Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Case Complexity-Based Incentives: How Payment Reform Strengthens Tiered Healthcare in China

Xinyi Huang
Sun Yat-sen University
Yingmei Li
The Seventh Affiliated Hospital of Sun Yat-sen University
Zilong Zheng
The Seventh Affiliated Hospital of Sun Yat-sen University
Xiaolong Wu
Sun Yat-sen University
Mingming Xu
xumm27@mail.sysu.edu.cn

Sun Yat-sen University

Article

Keywords: Health Insurance Payment Reform, Case-Mix Complexity, Tiered Healthcare System, Interrupted Time Series Analysis, Causal Forest

Posted Date: March 8th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8078362/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Healthcare systems worldwide continue to struggle with hospital-centered utilization and insufficient primary care capacity. In China, the Diagnosis–Intervention Packet (DIP) reform introduces a regional global budget and complexity-based reimbursement, aiming to discourage low-severity admissions in tertiary hospitals while incentivizing the treatment of resource-intensive cases. This study evaluates whether and how the DIP reform reshaped hospital case-mix and strengthened the functioning of China's tiered healthcare system.

Using inpatient records from a tertiary hospital (2019–2020), we applied a two-group interrupted time series design to estimate the reform's causal impact on case complexity, measured by Relative Weight (RW). We further employed a causal forest model to examine heterogeneity in treatment effects across patient groups. Results indicate a significant post-reform increase in RW among patients covered under the DIP scheme, reflecting a shift toward more complex case-mix and enhanced hospital specialization. Heterogeneity analysis shows that age and department accounted for nearly 90% of variation in treatment effects. The most pronounced increases were observed among children and adolescents (0–18 years) and in neonatology and pediatrics, suggesting that the reform rebalanced pediatric case-mix within tertiary hospitals, increasing the relative share of higher-complexity cases compared to routine, low-severity cases.

These findings demonstrate that complexity-based payment incentives can realign hospital behavior and advance functional differentiation in a tiered healthcare system. To ensure sustainable impact, payment reforms should be accompanied by strengthened primary and community care capacity, safeguarding quality and equity as patient flows are redistributed.

1. Introduction

Across both advanced and emerging economies, healthcare systems are confronted with enduring structural inefficiencies. A pervasive challenge—observed from the United States and Australia to China and India—is the excessive concentration of patients in tertiary hospitals, while primary and community care facilities remain underutilized [1–5]. This imbalance contributes to fragmented service delivery and inflated healthcare costs [6]. The tiered healthcare model, which allocates patients to appropriate levels of care according to medical necessity, has been widely promoted as a systemic solution to these problems [7–8]. Yet, translating this model from policy aspiration to institutional reality remains difficult. Misaligned provider incentives and deep-rooted patient preferences continue to reinforce hospital-centered utilization patterns [9–10]. Against this background, reforming health insurance payment systems has emerged as a potential policy lever to realign incentives, reshape provider behavior, and redirect patient flows in support of tiered care.

Over the past several decades, global health financing reforms have increasingly moved away from traditional **Fee-for-service (FFS)** models toward payment mechanisms designed to contain costs and improve efficiency by reshaping provider incentives. The FFS approach, which rewards service volume, has long been criticized for encouraging overtreatment and escalating healthcare costs [11]. By rewarding service volume rather than value, FFS created distorted incentives that undermined efficiency and equity in healthcare systems worldwide [12]. Empirical studies have shown that under FFS, providers often prioritize high-frequency, revenue-generating services such as imaging, laboratory tests, or minor procedures, even when their clinical necessity is limited [13]. This behavior not only inflates costs but also exposes patients to unnecessary interventions, raising concerns about safety and quality of care. Moreover, FFS tends to reinforce hospital-centered utilization patterns, as tertiary hospitals compete to attract even mild or routine cases in order to maximize admissions. Such dynamics exacerbate overcrowding in higher-level facilities while leaving primary and community care underutilized [14]. In addition, the payment uncertainty associated with FFS—where revenue is tied directly to fluctuating patient volumes—creates financial instability for providers and weakens incentives for preventive care or long-term disease management. Taken together, these shortcomings illustrate why FFS is widely regarded as unsustainable: it distorts provider behavior, undermines rational resource allocation, and perpetuates inefficiencies that health systems across both high- and middle-income countries have sought to correct through reform [15].

In response, a range of prospective payment models have been introduced across high- and middle-income countries. One of the most influential has been the **Diagnosis-Related Groups (DRGs)** system, which represents a fundamental departure from volume-based reimbursement. Under DRGs, hospitals receive a fixed payment for each patient admission, determined by diagnostic category and the expected level of resource consumption [16]. This classification system groups patients with similar clinical conditions and anticipated treatment costs, thereby standardizing reimbursement across institutions. By design, DRGs shift financial risk from insurers to providers, compelling hospitals to manage expenditures within the predetermined payment level and discouraging unnecessary extensions of care [17]. The system has been widely adopted in the United States—particularly within the Medicare program since the 1980s—and subsequently diffused across Europe, where it has been credited with reducing average length of stay, curbing cost escalation, and promoting greater efficiency in hospital management. Beyond efficiency gains, DRGs also encouraged hospitals to improve internal cost accounting and resource allocation, fostering more transparent performance monitoring [18–19]. Nevertheless, concerns remain regarding unintended consequences. Because reimbursement is fixed, hospitals may be incentivized to avoid admitting patients with unusually complex conditions that exceed the expected resource envelope, or to discharge patients prematurely in order to remain within budget constraints [18]. Critics argue that such behaviors could compromise service quality and equity, particularly for vulnerable populations requiring intensive care. Thus, while DRGs have proven effective in containing costs and rationalizing hospital operations, they also highlight the delicate balance between efficiency incentives and the safeguarding of patient outcomes [17, 20].

Another reform strategy is **Bundled Payments**, which provide a single, comprehensive reimbursement covering all services related to a specific episode of care, such as a surgical procedure, maternity care, or chronic disease management [21]. Unlike fee-for-service, which fragments payments across individual encounters, bundled payments link reimbursement to the entire continuum of treatment, thereby encouraging coordination among providers and discouraging redundant or unnecessary interventions [22]. By aligning financial incentives across surgeons, anesthesiologists, nurses, and post-acute care providers, this model fosters integrated care pathways and strengthens accountability for patient outcomes. International experience illustrates its potential: in the United States, bundled payment pilots under Medicare have demonstrated reductions in hospital readmissions and improvements in care

standardization for orthopedic procedures [23]; in European countries such as the Netherlands and Sweden, bundled payments have been applied to chronic conditions like diabetes, leading to more consistent monitoring and better adherence to treatment guidelines [24]. Evidence suggests that bundled payments can reduce duplicative testing, streamline resource use, and enhance patient satisfaction by providing more seamless care. Nevertheless, important challenges remain. Because bundled payments require detailed episode definitions and complex negotiations among multiple providers, they often entail significant administrative burden and higher transaction costs [25]. Moreover, the model may be less effective for patients with multimorbidity or unpredictable treatment trajectories, where rigid episode boundaries fail to capture the diversity of care needs [26]. In such cases, bundled payments risk under-compensating providers and may inadvertently compromise equity in service provision [27]. Thus, while bundled payments represent a promising strategy for improving integration and efficiency, their design and implementation must be carefully tailored to balance administrative feasibility with the realities of complex patient populations.

A third major approach is **Capitation**, in which providers receive a fixed payment per patient over a defined period, regardless of the number of services delivered [28]. This model represents a fundamental shift from volume-based reimbursement to population-based financing, as revenue is tied to the size of the enrolled patient pool rather than the frequency of encounters. By transferring financial risk to providers, capitation aims to control costs and encourage preventive care, since providers benefit financially when patients remain healthy and require fewer high-cost interventions [29]. When combined with performance-based incentives, capitation can promote quality improvements, particularly in chronic disease management, where proactive monitoring and early intervention reduce long-term expenditures. International experience illustrates its potential: in the United Kingdom, general practitioners under the National Health Service are largely funded through capitation, which has strengthened their role in primary care and preventive services [30]; in the United States, health maintenance organizations (HMOs) have applied capitation to encourage integrated care delivery [31]. Nevertheless, significant challenges persist. Without adequate risk adjustment, providers may be reluctant to accept patients with complex or unpredictable conditions, leading to inequities in access. Administrative burdens are also considerable, as capitation requires detailed contracts, continuous data reporting, and monitoring of performance indicators. Voluntary enrollment can undermine stability if healthier patients opt out, leaving providers with disproportionately high-risk populations [32]. Moreover, uncertainty in the distribution of performance bonuses, or their outright withholding, may weaken incentives and erode trust between payers and providers. Taken together, these limitations highlight that while capitation offers a promising pathway for cost containment and preventive care, its effectiveness depends critically on robust risk adjustment, transparent incentive structures, and sufficient investment in primary care infrastructure [33–34].

Despite the diversity of these payment reforms, most existing studies have concentrated on their immediate effects on expenditure control, service utilization, and efficiency metrics [35–36]. For example, evaluations often measure reductions in hospital admissions, shorter lengths of stay, or changes in per capita spending. While these outcomes are important, they provide only a partial view of the broader systemic implications of payment reform. Far less attention has been devoted to examining how such reforms influence the functioning of tiered healthcare systems, particularly the redistribution of patients across primary, secondary, and tertiary levels of care.

Understanding these broader impacts is crucial. Payment reforms not only alter provider incentives but also shape patient flows, institutional roles, and equity in access to care [37]. For instance, DRGs may encourage tertiary hospitals to concentrate on complex cases, while capitation may strengthen the role of primary care providers in managing routine conditions. Bundled payments, by fostering coordination, may blur traditional boundaries between care levels. Yet empirical evidence on these systemic effects remains limited, especially in low- and middle-income countries where tiered healthcare delivery is a central policy goal. Addressing this gap is essential to evaluate whether financing reforms truly advance efficiency, specialization, and equity within healthcare systems [38].

Recently, some studies have begun to explore the link between payment reform and referral behavior under the tiered healthcare system. They found that, under regional global budget systems, tertiary hospitals have been observed to increase referrals to primary care facilities to manage total expenditures, resulting in rising patient volumes at the grassroots level [39–41]. However, these findings provide only indirect evidence of system-wide functional differentiation. Critically, existing studies have rarely examined how patients of differing disease severity are redistributed across healthcare tiers, leaving unanswered the central question of whether more complex cases are indeed concentrated in tertiary hospitals while mild conditions are managed in primary care. Addressing this empirical gap is essential to understanding whether payment reforms truly advance the goals of a tiered and efficient healthcare system.

In China, the long-standing dominance of the FFS model tied hospital revenues directly to the volume of admissions and procedures, creating strong incentives for tertiary hospitals to compete for even mild or routine cases [42]. This distorted the intended functional hierarchy of the healthcare system and contributed to both inefficiency and inequity, as large hospitals became overcrowded while primary and community facilities remained underutilized. Empirical studies have shown that under FFS, tertiary hospitals frequently admitted patients with common respiratory infections or minor injuries that could have been effectively managed at lower levels of care, thereby consuming scarce specialist resources and inflating medical expenditures. To address these challenges, China introduced the **Diagnosis-Intervention Packet (DIP)** reform—a novel payment mechanism embedded within a regional global-budget framework, which is conceptually similar to DRGs but tailored to China's institutional context. Unlike DRGs, which classify patients into diagnostic groups, DIP assigns disease-specific weights that reflect case complexity and resource intensity, thereby linking hospital reimbursement more directly to the severity of conditions treated [43–44]. At the same time, the regional global-budget design caps total medical insurance expenditures, ensuring fiscal sustainability while incentivizing hospitals to optimize their case mix. By financially rewarding the treatment of complex, resource-intensive cases, the reform seeks not only to contain escalating medical costs but also to strengthen the functional differentiation of healthcare institutions—an essential component of China's broader effort to build an effective tiered healthcare delivery system [45]. Early pilot experiences in cities such as Shenzhen and Guangzhou suggest that DIP has begun to redirect routine cases toward primary care while reinforcing the specialization role of tertiary hospitals, offering valuable lessons for the nationwide rollout of payment reform.

The present study leverages the DIP reform as a policy context to examine how payment incentives can facilitate the functioning of a tiered healthcare system. Using inpatient data from a tertiary hospital, we apply a two-group interrupted time series approach to estimate the reform's overall impact on hospital case-mix, and employ a causal forest model to uncover heterogeneity across patient groups. By linking a complexity-based payment mechanism to shifts in hospital specialization, this study provides new empirical evidence on how health insurance payment reform can realign provider behavior and support tiered care delivery. Beyond the Chinese context, the findings offer insights for low- and middle-income countries seeking to strengthen healthcare system efficiency through financing reforms.

2. Method

2.1 Data sources and variables

We analyzed inpatient records from a large tertiary hospital in Shenzhen, China, covering January 2019 to December 2020. This institution is one of the major referral centers in the region, with more than 2,000 beds and a wide range of specialized departments, including pediatrics, cardiology, oncology, and surgery. As a teaching hospital affiliated with a medical university, it attracts patients not only from Shenzhen but also from surrounding cities and provinces, making it representative of the challenges faced by tertiary hospitals under China's tiered healthcare system. The hospital maintains a comprehensive electronic medical record (EMR) system, which provided detailed patient-level information on diagnoses, procedures, and demographic characteristics. The two-year observation window captures 12 months before and 12 months after the launch of the DIP payment reform in January 2020, yielding 24 monthly data points that allow us to examine both immediate and gradual changes in hospital behavior.

The key outcome variable was Relative Weight (RW), the index used in the DIP system to reflect case complexity and resource intensity, and to determine the level of reimbursement for each case. RW is calculated based on disease classification codes, treatment procedures, and expected consumption of medical resources such as diagnostic tests, medications, and length of stay. A higher RW value indicates that the case is more severe or requires more intensive resource input, thereby justifying higher reimbursement under the global budget framework. In practice, RW serves as a proxy for case-mix complexity, enabling us to assess whether the reform incentivized tertiary hospitals to concentrate on complex, resource-intensive patients rather than routine admissions. This measure is particularly important in the Chinese context, where hospitals previously competed for low-severity cases under the fee-for-service model, often admitting patients with minor conditions to secure revenue.

Additional covariates were collected at the patient level to control for demographic and socio-economic differences that might influence case complexity. These included age, sex, marital status (married versus others), occupation type (employed, agriculture, self-employed, retired, and unemployed), and the department of admission. Age and department are especially relevant, as they strongly shape disease profiles and treatment intensity—for example, pediatric and neonatal departments often face distinct case-mix patterns compared to internal medicine or surgery. Occupation type was included to capture socio-economic variation that may affect health status and access to care. Marital status was considered as a proxy for social support, which can influence hospitalization decisions. Incorporating these covariates allows us to disentangle the reform's impact from underlying patient characteristics, ensuring that observed changes in RW reflect policy effects rather than demographic shifts.

2.2 Study Design

We employed a two-group interrupted time series (ITS) design to estimate the impact of the DIP reform on case complexity (RW). This approach was chosen to align with both the temporal structure of the data and the analytic goals of the study [46]. The dataset includes dense monthly observations across 24 consecutive periods surrounding a clearly defined policy implementation point, which allows for explicit modeling of both immediate level changes and post-intervention slope changes—core estimates in ITS.

Compared with a conventional difference-in-differences framework that focuses on average before–after differences, the ITS approach captures the temporal dynamics of policy effects, making it particularly suited for evaluating gradual or time-dependent behavioral responses. In addition, ITS directly accounts for potential serial correlation and short-term shocks (e.g., COVID-19 disruptions) that may otherwise bias estimates.

Patients were classified into a treatment group (those covered under the DIP-based reimbursement scheme) and a control group (patients without Shenzhen health insurance who were not subject to the DIP scheme during the same period). This quasi-experimental design allows for comparison of changes between groups while controlling for pre-existing trends, thereby strengthening causal inference. The inclusion of a control group also helps to distinguish reform effects from broader secular changes in case complexity.

To improve model validity, several diagnostic and corrective steps were undertaken. Model diagnostics indicated autocorrelation in the initial two-group ITS specification. To address this issue, a lagged dependent variable was included in the model, which effectively corrected the autocorrelation as confirmed by the Durbin–Watson test. Moreover, the dependent variable was log-transformed to address right-skewness in the distribution, ensuring that estimates were not unduly influenced by extreme values. We also tested alternative specifications with different lag structures and found results to be robust.

The final two-group ITS model thus incorporated: (i) group indicators, (ii) time trends, (iii) policy intervention dummies, (iv) interaction terms for group-specific slopes and level changes, and (v) a lagged dependent variable. This specification allows us to estimate both the immediate and gradual impacts of the DIP reform while controlling for baseline differences and temporal autocorrelation. The two-group ITS model was specified as:

$$\ln Y_{i,t} = \beta_0 + \beta_1 Time_{i,t} + \beta_2 DIP_{i,t} + \beta_3 Post_{i,t} + \beta_4 Group_{i,t} + \beta_5 Group_{i,t} * Time_{i,t} + \beta_6 Group_{i,t} * DIP_{i,t} + \beta_7 Group_{i,t} * Post_{i,t} + \gamma$$

Where Y is the RW measured at the individual level, and $Y_{i,t-1}$ denotes the one-period lag of the dependent variable. $Time$ is the continuous monthly variable (0–23). DIP is a binary variable marking the onset of the DIP reform (0 = pre-reform; 1 = post-reform). $Post$ is the time trend after reform (0 = pre-

reform; 1–12 = post-reform). $Group$ is a binary variable representing the treatment and control groups. X represents the covariates. ϵ is the residual term. β_1 to β_4 capture baseline trends for the control group. Interaction terms β_5 to β_7 capture differences between treatment and control groups, including level shifts (β_6) and slope changes (β_7) attributable to the reform.

2.3 Heterogeneity Analysis: Causal Forest

To examine whether the impact of the DIP reform varied across patient subgroups, we employed a causal forest model—a machine learning method specifically designed to estimate Conditional Average Treatment Effects at the individual level. Unlike traditional linear models that assume a homogeneous treatment effect, the causal forest allows treatment impacts to vary flexibly with patient characteristics, thereby capturing complex and nonlinear heterogeneity [47].

The method builds an ensemble of decision trees trained on random subsamples, recursively partitioning patients by covariates to identify subgroups with differing responses to the reform. Averaging across trees yields stable individualized treatment effect estimates while reducing overfitting. The estimated individualized treatment effects could be further aggregated to quantify how the reform's impact differed across certain demographic and clinical subgroups [48].

In practice, causal forests extend the random forest framework by incorporating treatment assignment into the splitting criteria, ensuring that partitions are formed not only to maximize predictive accuracy but also to differentiate treatment effects. Each tree is grown on a bootstrap sample, and splitting rules are chosen to maximize heterogeneity in treatment responses across branches. By combining hundreds of such trees, the model produces robust estimates of individualized treatment effects that are less sensitive to noise in any single partition.

This approach is particularly suitable for evaluating health policy reforms such as DIP, where treatment effects may depend on multiple interacting factors (e.g., age, department, occupation). Unlike parametric models, causal forests do not require specifying functional forms or interaction terms a priori, thereby avoiding model misspecification. Moreover, the method provides variable importance measures that quantify the relative contribution of each covariate to treatment effect heterogeneity, offering additional insights into which patient characteristics drive differential impacts.

3. Results

3.1 Trends in case complexity over time

Figure 1(a) displays the temporal trends in mean RW values across three groups: the overall population (red line), the treatment group (blue line), and the control group (green line). All three lines span a 24-month period from January 2019 to December 2020, with a vertical dashed line marking the implementation of the DIP reform in January 2020. Prior to the reform, the treatment and control groups exhibited broadly parallel trends in RW, supporting the parallel-trend assumption. The control group reflects natural fluctuations in case complexity in the absence of reform, including a temporary rise and fall in 2020 likely attributable to disruptions caused by the COVID-19 pandemic.

In contrast, the treatment group shows a distinct upward shift in RW following the reform, indicating a rise in case complexity relative to the control group. This trend is further illustrated in Figure 1(b), which presents box plots comparing RW values before and after the reform. The treatment group's RW distribution shifted upward post-reform, with a higher median and narrower interquartile range, suggesting both an increase in complexity and greater consistency in case-mix. A horizontal dashed line marks the overall median RW value (57.5) across the dataset. Department-level variations in RW changes are detailed in the Appendix, with pediatric and surgical departments showing particularly pronounced increases.

3.2 Results of the two group time series analysis

The two group time series analysis estimates are shown in Table 1. The model controls for a range of patient-level covariates—including age, sex, marital status, occupation, and department of admission—as well as a lagged dependent variable to account for autocorrelation in RW values. The results indicates that RW values (β_4) and time trends (β_5) between the treatment and control groups were not significantly different prior to the reform, supporting the parallel-trend assumption. The average RW increased significantly by approximately 1.01% (=($e^{0.010}$−1)×100%) per month following the reform (β_7), but not immediately after the reform (β_6). It suggests a systematic shift toward more complex cases in tertiary hospitals.

Table 1
Results of the two group time series analysis

	Estimate	Standard Error
Time ($\varvec{\beta_1}$)	0.003*	0.002
DIP ($\varvec{\beta_2}$)	-0.006	0.018
Post ($\varvec{\beta_3}$)	-0.007**	0.003
Group ($\varvec{\beta_4}$)	0.023	0.017
Time×Group ($\varvec{\beta_5}$)	-0.005**	0.002
DIP×Group ($\varvec{\beta_6}$)	0.018	0.023
Post×Group ($\varvec{\beta_7}$)	0.010***	0.003
Intercept ($\varvec{\beta_0}$)	1.570***	0.082
$\ln\left(\varvec{Y}_{\varvec{i},\varvec{t}-1}\right)$ $\varvec{\gamma}$	0.548**	0.004
Covariates	**Controlled**	

Notes. Significance levels: *p < 0.1; **p < 0.05; ***p < 0.01.

3.3 Heterogeneity of reform effects at the individual level

The causal forest model identified individualized treatment effects for all observations, revealing significant heterogeneity across patients ($p = 2.2 \times 10^{-16}$). Figure 2 presents the Conditional Average Treatment effect on the Treated (CATT) aggregated by single covariates, while Fig. 3 displays subgroup estimates jointly stratified by both age and department.

Variable importance analysis showed that age and department contributed the most to the observed heterogeneity, accounting for 53.6% and 36.2% of the total variance, respectively—together explaining nearly 90% of the heterogeneity in reform effects. This underscores that variations in reform impact are primarily driven by differences across age groups and clinical departments.

When stratified by age (Fig. 2), patients aged 0–18 years exhibited the highest average treatment effect, with RW increasing by more than 2.68 units post-reform. The 19–44 age group showed moderate increases, while patients aged 45–64 and 65 + experienced smaller changes, with RW values remaining close to pre-reform levels. At the departmental level, the largest increases in RW were observed in neonatology and pediatrics, where the mean treatment effect was substantially higher than in other specialties. Departments such as dermatology, emergency medicine, and ICU also showed positive but less pronounced shifts, whereas community medicine and rehabilitation exhibited minimal changes, with RW distributions largely overlapping before and after the reform.

Joint stratification by age and department (Fig. 3) confirmed that the most substantial RW increases occurred among patients aged 0–18 years across pediatric-related departments. In neonatology, RW values rose sharply after the reform, with a clear upward shift in the distribution. Pediatrics displayed a similar pattern, with consistent increases across the entire age subgroup. Other specialties, including neurology, gastroenterology, and cardiology, showed smaller but still positive changes among younger patients, while older age groups exhibited relatively stable RW values across departments.

Overall, the subgroup analyses demonstrate that the post-reform increases in RW were concentrated among younger patients and pediatric-related departments, with smaller or negligible changes observed in adult and non-specialized departments. These results highlight the presence of strong heterogeneity in reform effects across demographic and clinical subgroups, as captured by the causal forest model.

4. Discussion

This study provides empirical evidence on how complexity-based payment incentives can realign hospital behavior and support the functioning of a tiered healthcare system. Following the introduction of the DIP reform, tertiary hospitals exhibited a significant increase in case complexity, as measured by RW. This upward shift in case-mix suggests that hospitals did not engage in risk avoidance or reduce the admission of severe patients—concerns commonly associated with prospective payment systems such as DRGs [16–18]. Instead, the reform encouraged tertiary hospitals to concentrate resources on more

specialized, high-intensity care, while reducing the relative share of low-severity cases [49]. These findings indicate that the DIP policy effectively reshaped clinical decision-making in ways consistent with the intended role of tertiary hospitals in China's tiered care model.

The observed shift in hospital behavior can be explained by the reform's complexity-adjusted reimbursement mechanism. Under the framework of regional global budgeting, the total pool of medical insurance funds within a given area is predetermined, and each hospital's share is allocated according to the aggregated relative weight (RW) of the cases it treats [50]. The RW value reflects the relative resource intensity and clinical complexity of different disease categories, thereby linking reimbursement not to service volume but to case severity. This design fundamentally alters institutional incentives by rewarding hospitals for treating complex patients rather than maximizing admissions.

For tertiary hospitals, which operate with higher fixed and technological costs, the marginal financial return from admitting patients with complex conditions becomes more favorable than treating low-severity cases [51]. This shift is driven by several factors. First, the RW assigned to severe cases is higher, meaning that reimbursement under the global budget system more accurately reflects the substantial resource consumption required for advanced diagnostics, intensive therapies, and specialized personnel. Second, tertiary hospitals face structural cost pressures—including expensive equipment, highly trained staff, and large-scale facilities—that cannot be covered adequately by treating low-RW, routine cases. By focusing on complex patients, hospitals can better align their revenue streams with their cost structures. Third, the reform reduces the incentive to pursue volume-driven admissions, since low-severity cases contribute little to overall reimbursement but still consume capacity. Concentrating on severe cases thus allows tertiary hospitals to preserve resources for patients whose treatment justifies the higher reimbursement levels. Finally, this adjustment also strengthens their functional positioning within the tiered healthcare system, reinforcing their role as centers of specialized care while redirecting mild conditions to primary and secondary facilities.

Beyond the immediate financial incentives, the DIP reform also reshaped hospital management practices. Departments were compelled to improve internal cost accounting, strengthen clinical pathways, and adopt more rigorous triage mechanisms to ensure that admissions aligned with reimbursement priorities. This managerial transformation highlights that payment reform can exert influence not only on admission decisions but also on broader organizational governance, fostering a culture of efficiency and accountability.

In this way, the reform encourages hospitals to optimize their case mix rather than expand service volume, aligning hospital-level strategies with system-level objectives of efficiency, specialization, and rational patient distribution. RW values thus serve not only as a technical parameter in reimbursement but also as a critical indicator for evaluating whether tertiary hospitals are fulfilling their designated role within the tiered healthcare delivery system.

Heterogeneity analysis further indicates that the impact of the reform was especially salient in pediatric departments and among patients aged 0–18 years. Before the introduction of complexity-based payment, parental decision-making played a critical role in shaping care-seeking patterns [52]. Strong parental risk aversion, combined with a cultural preference for tertiary hospitals, often led families to bypass primary and secondary facilities even for mild or routine pediatric conditions [53]. As a result, tertiary hospitals experienced a disproportionately high volume of low-severity pediatric admissions. This influx of uncomplicated cases diluted the overall case complexity profile of tertiary departments, undermining their intended role as centers for specialized and resource-intensive care.

The reform altered this dynamic by linking reimbursement more closely to case complexity. Under the new incentive structure, pediatric departments became more selective in admissions, prioritizing children with severe illnesses or conditions requiring advanced technological and clinical resources. Routine pediatric illnesses, such as common respiratory infections or minor injuries, were increasingly managed in primary or secondary care settings, where costs are lower and services are more appropriately aligned with patient needs. This redistribution of patient flows not only improved efficiency but also reinforced the functional positioning of different tiers within the healthcare system.

The pediatric case illustrates how financial incentives interact with cultural norms and behavioral preferences. In contexts where families traditionally equate tertiary hospitals with higher quality, reforms like DIP can gradually shift perceptions by signaling that mild conditions are better managed at lower levels of care. This cultural recalibration is critical for sustaining tiered healthcare, as it reduces inappropriate demand pressures on tertiary hospitals and strengthens trust in primary care providers.

The rebalancing of pediatric case-mix illustrates how payment reform interacts with existing cultural and behavioral factors to correct structural distortions in hospital admissions. By discouraging the concentration of mild cases in tertiary hospitals, the reform promoted a more rational allocation of resources and strengthened the pathways for tiered healthcare delivery. In this way, the pediatric experience provides a clear example of how financial incentives can reshape care-seeking behavior, enhance system efficiency, and ensure that tertiary hospitals fulfill their role in managing complex and high-risk patients.

This study contributes to the international literature in three ways. First, it demonstrates that payment reform can influence not only expenditure and healthcare utilization, but also the promotion of a tiered healthcare system. Second, by applying causal forest models, it highlights that reform effects are not uniform and depend on demographic and clinical characteristics—an insight often masked in aggregate evaluations. Third, it provides evidence from a large middle-income country, offering policy-relevant lessons for health systems facing overcrowded tertiary hospitals and underutilized primary care.

International comparison further underscores the distinctiveness of DIP. While DRGs in the United States and Europe have successfully reduced average length of stay, they have also raised concerns about risk avoidance and premature discharge [18–19]. Bundled payments, implemented in countries such as the Netherlands and Sweden, improved care integration but struggled with multimorbidity [23–24]. Capitation models, widely used in the United Kingdom, incentivized preventive care but often lacked adequate risk adjustment [30–31]. Against this backdrop, DIP represents a hybrid approach: it combines global budget discipline with complexity-based differentiation, making it particularly suited to China's challenge of overcrowded tertiary hospitals and underutilized primary care [45].

From a policy perspective, payment reforms that rely on complexity-based incentives should be accompanied by investments in primary and community care. As routine cases shift away from tertiary hospitals, insufficient capacity at the primary and secondary level could lead to care quality concerns, delayed treatment, or increased emergency department utilization. Strengthening the health workforce, diagnostic capabilities, and two-way referral mechanisms is therefore essential to ensure that case reallocation is clinically appropriate, and sustainable.

Moreover, successful implementation requires robust information systems to track RW values, monitor patient flows, and ensure transparency in fund allocation. Without reliable data infrastructure, hospitals may struggle to adjust case-mix strategies or demonstrate compliance with reform objectives. Policymakers should also consider equity safeguards, ensuring that vulnerable populations—such as rural residents or patients with rare diseases—are not disadvantaged by complexity-based reimbursement.

Limitations should be acknowledged. The study is based on data from a single tertiary hospital and reflects short-term effects; broader generalization should be made cautiously. Future work should incorporate multi-center data, examine longer-term reform impacts, and assess outcomes related to care quality, equity, and patient experience.

In addition, future research could explore patient perspectives, examining whether reforms alter trust in primary care and satisfaction with referral pathways. Qualitative studies may provide insights into how cultural expectations and risk perceptions evolve under new payment incentives. Such evidence would complement quantitative analyses, offering a more holistic understanding of how financing reforms reshape healthcare systems.

5. Conclusions

This study shows that China's DIP reform increased case complexity in tertiary hospitals and promoted a more rational alignment between case severity and institutional function. The reform rebalanced hospital case-mix, particularly in pediatrics, where pre-reform care-seeking patterns had led to disproportionate tertiary utilization for routine conditions. By linking reimbursement to case complexity, the reform strengthened the specialization role of tertiary hospitals within the tiered care system. To ensure sustained improvements, payment reforms should be coupled with continued reinforcement of primary and community care capacity, enabling patient redirection to occur safely and effectively.

Beyond these immediate findings, the study contributes to the broader literature by demonstrating that payment reform can influence not only expenditure and utilization but also the structural functioning of a tiered healthcare system. Methodologically, the use of interrupted time series and causal forest analysis highlights the importance of examining heterogeneous effects across patient groups, offering a more nuanced understanding of reform impacts. From a policy perspective, the DIP reform illustrates how complexity-based incentives can correct distortions created by fee-for-service models, providing lessons for other middle-income countries struggling with overcrowded tertiary hospitals and underutilized primary care.

Future research should extend this analysis to multi-center datasets and longer time horizons, while incorporating measures of care quality, equity, and patient satisfaction. Such evidence will be critical to assess whether the observed improvements in case-mix are sustained and whether reforms truly advance the goals of efficiency, specialization, and fairness in healthcare delivery. In this way, the DIP experience not only informs China's ongoing health system transformation but also offers valuable insights for global debates on financing reforms and the pursuit of universal health coverage.

Declarations

Competing interests

The authors declare that there are no competing interests.

Data Availability

The hospital records utilized in this study were obtained under institutional license and are subject to strict privacy and confidentiality restrictions. Consequently, these data are not publicly accessible. Access to the dataset may, however, be granted by the authors upon reasonable request and with the explicit permission of the relevant institution.

Ethical Approval

This study is based exclusively on secondary analysis of anonymized hospital records obtained under institutional license. No direct involvement of human participants or identifiable personal data was required. In accordance with relevant guidelines and regulations, ethical approval was not applicable for this research.

Informed Consent

This study was conducted using secondary analysis of anonymized hospital records obtained under institutional license. As the research did not involve direct contact with human participants or the collection of identifiable personal information, informed consent was not applicable. In accordance with relevant national regulations and institutional policies governing the use of anonymized secondary data, the requirement for informed consent was formally waived.

Author Contributions

A. performed data curation, formal analysis, methodology, and wrote the original draft. B. contributed to data curation, literature review, and writing of the original draft. C. was responsible for data curation. D. contributed to conceptualization, writing of the original draft, and review and editing. E. contributed to editing and supervision. All authors reviewed and approved the final manuscript.

Funding Statement

This work was supported by The Medical Scientific Research Foundation of Guangdong Province; Guangdong Philosophy and Social Sciences Planning Project; Guang Dong Basic and Applied Basic Research Foundation; National Natural Science Foundation for Young Scientist Program.

References

1. OKUNLOLA K A, Oreofe I, OHAERI B, et al. Challenges Facing Primary Health Care Services Utilization and Possible Solutions[J]. 2025.
2. Janke A T, Melnick E R, Venkatesh A K. Monthly rates of patients who left before accessing care in US emergency departments, 2017-2021[J]. JAMA Network Open, 2022, 5(9): e2233708-e2233708.
3. Freijser L, Annear P, Tenneti N, et al. The role of hospitals in strengthening primary health care in the Western Pacific[J]. The Lancet Regional Health–Western Pacific, 2023, 33.
4. Zhang H, Zhang S. Reasons for underutilization of community care facilities for the elderly in China[J]. BMC geriatrics, 2024, 24(1): 791.
5. Akhtar M H, Ramkumar J. Primary Health Center: Can it be made mobile for efficient healthcare services for hard to reach population? A state-of-the-art review[J]. Discover Health Systems, 2023, 2(1): 3.
6. Hanson K, Brikci N, Erlangga D, et al. The Lancet Global Health Commission on financing primary health care: putting people at the centre[J]. The Lancet Global Health, 2022, 10(5): e715-e772.
7. Ackley C A. Tiered cost sharing and health care demand[J]. Journal of Health Economics, 2022, 85: 102663.
8. Feng Q Q, Ao Y B, Chen S Z, et al. Evaluation of the allocation efficiency of medical and health resources in China's rural three-tier healthcare system[J]. Public Health, 2023, 218: 39-44.
9. Li B, Chen J. Barriers to community-based primary health care delivery in urban China: a systematic mapping review[J]. International Journal of Environmental Research and Public Health, 2022, 19(19): 12701.
10. Wang X, Seyler B C, Han W, et al. An integrated analysis of spatial access to the three-tier healthcare delivery system in China: a case study of Hainan Island[J]. International Journal for Equity in Health, 2021, 20(1): 60.
11. Ikegami N .Fee-for-Service Payment – An Evil Practice that Must be Stamped Out?[J]. International Journal of Health Policy and Management,2015,4(2):57-59.
12. Elodie A, Hamed M, Shima N. Bundled Payment vs. Fee-for-Service: Impact of Payment Scheme on Performance[J]. Management Science,2017,63(5):1606-1624.
13. Guo P, Tang C S, Wang Y, et al. The impact of reimbursement policy on social welfare, revisit rate, and waiting time in a public healthcare system: Fee-for-service versus bundled payment[J]. Manufacturing & Service Operations Management, 2019, 21(1): 154-170.
14. Di Guida S, Gyrd-Hansen D, Oxholm A S. Testing the myth of fee-for-service and overprovision in health care[J]. Health economics, 2019, 28(5): 717-722.
15. Gosden T, Forland F, Kristiansen I, et al. Capitation, salary, fee-for-service and mixed systems of payment: effects on the behaviour of primary care physicians[J]. Cochrane Database of Systematic Reviews, 1996, 2011(10).
16. Zheng X, Xiao X, Shen K, et al. Impact of Diagnosis-Related Groups (DRG) reform on cost homogeneity of treatment for patients with malignant tumours[J]. Scientific Reports, 2024, 14(1): 21212.
17. Böcking W, Ahrens U, Kirch W, et al. First results of the introduction of DRGs in Germany and overview of experience from other DRG countries[J]. Journal of Public Health, 2005, 13(3): 128-137.
18. Or Z. Implementation of DRG Payment in France: Issues and recent developments[J]. Health policy, 2014, 117(2): 146-150.
19. Polyzos N, Karanikas H, Thireos E, et al. Reforming reimbursement of public hospitals in Greece during the economic crisis: implementation of a DRG system[J]. Health policy, 2013, 109(1): 14-22.
20. Wild V, Carina F, Frouzakis R, et al. Assessing the impact of DRGs on patient care and professional practice in Switzerland (IDoC): a potential model for monitoring and evaluating healthcare reform[J]. Swiss medical weekly, 2015, 145.
21. Agarwal R, Liao J M, Gupta A, et al. The Impact Of Bundled Payment On Health Care Spending, Utilization, And Quality: A Systematic Review: A systematic review of the impact on spending, utilization, and quality outcomes from three Centers for Medicare and Medicaid Services bundled payment programs[J]. Health Affairs, 2020, 39(1): 50-57.
22. Liao J M, Navathe A S. The path ahead for bundled payments[J]. JAMA, 2022, 328(16): 1592-1594.
23. Crowley A P, Kilaru A S, Huang Q E, et al. Physician and Hospital Performance in Medicare's Updated Bundled-Payment Model for Joint Replacement[C]//JAMA Health Forum. American Medical Association, 2025, 6(7): e251930-e251930.
24. Karimi M, Tsiachristas A, Looman W, et al. Bundled payments for chronic diseases increased health care expenditure in the Netherlands, especially for multimorbid patients[J]. Health policy, 2021, 125(6): 751-759.

25. Mechanic R E. Opportunities and challenges for episode-based payment[J]. New England Journal of Medicine, 2011, 365(9): 777-779.
26. Hussey P, Mulcahy A W, Schnyer C, et al. Bundled payment: Effects on health care spending and quality[J]. Santa Monica, CA: RAND August, 2012.
27. Tol J, Swinkels I C S, Struijs J N, et al. Integrating care by implementation of bundled payments: results from a national survey on the experience of Dutch dietitians[J]. International journal of integrated care, 2013, 13: e055.
28. Sambodo N P, Bonfrer I, Sparrow R, et al. Effects of performance-based capitation payment on the use of public primary health care services in Indonesia[J]. Social science & medicine, 2023, 327: 115921.
29. Basu S, Phillips R S, Song Z, et al. High levels of capitation payments needed to shift primary care toward proactive team and nonvisit care[J]. Health Affairs, 2017, 36(9): 1599-1605.
30. Connolly S, Bevan G, Mays N. Funding and performance of healthcare systems in the four countries of the UK before and after devolution: a longitudinal analysis of the four countries, 1996/97, 2002/03 and 2006/07, supplemented by cross-sectional regional analysis of England, 2006/07[J]. 2010.
31. Luft H S. Health maintenance organizations: Dimensions of performance[M]. Transaction Publishers, 1987.
32. WORK H. Capitated Payments[J]. Center for Evidence-based Policy September, 2022: 2.
33. Barber S L, Mathauer I, Rosenberg M, et al. Purchasing primary care services for quality chronic care: Capitation with performance payments in four countries[J]. The International Journal of Health Planning and Management, 2025.
34. Zhang J, Yan J, Shi Y, et al. Does capitation payment influence healthcare service behavior in county medical community? Evidence from patients with diabetes in rural China[J]. BMC Public Health, 2025, 25(1): 1722.
35. Ndayishimiye C, Tambor M, Dubas-Jakóbczyk K. Barriers and facilitators to health-care provider payment reform–a scoping literature review[J]. Risk Management and Healthcare Policy, 2023: 1755-1779.
36. Ghazaryan E, Delarmente B A, Garber K, et al. Effectiveness of hospital payment reforms in low-and middle-income countries: a systematic review[J]. Health policy and planning, 2021, 36(8): 1344-1356.
37. Qu Yixian. Will you be discharged when your quota is up? National Health Insurance Bureau explains health insurance fee control[N]. Economic Observer,2024-04-15(005). DOI: 10.28421/n.cnki.njjgc.2024.000411. [in Chinese]
38. Li Xiangfei, Han Yibing, Yang Wei. A study of policy synergies in 71 DIP pilot cities[J]. China Health Economics,2024,43(05):37-42. [in Chinese]
39. Liu X, Lyu H, Jin H. Bridging care hierarchies through regional global budget payment: evidence from county medical communities in Zhejiang Province, China[J]. Health Policy and Planning, 2025, 40(8): 843-853.
40. Post B. Association of a Medicare outpatient payment reform with hospital-primary care integration: Heterogeneity across markets and physicians[J]. Medical Care, 2021, 59(12): 1075-1081.
41. Crook H L, Zheng J, Bleser W K, et al. How are payment reforms addressing social determinants of health[J]. Policy Implications and Next Steps. Duke Margolis Center for Health Policy, 2021.
42. Li X, Zhang L, Li Z, et al. Patient choice and willingness toward gatekeepers as first-contact medical institutions in Chinese tiered healthcare delivery system: a cross-sectional study[J]. Frontiers in public health, 2021, 9: 665282.
43. Zeng X, Huang W, Yu W. Establishment and Implementation of the DIP Core Disease Classification Management Model [J]. Health Economics Research, 2023, 40(10): 66-73. [in Chinese]
44. Zhang Y, Xu S, Tan G. Unraveling the effects of DIP payment reform on inpatient healthcare: insights into impacts and challenges[J]. BMC Health Services Research, 2024, 24(1): 887.
45. General Office of the State Council. Guiding Opinions of the General Office of the State Council on Promoting the Construction of a Hierarchical Medical System [EB/OL]. (2015-09-11) [2025-10-15].
46. Turner S L, Forbes A B, Karahalios A, et al. Evaluation of statistical methods used in the analysis of interrupted time series studies: a simulation study[J]. BMC medical research methodology, 2021, 21(1): 181.
47. Athey S, Imbens G. Recursive partitioning for heterogeneous causal effects[J]. Proceedings of the National Academy of Sciences, 2016, 113(27): 7353-7360.
48. Venkatasubramaniam A, Mateen B A, Shields B M, et al. Comparison of causal forest and regression-based approaches to evaluate treatment effect heterogeneity: an application for type 2 diabetes precision medicine[J]. BMC Medical Informatics and Decision Making, 2023, 23(1): 110.
49. Ding Y, Jiang L. The Impact of DRG and DIP-Based Multi-Modal Health Insurance Payment Reform on Hospital Financial Management and Optimization Strategies[J]. Chinese Health Economics, 2023, 42(04): 84-86. [in Chinese]
50. Lai Y, Fu H, Li L, et al. Hospital response to a case-based payment scheme under regional global budget: the case of Guangzhou in China[J]. Social Science & Medicine, 2022, 292: 114601.
51. Maeda, J. L., Raetzman, S. O. and Friedman, B. S. What Hospital Inpatient Services Contributed the Most to the 2001-2006 Growth in the Cost Per Case[J]. Health Services Research, 2012, 47(5): 1814–1835.
52. Nicholson E, McDonnell T, Conlon C, et al. Parent's preferences for unscheduled paediatric healthcare: A discrete choice experiment[J]. Health Expectations, 2023, 26(5): 1931-1940.
53. Feldman A G, Sokol R J. Neonatal cholestasis: updates on diagnostics, therapeutics, and prevention[J]. Neoreviews, 2021, 22(12): e819-e836.

Figures

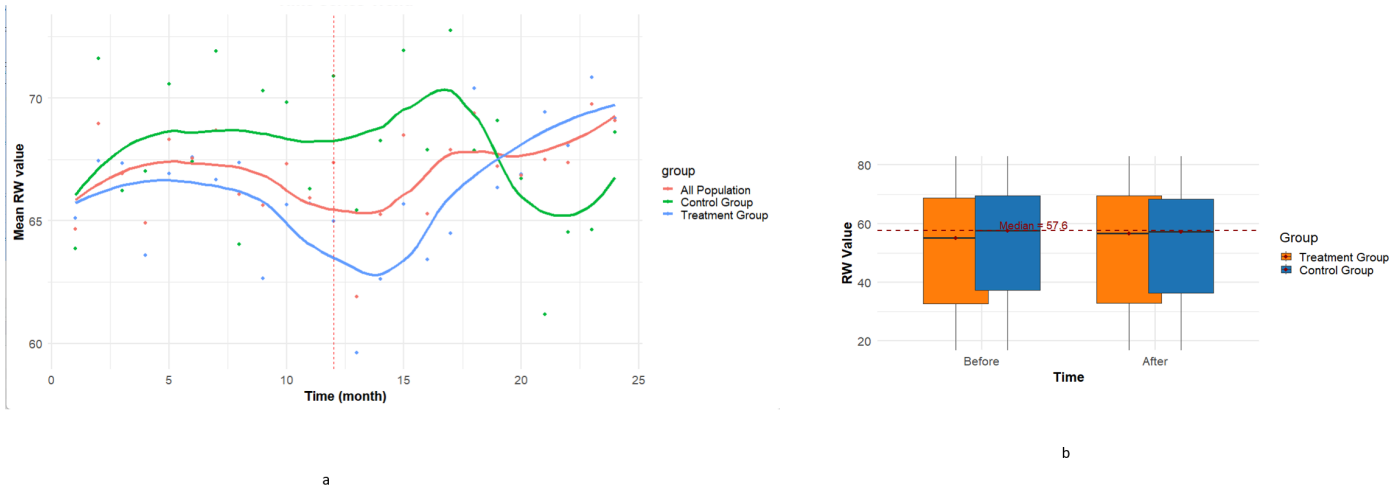


Figure 1

Temporal trends (a) and distribution of case complexity (b) for the treatment and control groups

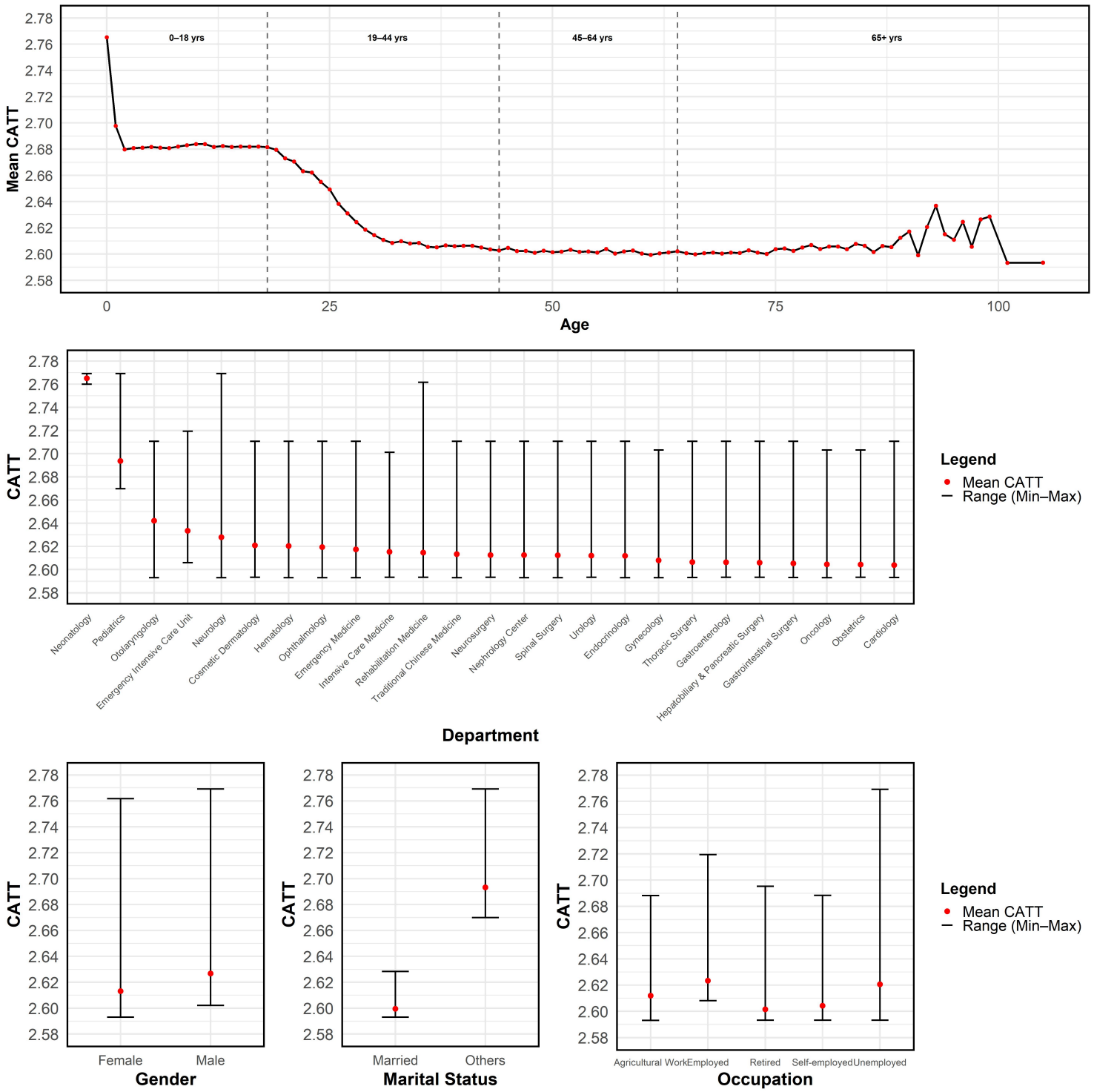


Figure 2

Aggregated individual treatment effects stratified by age, department, gender, marital status, and occupation

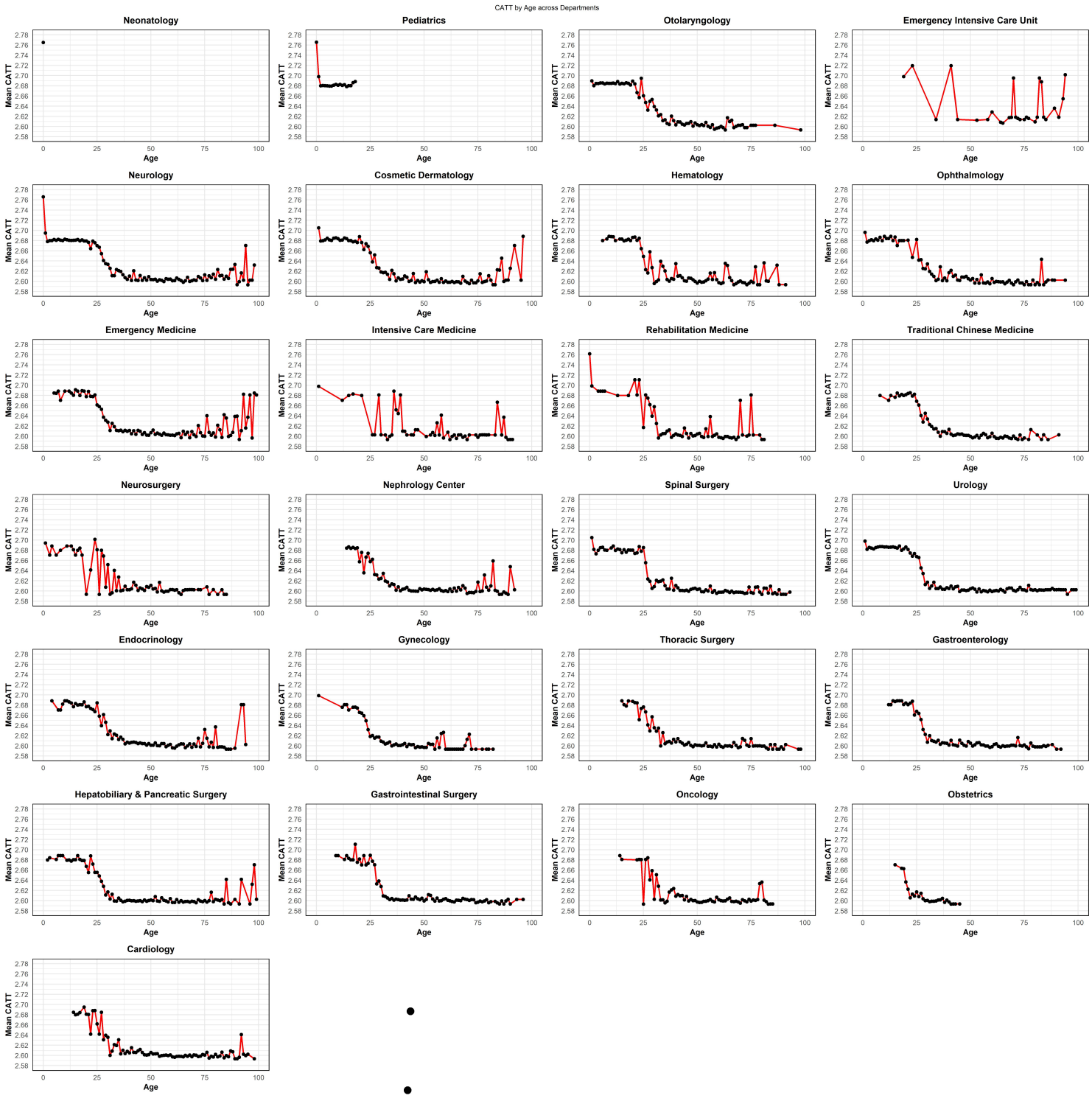


Figure 3

Aggregated individual treatment effects jointly stratified by age and department

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Appendix.docx