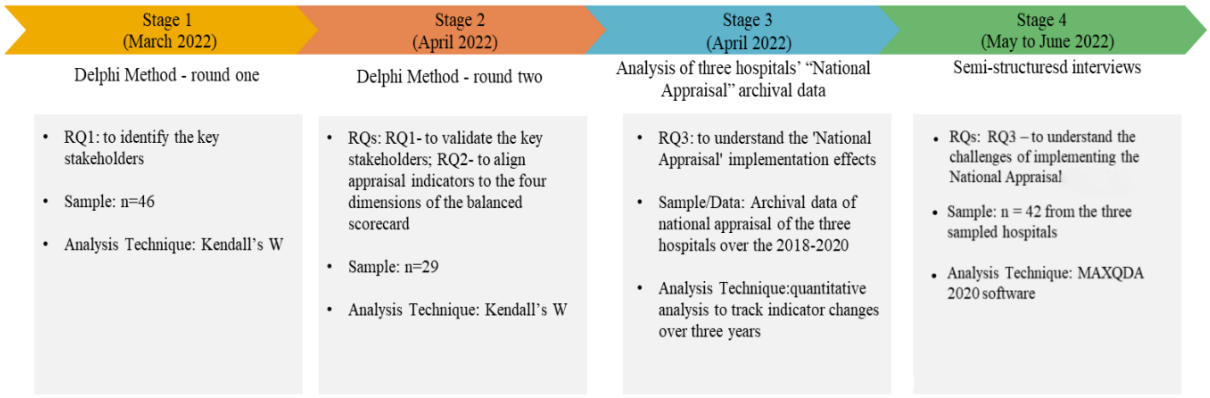
3. Research Methods

3.1 Research Methods and Procedure

This study employs a mixed methods research design that synergizes qualitative and quantitative data to enhance research validity.[34] Specifically, this study utilizes both methodological triangulation, incorporating the Delphi method, quantitative analysis of archival data, and semi-structured interviews with key stakeholders from selected hospitals, and data triangulation, drawing from expert consultations, three years of archival data, and interviews with 41 key stakeholders across three hospitals. This triangulation approach enhances research validity by gathering diverse methods and comprehensive data.[35]

As depicted in Figure 3, the overall procedure comprises four stages. The initial stage involved expert consultations using the Delphi method to address RQ1, identifying key stakeholders. Subsequently, a second round of the Delphi method was utilized in stage two to validate the key stakeholders identified in RQ1, and address RQ2, aligning 'National Appraisal' indicators with relevant stakeholders and the Balanced Scorecard (BSC) dimensions.

The third stage analyzed archival data (2018-2020) from three sample hospitals to explore changes over time in key appraisal indicators during the 'National Appraisal' implementation, answering RQ3. Finally, stage four encompassed interviews with key stakeholders from these hospitals, extending insights of RQ1 to RQ3 from the preceding stages. Specifically, it explored the positive effects of 'National Appraisal' implementation, and concurrent challenges. Subsequently, we elaborate on each stage of this process.



Note: RQ = Research Question

Fig. 3 Research Procedure

3.1.1 The Delphi Method

The Delphi method, recognized for consolidating expert opinions and achieving group consensus, serves as the foundation for this study's first and second research questions.[35] Engaging pertinent experts, this research systematically collects their insights through two rounds of questionnaires to identify key stakeholders and align the national appraisal indicators with these stakeholders and their corresponding BSC dimensions.

- Key Stakeholders and Delphi Method Round One

Through an extensive professional network, this study recruited 46 experts across ten provinces in China between February and March 2022. Questionnaires were distributed to 46 experts, and all 46 were successfully retrieved, yielding a 100% recovery rate. The selection criteria for these experts included: (1) a minimum of ten years of experience in clinical medicine, hospital management, or related domains; (2) holding senior professional titles; (3) leadership experience in managing departments or hospitals.

The selected experts, including 30 (65%) hospital vice presidents or presidents, 13 (18%) hospital department heads and 3 (7%) medical association leaders, had an average age of 52 years old and an average tenure of 18 years in their current organizations. Among them, 70% held a master's degree or higher.

The initial round of expert consultations was to identify primary key stakeholders—including health authorities, hospital executives, patients and their families, hospital staff, competitors, suppliers, industry associations, the public, and the media—based on their prominence in existing literature. A questionnaire was formulated to collect expert insights, employing a scoring system to gauge their agreement level regarding the importance of each stakeholder identified in the literature review. The scoring scale ranged from Strongly Disagree (1 point) to Strongly Agree (6 points), enabling experts to indicate their level of agreement.

- Key Stakeholders, National Appraisal Indicators and Delphi Method Round Two

For the second round of the Delphi method, a refined cohort of 32 hospital executives of vice presidents or presidents—was selected from the initial 46 participants involved in the first Delphi round. Out of these, 29 were successfully retrieved, resulting in a recovery rate of 90.6%.

During this phase, the experts further identified the key stakeholders and sequentially evaluated each of the 26 performance appraisal indicators' relevance to the key stakeholders identified in the first round. Furthermore, these experts categorized each of the 26 indicators into the appropriate dimension of the BSC framework.

- Delphi Method Data Analysis

Kendall's coefficient of concordance (Kendall's W)[37] was employed in both rounds of expert consultations to gauge the agreement among experts regarding key stakeholders. Generally, expert consistency can be considered good if the Kendall's W coefficient reaches 0.3 or above.[35,38] In this study, the calculation of Kendall's W employed SPSS Statistics 26.0's nonparametric test module.

In the second round of expert consultations, the pertinence of each key performance appraisal indicator to the interests of key stakeholders was assessed by considering the scores allocated to each indicator by said stakeholders. Additionally, each indicator garnered a specific count

of expert endorsements across the four dimensions of the BSC, into which experts categorized the indicators. The dimension with the high number of endorsements was designated as the suitable BSC dimension for that indicator. This approach allowed an analysis of the "National Appraisal" key indicators from the BSC perspective.

3.1.2 Archival Data of "National Appraisal" Implementation in Three Hospitals

Analyzing the "National Appraisal" archival data from three tertiary public hospitals spanning 2018-2020 provides a comprehensive view of performance patterns within a BSC framework. This analysis unveils the implementation effects of the "National Appraisal", unveiling underlying patterns and developmental trends.

With over 11,804 public hospitals in China by 2022, with significant disparities across different levels,[5] this study focuses on hospitals situated in prefecture-level cities, of which China had 334 by March 2022. Prefecture-level hospitals often grapple with a "sandwiched" position between technologically advanced provincial hospitals and state-supported county-level hospitals, facing huge development bottlenecks. The implementation of the “National Appraisal” had the most prominent impact on this group of hospitals, making them a good representative for studying the implementation effects of the “National Appraisal”. The profile of the sample hospitals is summarized in Table 1.

Table 1 Sample Hospitals Profile

Name	Founding Year	Number of Beds	Number of Departments	Number of Staff	Annual Emergency Visits (10,000 person-times)	Annual Discharges (10,000 person-times)	Annual Revenue (Billion Dollar)
CZ	1937	2200	57	3064	239.19	10.27	3.62
QY	1939	1780	52	3169	182.12	7.58	3.16
GZ	1939	2800	59	3500	130	13.6	2.75

The "National Appraisal" mandated unified performance evaluations nationwide. Tertiary public hospitals were required to analyze and assess the previous year's performance by September 2019, using the national indicator system. Consequently, inaugural national appraisal data pertained to 2018, with subsequent annual assessments founded on the preceding year's data. To analyze the impact and performance changes post-initiation, this study delved into the 2018-2020 national assessment data of three hospitals.

3.1.3 Semi-Structured Interviews with Sample Hospital Stakeholders

Face-to-face interviews with key stakeholders of the three sample hospitals were conducted during May and June 2022.

- Interviewee Selection

Semi-structured interviews were conducted with 41 interviewees drawn from the aforementioned stakeholders of the three sample hospitals, including hospital administrators, clinicians, patients and health authority (Table 2).

- Interview Questions

Based on the three research questions, an interview outline was formulated prior to the interviews and refined progressively. Universal questions across interviews encompassed: 1) How has the "National Appraisal" positively impacted stakeholders? 2) What challenges do tertiary public hospitals face in implementing the "National Appraisal" at the prefecture level? 3) From the stakeholder viewpoint, do current appraisal indicators harmonize stakeholder interests, and what indicators require enhancement? 4) What can be improved in the "National Appraisal"?

Table 2 Demographic characteristics of the interviewees and interview data

Interviewee	Stakeholder	Specialty/ Department	Education background	Years of service in the current position	Date of interview	Length of interview
CZ Hospital	Hospital administrator (7)	medicine (5) pharmacy (1) IT technician (1)	doctor (4) master (1) bachelor (2)	Average 3.3 years	June 1, 2022	Average 47min 3s
	Clinicians (4)	clinical medicine (4)	master (1)	Average 8 years	June 1, 2022 June 2, 2022	Average 35min 46s
	Representative of a health authority (1)	Deputy Director of the Health Commission (1)	doctor (1)	Average 5 years	June 2, 2022	Average 23min 30s
	Patient (4)	Orthopedics (1) Gastroenterology (1) Gastrointestinal (2)			June 2, 2022	Average 7min 7s
GZ Hospital	Hospital administrator (5)	medicine (3) IT technician (1) nursing (1)	master (3) bachelor (2)	Average 6.6 years	May 10, 2022 May 11, 2022	Average 42min 51s
	Clinicians (2)	clinical medicine (1) sports medicine (1)	bachelor (2)	Average 28 years	May 10, 2022	Average 38min 45s
	Representative of a health authority (1)	Chief of the Institutional Reform Section of the Health Commission (1)	doctor (1)	Average 3 years	May 11, 2022	Average 53min 33s
	Patient (5)	Stomatology (1) Neurology (1) Endocrinology (1) General practice (1) Spine surgery (1)			May 10, 2022 May 11, 2022	Average 4min 1s
QY Hospital	Hospital administrator (4)	medicine (4)	doctor (1) master (1) bachelor (2)	Average 5.75 years	May 20, 2022 May 21, 2022	Average 42min 23s
	Clinicians (4)	clinical medicine (4)	master (4)	Average 7.25 years	May 20, 2022 May 21, 2022	Average 30min 55s
	Patient (4)	Neurology (2) Cerebrovascular (2)	unknown	n/a	May 21, 2022	Average 3min 53s

- Interview Data Analysis

MAXQDA 2020 software was employed to analyze interview data. Employing open coding, keywords were extracted from data records, leading to Level 1 codes. Conceptually similar codes were grouped to form Level 2 codes. Open coding condensed data hierarchically, maintaining original meaning. This procedure yielded 554 Level 1 codes. Level 2 coding was subsequently executed based on Level 1 codes, and Level 2 codes were further synthesized into five major thematic clusters in the results section, reflecting the dominant patterns and challenges identified in the interviews.

3.2 Validity and Reliability

This study prioritized maintaining validity and reliability across the entire research process, encompassing research design, data collection, and analysis, through the following means.

Validity: First, extensive literature review enabled the researcher to formulate a robust framework and research questions by amalgamating established BSC and stakeholder theories (theory triangulation). The study then employed a mixed methodology, integrating quantitative and qualitative approaches (methodological triangulation). Incorporating multiple data sources (data triangulation) simultaneously or sequentially enhances study validity and rigor.[39] The quantitative aspect entailed two-round expert surveys, iteratively validating the outcomes. Concurrently, interviews engaged diverse stakeholders, corroborating viewpoints across roles and perspectives. The study harmonized archival data on hospital performance appraisal over three years and interview results, reinforcing one another (data triangulation).

Reliability: Meticulous steps were taken from the outset to ensure the reliability of this study. A robust research protocol was meticulously crafted, drawing on established principles outlined by researchers such as Yin[40] (2009) and Gibbert[41] et al. (2008). Transparency was prioritized, with every aspect of the research process meticulously documented and clarified. This commitment to transparency is evident in the detailed presentation of procedures.

4 Results

4.1 RQ1: The key stakeholders and their relevance in “National Appraisal" indicators

4.1.1 Delphi Method Round One

Expert consultation results are reported in Table 3. The mean scores assigned to the 9 key stakeholders identified from the literature varied between 5.43 and 3.87 points. Experts concurred that the primary five key stakeholders were hospital executives, health care workers (HCWs), patients, health authorities, and the public.

Table 3 Stakeholder scores assigned by experts (round one)

	Position			Professional Background			
Expert	Hospital vice presidents and presidents, mean (SD) (N=30)	Hospital department heads, mean (SD) (N=13)	*p* value	Clinical managers, mean (SD) (N=27)	Managers, mean (SD) (N=19)	*p* value	Total mean (SD) (N=46)
Hospital executives	5.47 (0.82)	5.31 (0.63)	0.533	5.33 (0.83)	5.58 (0.61)	0.279	5.43 (0.75)
HCWs	5.20 (1.06)	5.15 (0.80)	0.902	5.26 (0.86)	4.89 (1.45)	0.291	5.11 (1.14)
Patients	5.07 (0.94)	4.46 (1.13)	0.123	5.00 (0.96)	4.63 (1.42)	0.299	4.85 (1.17)
Health authorities	4.80 (1.16)	5.00 (1.00)	0.617	4.93 (1.17)	4.58 (1.26)	0.344	4.78 (1.21)
The public	4.60 (0.81)	4.46 (1.05)	0.647	4.63 (0.84)	4.47 (0.96)	0.562	4.57 (0.89)
Competitors	4.47 (0.97)	4.77 (1.09)	0.402	4.41 (0.93)	4.53 (1.39)	0.729	4.46 (1.13)
Industry associations	4.23 (1.41)	3.77 (1.74)	0.354	4.11 (1.42)	4.21 (1.62)	0.827	4.15 (1.49)
The media	3.87 (1.17)	4.15 (1.21)	0.486	4.15 (1.13)	3.53 (1.31)	0.092	3.89 (1.23)
Suppliers	4.00 (1.17)	3.69 (1.44)	0.482	4.04 (0.98)	3.63 (1.64)	0.300	3.87 (1.29)

** The p-value is the result of t-test.*

The consistency of the appraisal results was further affirmed by the Kendall's W analysis. The overall Kendall's W for all nine indicators across the 46 questionnaires stood at 0.273 ($p < .001$). After excluding "the media" which ranked last, the Kendall's W for the top eight indicators was 0.263 ($p < .001$). This process was repeated, resulting in a Kendall's W of 0.222 ($p < .001$) for the first five. A noticeable drop occurred when considering the first four indicators, with a Kendall's W of 0.116 ($p < .001$). Considering the number of key indicators and the Kendall's W values, selecting the initial five indicators as key stakeholders appears most reasonable: hospital executives, HCWs, patients and their families, health authorities, and the public.

To further validate the results, similar analyses were conducted for different subgroups (clinical managers vs managers without clinical roles) and the results were reported in Table 3. There existed no statistically significant disparity in the scores attributed to each key stakeholder among managers of diverse levels and with or without clinical roles, indicating a substantial alignment of opinions. The consistency of scores for the subgroups supports the selection of the top five indicators as key stakeholders.

4.1.2 Delphi method round two

In round two, respondents were tasked with not only identifying the key stakeholders, but also categorizing the "National Appraisal" indicators into the four dimensions of the Balanced Scorecard (BSC) and evaluating the degree of relevance between these key indicators of the "National Appraisal" and the key stakeholders of public hospitals. For the first task, the unanimous consensus among experts was that the five key stakeholders remained the same as the findings in the first round of consultation. The mean scores were as follows: patients at 3.55, hospital executives at 3.55, HCWs at 3.34, health authorities at 2.55, and the public at 2.00. The Kendall's W coefficient stood at 0.317 ($p < .001$).

For the second task, the findings (Table 5) revealed that among the 26 key indicators, 5 indicators were classified under the financial dimension (consensus between 86.20%-100%),

6 under the customer (patients) (consensus between 55.20%-96.60%), 10 under the internal process (consensus between 79.30%-96.60%), and 5 under the learning and growth (consensus between 56.60%-96.60%). Notable divergences emerged in experts' opinions regarding certain indicators within the customer dimension. Specifically, less than 60% of experts concurred that the cost-control indicators belonged to the customer dimension, with the remaining 40% advocating for their classification within the financial dimension. These discrepancies hint the "conflict of interests" between patients and public hospitals.

In summary, the experts' opinions on the indicators within the financial dimension were found to exhibit the highest level of consistency in terms of endorsement percentage, while opinions on indicators within the customer dimension showed the least consistency. The endorsement percentages across the four dimensions averaged 93% for financial, 70% for customer, 90% for internal process, and 81% for learning and growth.

The analysis also extended to determining the most relevant key stakeholders for each performance appraisal indicator. By assessing the scores allocated by experts to indicate the relevance of each appraisal indicator to key stakeholders, a calculation was executed to ascertain the most pertinent stakeholders for each performance appraisal indicator (Table 4). Among the 26 indicators, 14 were deemed most relevant to hospital executives, 3 to HCWs, 9 to patients, but none to health authorities and the public.

Table 4 Performance Appraisal Indicators and Their Relevance to Stakeholders

Indicators	Hospital executives		HCWs		Patients		Health authorities		The public		Most relevant stakeholders
	Total score	Mean score	Total score	Mean score	Total score	Mean score	Total score	Mean score	Total score	Mean score	
4. Proportion of discharged patients who underwent a surgery	34	2.62	32	2.67	5	1.67	2	2	0	0	Hospital manager
6. Proportion of discharged patients who underwent a Level 4 surgery	37	2.53	31	2.82	5	1.67	0	0	0	0	Hospital manager
10. Single disease quality control	27	2.33	24	2.67	20	2.86	2	2	0	0	Hospital manager
13. Number of clinical test parameters accredited by national external quality assessment	45	2.5	6	2.33	15	2.5	2	2	3	3	Hospital manager
17. Defined daily dose (DDDs)	33	2.83	17	2.57	18	2.25	0	0	5	2.5	Hospital manager
24. Function application level of the electronic medical record	48	2.72	15	2.5	6	3	6	2	0	0	Hospital manager
31. Proportion of medical service revenue to medical revenue	58	2.68	16	2.67	0	0	2	2	0	0	Hospital manager
33. Proportion of personnel expenditure to business expenditure	42	2.8	31	2.82	3	3	6	3	0	0	Hospital manager
34. Energy consumption expenditure per ten thousand yuan of income	64	2.67	3	3	3	3	2	2	3	1.5	Hospital manager
35. Balance of income and expenditure	63	2.63	4	2	3	3	4	2	0	0	Hospital manager
36. Asset-liability ratio	58	2.68	0	0	3	3	13	2.33	0	0	Hospital manager
45. The proportions of anesthesiologists, pediatricians, intensive care specialists, pathologists and TCM practitioners	27	2.33	26	2.17	4	2	6	2	0	0	Hospital manager
46. Doctor-nurse ratio	40	2.56	24	2.4	4	2	2	2	0	0	Hospital manager
50. Fund for scientific research projects per 100 health technicians	42	2.63	21	2.63	0	0	3	1.33	2	1	Hospital manager

Indicators	Hospital executives		HCWs		Patients		Health authorities		The public		Most relevant stakeholders
5. Proportion of discharged patients who underwent a minimally invasive surgery	28	2.55	30	2.58	13	2.17	0	0	0	0	HCWs
48. The pass rate of hospital residents taking the Qualification Examination for Doctors for the first time	21	2.1	37	2.47	0	0	4	2	1	1	HCWs
55. Satisfaction of health care workers	11	2.75	69	2.88	0	0	3	3	0	0	HCWs
8. Incidence of postoperative complications	16	2.83	26	2.6	37	2.92	0	0	0	0	Patients
9. Surgical site infection rate in sterile procedures	16	2.83	26	2.45	33	2.83	0	0	0	0	Patients
14. Mortality rate of cases in the low-risk group	17	2.57	29	2.64	32	2.91	0	0	0	0	Patients
38. Rate of increase in outpatients' medical expense per visit	17	2.57	6	2	48	2.67	2	2	0	0	Patients
39. Rate of increase in outpatients' medication expense per visit	13	2.33	9	2.25	48	2.67	2	2	0	0	Patients
40. Rate of increase in average medical expense per inpatient	14	2.8	8	2.67	46	2.61	6	2.33	0	0	Patients
41. Rate of increase in average medication expense per inpatient	14	2.8	10	2.5	45	2.65	6	2.33	0	0	Patients
53. Outpatient satisfaction	3	3	6	2.33	62	2.86	3	3	6	3	Patients
54. Inpatient satisfaction	3	3	17	2.57	52	2.89	3	3	6	3	Patients

The above results yield insights into addressing RQ1. This study identifies five primary stakeholders of tertiary public hospitals: hospital executives, HCWs, patients, health authorities, and the public. Among them, only three stakeholders – hospital executives, HCWs, patients – are intricately linked with the performance appraisal indicators of tertiary public hospitals with the absence of health authorities and the public.

4.2 RQ2: Categorization of the key indicators of the "National Appraisal" into the four dimensions of the balanced scorecard

Several findings contribute to addressing RQ2. From the perspective of BSC dimensions, each key indicator of the "National Appraisal" falls into specific categories (Table 5). The financial dimension encompasses 5 indicators, all of which strongly resonate with the concerns of hospital executives. The customer dimension comprises 6 indicators, all of which are primarily associated with patients' interests. The internal process dimension encompasses 10 indicators, of which six are most pertinent to hospital executives, one to HCWs, and three to patients. Finally, the learning and growth dimension is gauged by five indicators, with three being most relevant to hospital executives and two to HCWs.

From stakeholder perspective, hospital executives encompass all five indicators within the financial dimension, six in the internal process, and three in the learning and growth. For HCWs, one indicator from the internal process and two from the learning and growth stand out. Patients find all six customer indicators and three internal process most relevant. Notably, the appraisal indicators primarily cater to the concerns of hospital executives and patients, with fewer indicators aligning with the interests of HCWs, and none correlating with the roles of health authorities and the interests of the public. The results are illustrated in Figure 4.

Table 5 Distribution of 26 indicators among key stakeholders and the BSC dimensions

Dimension	Hospital executives N (%)	Health care workers N (%)	Patients N (%)	Health authorities N (%)	The Public (%)	Total
Financial	5 [31, 33, 34, 35, 36] (100%)	0	0	0	0	5 (19%)
Customer	0	0	6 [38, 39, 40, 41, 53, 54] (100)	0	0	6 (24%)
Internal process	6 [4, 6, 10, 13, 17, 24] (60%)	1 [5] (10%)	3 [8, 9, 14] (30%)	0	0	10 (38%)
Learning and growth	3 [45, 46, 50] (60%)	2 [48, 55] (40%)	0	0	0	5 (19%)
Total	14 (54%)	3 (11%)	9(35%)	0	0	26 (100)

Note: N indicates the number of indicators in the respective dimension; the numbers within the square brackets ([]) indicate the specific appraisal indicators of the “National Appraisal” system.

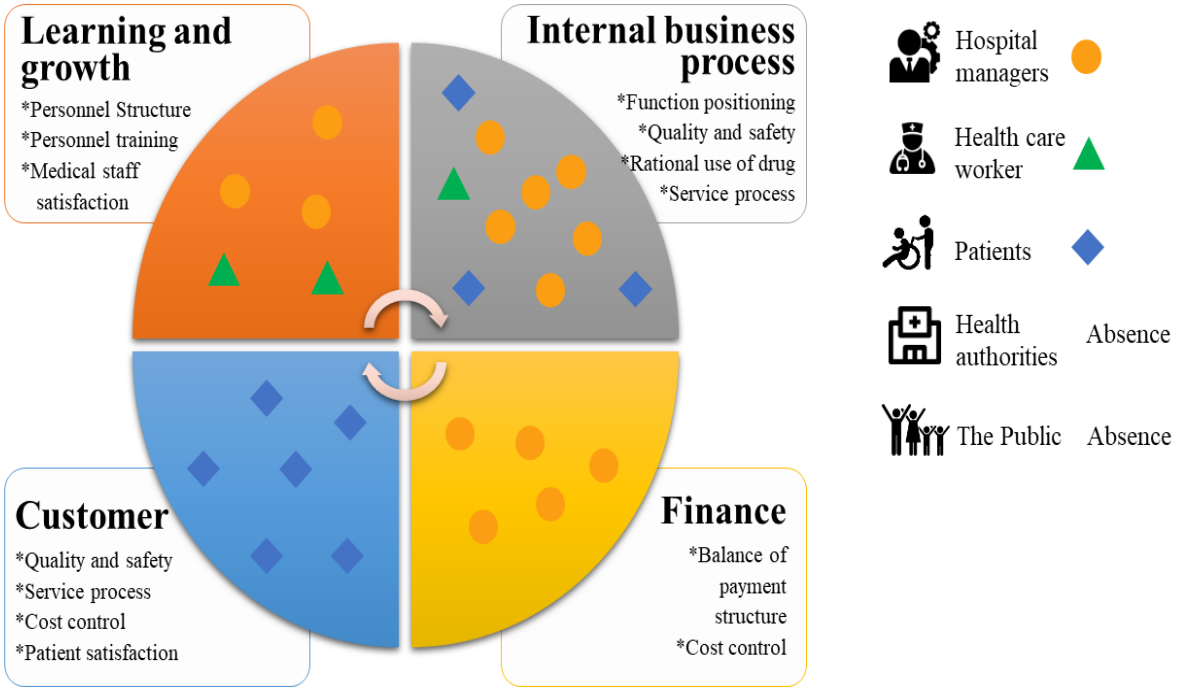


Fig. 4 The National Appraisal model from hospital stakeholders and the BSC perspective

4.3 RQ3: Implementation of the "National Appraisal": benefits and challenges

4.3.1. Findings from archival data

Derived from the implementation data of the 'National Appraisal' in three sample hospitals, the results are manifested in the four dimensions of the BSC obtained from the Delphi method round two. The observed trends between 2018 and 2020 in these indicators were summarized in Figure 5. Building on the BSC framework, key performance indicators (KPIs) were selectively prioritized based on national appraisal relevance, operational significance, and stakeholder interests, including hospital executives, healthcare workers (HCWs), and patients.

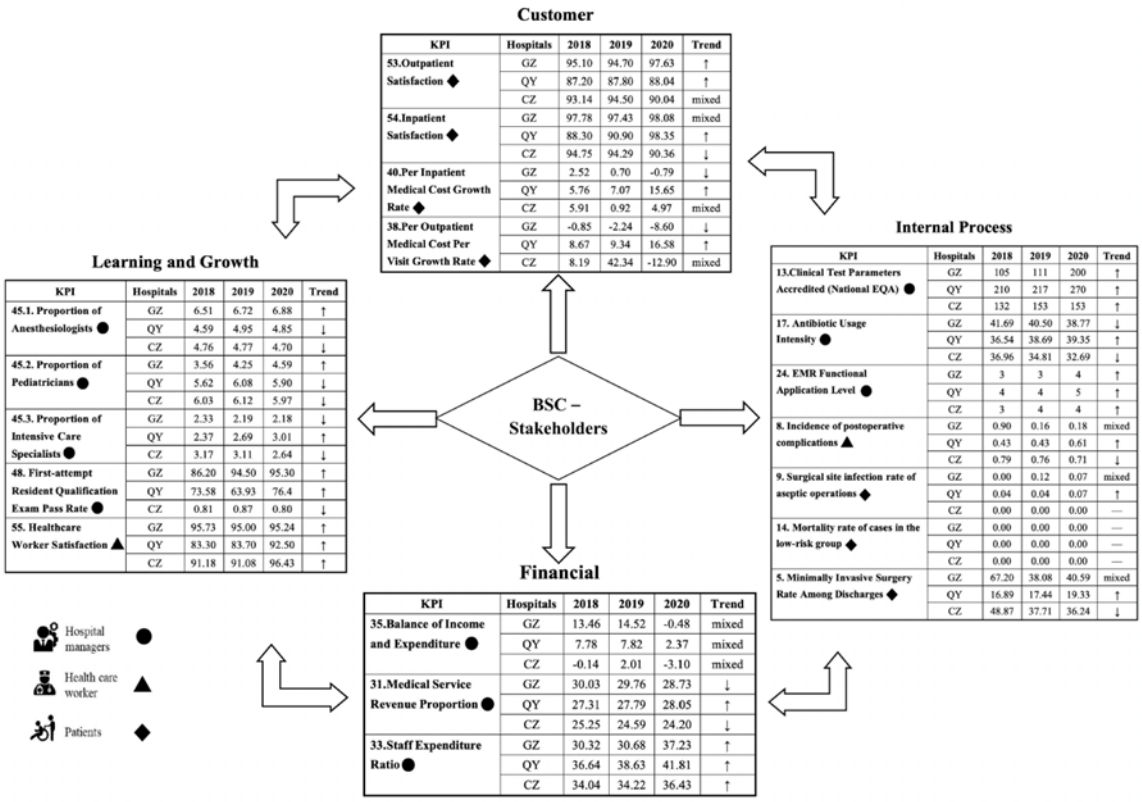


Fig. 5 - Implementation outcome of the "National Appraisal" in BSC framework

The **financial dimension** primarily serves as an indicator of the hospital's operational efficiency, revealing insights into the revenue and expenditure structure, cost control, debt situation, and financial sustainability. The performance data from the three sample hospitals revealed the following trends. A trend of an increased proportion of the compensation and benefits of HCWs to the overall expenditure of hospitals was observed (indicator 33), indicating a notable achievement in hospital's commitment to HCWs' well-being.

However, the trend for indicator 31 either remained stagnant or declined, indicating that hospitals heavily relied on income from drugs and consumables, placing less value on the medical services provided by HCWs in the hospitals' overall medical revenue. Additionally, the mixed findings of indicators 35 on balance of income and expenditure, reflecting overall financial viability across the three hospitals depicted a challenging financial situation.

The **customer/patient dimension** covers both cost accessibility and service satisfaction, reflecting key interests of patients. There is no evident downward trend in controlling average outpatient costs. Some hospitals continued to report a rise in average costs. This increase was influenced by rising diagnosis and treatment costs, which could be rationalized by hospitals' financial dependence on income from drugs and consumables mentioned above. Despite the rising cost, the patient satisfaction index (53 and 54) of the sample hospital has steadily increased due to benefits brought by the improved health care quality, safety and efficiency as discussed next.

The **internal process dimension** included 10 key indicators through Delphi expert consultations in this study.

Indicators (e.g. 5) measuring the functional positioning of hospitals exhibit a gradual upward trend, indicating active adoption of new medical technologies such as minimally invasive surgery. This reflects a commitment to enhancing the medical capability.

Indicators 8, 9, and 14 are closely linked to health care quality and safety, and all three indicators show steady progress, signaling ongoing improvement in these critical areas. Indicator 13 reflects the hospital's participation in national external quality evaluation projects, with a noticeable upward trend indicating advancements in standardization of examinations.

Indicator 17 assesses the intensity of antibacterial drug use in hospitals, with sample hospitals surpassing the national standard and two out of three sampled hospitals demonstrating a downward trend, signifying continuous improvement in the level of clinically appropriate drug use.

Indicator 24 evaluates the hospital's level of informatization construction and application. Sample hospitals have exceeded the national standard, and are consistently optimizing their informatization construction.

The **learning and growth dimension** serve as a reflection of the hospital's personnel structure, talent development, discipline construction, and the satisfaction of HCWs.

The results from sample hospitals reveal a significant increase in HCWs' satisfaction (indicator 55) which align with the upward trend in the compensations and benefits of HCWs (indicator 33 in financial dimension). There was a steady increase in the pass rate of hospital residents taking the qualification examination for doctors (indicator 48), indicating heightened efforts in talent development. The above indicators suggest the overall improvements in healthcare services, operational efficiency, and capability. These positive outcomes were also noted in the interviews as we discuss next.

4.3.2 Key themes from interviews

The interviews reveal the complex interaction between centrally mandated appraisal requirements and the actual conditions faced by tertiary public hospitals. The key themes are summarized below.

Theme 1: Alignment between hospital strategic goals and national appraisal indicators

Interviews revealed that the "National Appraisal" played a key role in clarifying hospital development goals, defining the functional positioning of tertiary public hospitals, and guiding reforms. However, many hospital managers reported challenges in aligning standardized indicators with local hospital strategies. The appraisal's "one-size-fits-all" design often conflicted with hospitals' specific service priorities and patient needs. As one hospital director noted, "*It's a one-size-fits-all assessment—we have to adjust our development goals every year to meet changing criteria, even if it means compromising the actual needs of patients.*"

Theme 2: Main benefits brought by the implementation of the "National Appraisal"

The "National Appraisal" has contributed to improved healthcare service quality, enhanced operational efficiency, strengthened discipline development, and better management

performance in hospitals. For hospital managers and health authorities, it provided clearer direction and stronger motivation to drive reforms that ultimately benefit patients. For doctors and nurses, it enhanced professional skills by offering more training opportunities and promoting standardized diagnosis and treatment. Moreover, there was a noted surge in motivation for scientific research, signaling a greater emphasis on its development. The "National Appraisal" helped upgrade medical care, increased their compensations and development opportunity, and enhance HCWs satisfaction. As a result, patients benefit from the implementation of the "National Appraisal" which were evidenced by the 13 patient interviewees. These patients acknowledged easy access to quality healthcare and advanced medical equipment and technology as well as more affordable expenses.

Theme 3: Challenges in implementation of the "National Appraisal"

The challenges and dilemmas associated with implementing performance appraisal can be broadly categorized into two groups. The first category highlights the difficulty of implementing the national appraisal due to lack of resources in these hospitals, such as funding, talents, and policy support. In prefecture-level cities, positioned as the "middle layer of the sandwich," challenges like "*a lack of professional talents*", and "*the need for more financial investment*" emerged as top concerns based on frequency mentioned by the interviewees. In particular, the "National Appraisal" has intensified financial pressure on hospitals to reduce patient costs, without accompanying government subsidies. As one hospital vice president complained, "*We are constantly told to reduce drug prices and service fees, but this is not sustainable without corresponding government financial support*." Regarding lack of talents, shortages in anesthesiologists, pediatricians, and nurses were prevalent.

The second challenge category concerns the discrepancy between uniform appraisal indicators and the realities within the local healthcare community. Respondents consistently expressed that employing a uniform set of indicators for performance appraisal across diverse hospitals

lacked practicality. During appraisal implementation, obstacles and conflicts arose between the "National Appraisal" indicators and the interests of patients, hospitals, and HCWs. Some indicators do not align well with actual hospital operations, adversely affecting doctors' patient care practices and consequently impacting patient well-being. For example, cost-control measures may strain hospital operations or contradict principles of clinical diagnosis and treatment. These conflicts can impede patients' access to necessary healthcare services. HCWs express significant concern regarding the discrepancy between indicators and patient needs, as well as the undue emphasis on scientific research assessment. Moreover, they note that the appraisal has, to some extent, disrupted standard treatment procedures.

5 Discussion

Through the lenses of stakeholder theory and the Balanced Scorecard (BSC) framework, this study examined how the “National Appraisal” was related to the interests of key stakeholders of public hospitals in China and the benefits and challenges of implementing the “National Appraisal”.

Our findings indicate that the "National Appraisal" adopts a "balanced scorecard" approach across four dimensions, albeit with an imbalance in the involvement of key stakeholders from public hospitals in China. The indicators encompass healthcare quality and safety, access, efficiency, and patient-centeredness, aligning with widely recognized performance indicators in developed countries.[20,23] Furthermore, we extend our analysis beyond the "National Appraisal" framework per se to examine its implementation effects. Our results showed that the “National Appraisal” resulted in significant positive outcomes.

This paper makes three key contributions. First, it addresses the limited literature on hospital performance appraisal by empirically examining the long-term implementation of the BSC in public hospitals, offering valuable insights into its effectiveness in healthcare settings over time.

In doing so, it not only enhances understanding of performance appraisal effectiveness in Chinese public hospitals, but also provides an empirically grounded framework for other countries pursuing balanced, collaborative, multi-stakeholder hospital performance reforms. Specifically, the proposed stakeholder-balanced scorecard integrative approach offers a compelling and practical model for performance measurement in complex, multi-actor systems like public hospitals and serves as a reference for hospital performance appraisal in other countries and institutional contexts.

Secondly, to the best of the authors' knowledge, this is the first study of China's "National Appraisal" system using stakeholder theory and the BSC framework, shedding light on its intricacies and offering a replicable, adaptable model for hospital performance management reform in other national contexts by demonstrating how to simultaneously identify and critically balance the interests of regulators, hospital managers, HCWs, patients, and the public, optimize indicator allocation under differing resource and institutional constraints, and embed dynamic feedback and local-adaptation mechanisms into the appraisal process.

Thirdly, the research employs a mixed-method approach, offering a comprehensive empirical study of the BSC appraisal approach and its implementation effects. By combining quantitative and qualitative methods, the study provides a holistic understanding of how the BSC-based appraisal influences hospital operations and performance, offering rich and informative insights for hospital executives and policymakers alike.

5.1 The "National Appraisal": a balanced scorecard-oriented

From the perspective of performance management in public hospitals, the "National Appraisal" has achieved significant and positive outcomes through the adoption and implementation of a balanced scorecard-oriented appraisal system in China's public hospitals. This appraisal system has succeeded by emphasizing non-financial indicators (81%) over financial indicators (19%),

thereby guiding public hospitals to prioritize "public interests" over "financial interests". The majority of non-financial indicators (38%) focus on internal processes, driving improvements in healthcare quality, safety, and efficiency. Additionally, 24% of the indicators are directed towards patients (customer dimension), successfully prioritizing patient interests in the performance appraisal process. While health care workers (HCWs)' interests are somewhat marginalized, comprising only 19% of indicators, they do benefit from enhanced compensation and development opportunities. Statistically, the indicators are distributed unbalanced, but many indicators are interrelated and mutually influential.[42] For example, both HCWs and patients benefit indirectly from the improved internal processes due to its widely covered indicators.

The review of performance management approaches in China's public hospitals indicates that the adoption of a "balanced approach" represents a natural evolution following various methodologies over the past seven decades. The introduction of the balanced "National Appraisal" marks a significant milestone in China's healthcare reform and development. Its comprehensive and balanced approach provides valuable insights and reference points for performance management in healthcare systems worldwide, contributing to the advancement of global healthcare standards.

5.2 Key stakeholders in Chinese public hospitals and their conflict of interests

Our study's most intriguing findings revolve around the identification of five key stakeholders in China’s public hospitals and the imbalanced roles and interests in the “National Appraisal”. Specifically, patients, hospital executives, and HCWs, health authorities and the public are key stakeholders. But the “National Appraisal” indicators only encompass these first three groups, neglecting the roles of health authorities and the interests of the public. This omission is problematic because health authorities oversee healthcare policies, regulations, and resource

allocation, particularly financial resources, while the public represents the broader community served and affected by hospital services. This findings highlight the need for stakeholder engagement in performance appraisal implementation in public hospitals.[3]

In the context of public hospitals in China, addressing and balancing the diverse needs of the stakeholders within the “National Appraisal” poses a significant challenge. In our study, conflicts among key stakeholders are most evident in financial matters involving patients, hospitals, and governments. In the second round Delphi expert consultation, discrepancies arose regarding the classification of indicators related to patient expense increase rates. These differences underscore the conflict between patient and hospital financial interests. Controlling patient expense increases is in conflict of hospitals’ interests, especially amid insufficient government funding for public hospitals. Similar findings were reflected in the archival data from three hospitals from 2018 to 2020 that exposed persistent financial struggles, including debt, losses, and operational strain. These hospitals heavily depended on revenue from drug and consumable sales. Rising diagnosis and treatment expenses underscore their reliance on such revenue sources. Local governments offered minimal support for the "National Appraisal" implementation, leading to an imbalance in power and responsibilities.

Lastly, at a micro level, conflicting interests may arise between policy-oriented indicators and the practical needs of patients. For example, the cost-control measures of the "National Appraisal" could potentially strain hospital operations and conflict with clinical diagnosis and treatment principles, thereby hindering patients' access to essential healthcare services and depriving HCWs professional autonomy and personal interest. In essence, when these measures aim to safeguard patients' interests, they inadvertently compromise patients' ultimate well-being.

Stakeholder conflicts in the “National Appraisal” system stem from structural imbalances in power and influence. In China’s healthcare governance, the National Health Commission (NHC) set the framework, performance targets, and indicators for tertiary public hospitals. Although they oversee and evaluate hospital performance post hoc through audits, their own responsibilities—such as providing financial support or timely feedback—are not assessed within the appraisal system. In effect, health authorities hold significant power over hospitals but bear little corresponding responsibility. Patients and the public, despite being the ultimate beneficiaries, are excluded from direct participation in indicator development and performance appraisal decisions. Their perspectives are reflected only indirectly through survey-based metrics like patient satisfaction, which often fail to capture their actual needs.

Hospitals, particularly managers and healthcare workers (HCWs), are thus pressured to “deliver more with less,” straining already limited resources. While managers bear the brunt of operational responsibility under the appraisal, HCWs are the most vulnerable, with their interests often marginalized or sacrificed. This asymmetry reflects a key principle of stakeholder theory: formal recognition does not guarantee influence over strategic processes. To maximize the benefits and ensure the sustainability of the “National Appraisal,” it is essential to address key challenges brought by the power dynamics among stakeholders: the imbalance of power and responsibility among health authorities, the mismatch between hospital duties and available resources, the exclusion of public participations, and limited external oversight.

5.3 Outcomes of “National Appraisal” Implementation

By adopting a balanced approach, China's "National Appraisal" in public hospitals achieves multiple benefits: clarifying hospital strategic goals and roles, shaping hospital priorities, and enhancing operational and healthcare performance as well as the satisfaction of patients and HCWs.

The study suggests that the implementation of the "National Appraisal" consistently succeeds in clarifying hospital strategic direction and objectives, in line with findings on BSC implementation in non-profit organization.[12] By defining key indicators across four vital areas (healthcare quality, operational efficiency, sustainable development, and satisfaction), health authorities have successfully used the "National Appraisal" to align the Chinese government's vision and strategic goal of prioritizing the public roles of public hospitals that had previously struggled with "for-profit" practices due to inadequate funding from governments. At the hospital level, using the "National Appraisal" as a framework and supported by hospital information systems, hospital executives are empowered to make informed decisions that encompass all aspects of performance. This approach has helped to streamline the complex process of running hospitals by focusing on healthcare quality, safety, and efficiency, providing more affordable patient care.[31] Overall, the "National Appraisal" system has been instrumental in shaping hospital priorities and clarifying their roles.

At the operational level, both archival data and stakeholder interviews highlight the positive impact of the "National Appraisal" system on health care services and hospital operations over time. Implementation led to the adoption of new medical technologies and IT integration, boosting operational efficiency and patient care in tertiary public hospitals. Subsequently, positive impacts contribute to the overall benefits of patients and lead to the improved patient satisfaction. It can be seen that the “National Appraisal” prioritizing patients’ interests and hospitals are guided to prioritize patient-centered health care delivery. Patients interviewed also report tangible benefits, including easy access to quality healthcare, advanced medical equipment, and controlled out-of-pocket expenses. This reinforces the notion in literature that the BSC result in tangible improvements in patient satisfaction.[3]

Lastly, the "National Appraisal" implementation has increased heath care workers’ compensation and development opportunities, and job satisfaction. Compared to the impact on

patients, the effect on HCWs is less pronounced which echoes the findings regarding BSC implementation in the literature.[3]

5.4 Recommendations

5.4.1 Stakeholder Engagement and Appraisal Indicators Refinement

As the architects of the "National Appraisal," health authorities wield the power to establish appraisal indicators and oversee medical practices and their accountability is essential for effective oversight. For example, the appraisal implementation process necessitates coordination with medical insurance policies.

To address these challenges, we propose extending the “National Appraisal” for Chinese public hospitals to include indicators evaluating the regional health authorities that finance and administrate the local public hospitals. This evaluation would encompass policy support, resource allocation efficiency, timely assessment feedback, collaboration with hospitals and other stakeholders to share best practices, and alignment with hospital goals set by NHC.

Given the limited financial resources, with China's only 5.59% of GDP used for health expenditure,[7] local governments, typically responsible for funding public hospitals, often provide minimal assistance. This imbalance in responsibilities underscores the need for active participation of local governments and health authorities in performance assessment, ensuring the necessary resources for effective implementation of the appraisal.

Furthermore, hospital executives require timely feedback on appraisal results to guide improvement efforts. Currently, health authorities have not fully utilized the collected data to steer hospital development. Therefore, there is a need to establish a feedback and support mechanism for hospitals by the health authorities to inform decision-making and drive improvements in hospital performance.[43]

Regarding the absence of the public in the appraisal process, we propose two methods to engage them. First, we advocate involving the public in hospital governance, particularly through a performance appraisal council, as we will discuss in more detail shortly. Second, we recommend adding indicators to assess public hospitals' commitment to community engagement within the appraisal framework. These indicators may encompass community engagement activities, health price transparency, and the inclusion of community representatives in hospital governance.

In addition, a prevalent theme emerging from the interviews underscores the necessity of adopting a diversified appraisal approach, considering the diverse circumstances of hospitals across different levels and geographical locations.[2,19] We propose tailoring performance appraisals by considering hospital's local development, population size, administrative level, and other relevant contextual factors. Moreover, there is a need to incorporate clinical considerations when formulating indicators.[3]

Finally, it's essential to ensure a more inclusive and comprehensive integration of HCWs' perspectives in both the design and execution of performance appraisals to ensure their voices are heard. Specifically, considering the issues identified in this study and the significant emphasis placed on HCWs by the WHO in assessing hospital performance,[16] we argue that additional factors such as mandating an adequate number of physicians to alleviate workload, and equitable compensation for permanent and non-permanent HCWs must be incorporated into the appraisal system to safeguard their interests. Moreover, it is equally crucial to involve HCWs in the design and implementation processes of performance appraisals.

5.4.2 Improving Governance

In assessing the performance of public hospitals, researchers[13] emphasize the critical role of governance and structure. These elements directly influence the availability of resources necessary to facilitate processes, consequently impacting healthcare outcomes.

As part of the stakeholder engagement strategy by “strategically” engaging community,[44] we advocate for the establishment of a "Performance Appraisal Council" (PAC) to oversee the hospital leadership and performance appraisal office in implementing the appraisal. The PAC offers numerous benefits. Firstly, it serves as a platform for stakeholder engagement by including representatives from the community, health professionals, and academia. This facilitates collaboration and integration among stakeholders, benchmarking at home and abroad, fostering a collective effort to achieve shared goals.

Additionally, the PAC helps strengthen hospital governance, particularly in performance management. Among the 26 indicators, 14 are most relevant to hospital executives, constituting the largest proportion (54%). This suggests that hospital executives bear significant responsibility[45] and are heavily involved in performance appraisal. Concentrating performance management solely on top leadership may hinder transparency, and information flow. Therefore, establishing a council may alleviate the burden on hospital executives, allowing them to focus on other strategic agendas while improving hospital governance. In short, the PAC can provide valuable oversight and support, ensuring effective performance management processes while promoting transparency and accountability throughout the organization.

5.5 Limitations and Future Research

This study has several limitations. Firstly, despite employing various methods and data sources for robust research, findings may not fully represent the complexity of public hospitals across

different regions in China, given the country's vast size. A follow-up study could expand the scope and depth of research to provide valuable insights.

Secondly, although archival data spanning three years was used to explore the impacts of "National Appraisal" implementation, it does not allow for establishing cause-and-effect relationships, as seen in other empirical research on BSC implementation.[12] While this approach offers insights into trends over time, it does not capture causal intricacies. Additionally, while the selected prefecture-level hospitals offer valuable insights into the real-world implementation of the appraisal, caution is needed when generalizing these findings to all tertiary public hospitals in China, because substantial differences in resource availability, and operational contexts exist across different levels of tertiary hospitals.

Lastly, key stakeholders interviewed were mainly hospital executives, with limited representation from health authorities. Thus, the results may not fully reflect the perspectives and interests of health authorities, compromising the generalizability of findings. Future research may compare perspectives from different stakeholders.

References

1. Zheng Y, Wang W, Liu W, Mingers J. A performance management framework for the public sector: The balanced stakeholder model. *Journal of the Operational Research Society*. 2018;70(4):568-580. doi:10.1080/01605682.2018.1448247
2. Chang LC, Lin SW, Northcott DN. The NHS Performance Assessment Framework: a "balanced scorecard" approach? *J Manag Med*. 2002;16(4-5):345-58. doi:10.1108/02689230210446526
3. Amer F, Hammoud S, Khatatbeh H, Lohner S, Boncz I, Endrei D. The deployment of balanced scorecard in health care organizations: is it beneficial? A systematic review. *BMC Health Services Research*. 2022/02/10 2022;22(1):170. doi:10.1186/s12913-022-07555-9
4. Kaplan RSN, D. P. The Balanced Scorecard: Measures That Drive Performance. *Harvard Business Review*. 2005;70(71-79)
5. (NHPC) NHPC. *National Report on Health Sector Performance Indicators 2003*. 2003. https://www.aihw.gov.au/getmedia/bd83a5e9-c4d6-497a-9f1b-380aa3fc71b9/nrhspi03-c01.pdf.aspx
6. Ma S, Xu X, Trigo V, Ramalho NJ. Doctor-patient relationships (DPR) in China. *J Health Organ Manag*. Mar 20 2017;31(1):110-124. doi:10.1108/JHOM-09-2016-0165
7. Bank W. *Current Health Expenditure (% of GDP) - China*. 2020. https://data.worldbank.org/indicator/SH.XPD.CHEX.GD.ZS?locations=CN&most_recent_value_desc=true&view=map
8. Zeng W, Ma S, Callan VJ, Wu L. Exploring the doctor-patient relationship as a challenge job demand: application of the job demands-resources model in a Chinese public hospital. *Psychol Health Med*. Sep 2022;27(8):1661-1671. doi:10.1080/13548506.2021.1916952
9. Commission GPH. *Brief Report on the Medical Resources and Services in Guangdong Province for the Year 2022 [2022 年广东省医疗卫生资源和医疗服务情况简报*. 2023. 2023 Jul 3. https://wsjkw.gd.gov.cn/zd_ylfwxxgk/content/post_4235145.html
10. He AJ. The doctor-patient relationship, defensive medicine and overprescription in Chinese public hospitals: evidence from a cross-sectional survey in Shenzhen city. *Soc Sci Med*. Dec 2014;123:64-71. doi:10.1016/j.socscimed.2014.10.055
11. Opinions on Strengthening Performance Appraisal of Tertiary Public Hospitals (2019).
12. Greiling D, Manville G. Balanced scorecard implementation in German non‐profit organisations. *International Journal of Productivity and Performance Management*. 2010;59(6):534-554. doi:10.1108/17410401011063939
13. Noto G, Lo Verso AC, Barresi G. What is the performance in public hospitals? A longitudinal analysis of performance plans through topic modeling. *BMC Health Serv Res*. Apr 9 2021;21(1):326. doi:10.1186/s12913-021-06332-4
14. Organization WH. *The World Health Report 2000: Health Systems—Improving Performance*. 2000.
15. Fekri O ME, Klazinga N. *Health System Performance Assessment in the WHO European Region: Which Domains and Indicators Have Been Used by Member States for Its Measurement?* 2018. 55. https://www.ncbi.nlm.nih.gov/books/NBK519096/
16. Veillard J, Champagne F, Klazinga N, Kazandjian V, Arah OA, Guisset AL. A performance assessment framework for hospitals: the WHO regional office for Europe PATH project. *Int J Qual Health Care*. Dec 2005;17(6):487-96. doi:10.1093/intqhc/mzi072

17. Health Do. *Annual Report and Accounts 2011-12 (HC 66)*. 2012. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/212977/23735_HC-66-DoH.pdf
18. Verzulli R, Jacobs R, Goddard M. Autonomy and performance in the public sector: the experience of English NHS hospitals. *Eur J Health Econ*. May 2018;19(4):607-626. doi:10.1007/s10198-017-0906-6
19. McLoughlin V, Leatherman S, Fletcher M, Owen JW. Improving performance using indicators. Recent experiences in the United States, the United Kingdom, and Australia. *Int J Qual Health Care*. Dec 2001;13(6):455-62. doi:10.1093/intqhc/13.6.455
20. Thomas Craig KJ, McKillop MM, Huang HT, George J, Punwani ES, Rhee KB. U.S. hospital performance methodologies: a scoping review to identify opportunities for crossing the quality chasm. *BMC Health Serv Res*. Jul 10 2020;20(1):640. doi:10.1186/s12913-020-05503-z
21. Kramers PG. The ECHI project: health indicators for the European Community. *Eur J Public Health*. Sep 2003;13(3 Suppl):101-6. doi:10.1093/eurpub/13.suppl_1.101
22. Oliveira HC, Rodrigues LL, Craig R. Bureaucracy and the balanced scorecard in health care settings. *Int J Health Care Qual Assur*. Apr 8 2020;ahead-of-print(ahead-of-print)doi:10.1108/IJHCQA-07-2019-0121
23. Sevim F, Aldogan EU. Evaluation of Health Systems Performance of OECD Countries Using MOORA Method. *Journal of Health Management*. 2024;26(1):172-183. doi:10.1177/09720634231215131
24. Li Y, He W, Yang L, Zheng K. A historical review of performance appraisal of public hospitals in China from the perspective of historical institutionalism. *Front Public Health*. 2022;10:1009780. doi:10.3389/fpubh.2022.1009780
25. Yuan H, Ferreira FAF. Sustainable competitive advantage in maternal and child health institutions: a dynamic capability approach. *Ann Oper Res*. May 10 2022:1-27. doi:10.1007/s10479-022-04738-9
26. Freeman REE, McVea J. A Stakeholder Approach to Strategic Management. *SSRN Electronic Journal*. 2001;doi:10.2139/ssrn.263511
27. Wu AW. *Study on the Relationship Between Stakeholder Orientation and Performance in Public Non-Profit Hospitals*. Doctoral dissertation. Renmin University of China; 2008.
28. Mitchell RK, Agle BR, Wood DJ. Toward a Theory of Stakeholder Identification and Salience: Defining the Principle of Who and What Really Counts. *The Academy of Management Review*. 1997;22(4)doi:10.2307/259247
29. Stainer A, Stainer L. Total productivity: a stakeholder perspective. *International Journal of Business Performance Management*. 2003;5(2/3)doi:10.1504/ijbpm.2003.003257
30. Xie STS, H.; Cao, G. Reflections on the Development of Public Welfare in Chinese Public Hospitals [我国公立医院公益性内涵发展的思考]. *Chinese Hospital Management (Chin Hosp Manag)*. 2017;37(9):1-3.
31. Bassani G, Leardini C, Campedelli B, Moggi S. The dynamic use of a balanced scorecard in an Italian public hospital. *Int J Health Plann Manage*. May 2022;37(3):1781-1798. doi:10.1002/hpm.3440
32. Catuogno S, Arena C, Saggese S, Sarto F. Balanced performance measurement in research hospitals: the participative case study of a haematology department. *BMC Health Serv Res*. Aug 3 2017;17(1):522. doi:10.1186/s12913-017-2479-6
33. Zelman, W. N., Pink, G. H., & Matthias, C. B. (2003). Use of the balanced scorecard in health care. *Journal of health care finance*, *29*(4), 1-16.
34. Tashakkori AT, C. *Mixed Methodology: Combining Qualitative and Quantitative Approaches*. Sage; 1998.

35. Carter N, Bryant-Lukosius D, DiCenso A, Blythe J, Neville AJ. The use of triangulation in qualitative research. *Oncol Nurs Forum*. Sep 2014;41(5):545-7. doi:10.1188/14.ONF.545-547
36. Dalkey N, Helmer O. An Experimental Application of the Delphi Method to the Use of Experts. *Management Science*. 1963;9(3):458-467.
37. Kendall MG, Smith BB. The Problem of m Rankings. *The Annals of Mathematical Statistics*. 1939;10(3):275-287. doi:10.1214/aoms/1177732186
38. Tomczak ET, M. The Need to Report Effect Size Estimates Revisited: An Overview of Some Recommended Measures of Effect Size. *Trends in Sport Sciences*. 2014;21(1)
39. Miles MBH, A. M. *Qualitative Data Analysis: An Expanded Sourcebook*. 2nd ed. Sage; 1994.
40. Aberdeen T. Yin, R. K. (2009). Case study research: Design and methods (4th Ed.). Thousand Oaks, CA: Sage. *The Canadian Journal of Action Research*. 2013;14(1):69-71. doi:10.33524/cjar.v14i1.73
41. Gibbert M, Ruigrok W, Wicki B. What Passes as a Rigorous Case Study? *Strategic Management Journal*. 2008;29(13):1465-1474.
42. Gartner JB, Lemaire C. Dimensions of performance and related key performance indicators addressed in healthcare organisations: A literature review. *Int J Health Plann Manage*. Jul 2022;37(4):1941-1952. doi:10.1002/hpm.3452
43. Østergaard-Nielsen MR. How does performance feedback shape public managers' inclination to involve frontline professionals? Evidence from a conjoint experiment. *International Public Management Journal*. 2024:1-17. doi:10.1080/10967494.2024.2406432
44. Knox S, Marin-Cadavid C, Oziri V. Stakeholder engagement-as-practice in public sector innovation. *International Public Management Journal*. 2024;28(1):153-168. doi:10.1080/10967494.2024.2423952
45. Elkomy S, Murad Z, Veleanu V. Does leadership matter for healthcare service quality? Evidence from NHS England. *International Public Management Journal*. 2020;26(2):147-174. doi:10.1080/10967494.2020.1828204