Performance differences in public general hospitals when responding to points counting payment based on diagnosis-related group: A robust multiple interrupted time series study in Wenzhou, China

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2023-073913
Article Type:	Original research
Date Submitted by the Author:	25-Mar-2023

Complete List of Authors:	Zhu, Tingting; Wenzhou Medical University, School of Public Health and Management Chen, Chun; Wenzhou Medical University, School of Public Health and Management Zhang, Xinxin; Wenzhou Medical University, School of Public Health and Management Yang, Qingren; Wenzhou Medical University, School of Public Health and Management Hu, Yipao; Health Commission of Wenzhou, Health Information Center Liu, Ruoyun; Wenzhou Medical University, School of Public Health and Management Zhang, Xiangyang; Wenzhou Medical University First Affiliated Hospital Dong, Yin; Health Community Group of Yuhuan People’s Hospital

Keywords:	Health policy < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Organisation of health services < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.
supplemental files.rar

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Performance differences in public general hospitals when responding to points counting payment based on diagnosis-related group: A robust multiple interrupted time series study in Wenzhou, China

Tingting Zhu[1†], Chun Chen[1†], Xinxin Zhang[1], Qingren Yang[2], Yipao Hu[3], Ruoyun Liu[1], Xiangyang Zhang[4*], Yin Dong[5*]

[1] School of Public Health and Management, Wenzhou Medical University, Wenzhou, China

[2] School of Innovation and Entrepreneurship, Wenzhou Medical University, Wenzhou, China

[3] Health Information Center, Health Commission of Wenzhou, Wenzhou, China

[4] Wenzhou Medical University First Affiliated Hospital, Wenzhou, China

[5] Health Community Group of Yuhuan People's Hospital, Taizhou, China

†These authors contributed equally to this work and share first authorship

*** Correspondence:**

Xiangyang Zhang

First Affiliated Hospital of Wenzhou Medical University, Wenzhou, 325035, China, Email: zxyanghero@126.com, TEL: +86 138 5770 2433

Yin Dong

Health Community Group of Yuhuan People's Hospital, Taizhou, 317600, China, Email: 1071259476@qq.com, TEL: +86 189 5767 8008

Keywords diagnosis-related group, points counting payment, performance evaluation, public general hospitals, robust multiple interrupted time series

Word count: 4493
Number of figures: 4
Number of tables: 4

Abstract

Objectives This study measures the differences in inpatient performance of public general hospitals after the policy of points counting payment based on diagnosis-related group (DRG) was implemented from the perspective of institution levels.

Design A robust multiple interrupted time series study using 2018–2021 data was conducted to identify the different performance before and after intervention implementation.

Setting Twenty two above-primary-level public general hospitals (8 tertiary institutions and 14 secondary institutions) in Wenzhou, China.

Intervention The payment policy intervention was officially implemented in January 2020.

Outcome measures Indicators were case-mix index (CMI), cost per hospitalization (CPH), average length of stay (ALOS), cost efficiency index (CEI), and time efficiency index (TEI). The outcomes employed the mean of these indicators.

Results Post implementation of the DRG policy, except for ALOS mean, the estimated change points of other outcomes ($p<0.05$) in tertiary and secondary institutions were inconsistent. Compared with before the change, the CMI mean turned to uptrend with the slope changing 0.0009540 ($p<0.01$) in tertiary institutions and changing 0.003137 ($p<0.0001$) in secondary institutions. Although the slope of CPH mean did not change ($p>0.05$), the uptrend of CEI mean in tertiary institutions alleviated ($p<0.05$) and furthered increase ($p<0.05$) in secondary institutions. Both the slopes of ALOS and TEI mean in secondary institutions changed ($p<0.05$) but did not change in tertiary institutions ($p>0.05$).

Conclusions Hospitals of different tiers react differently to the payment policy; tertiary institutions react faster than secondary institutions in terms of cost control. Focusing on the effect trend, tertiary institutions improve the medical capacity slower than secondary institutions in terms of CMI; implementing the policy can mitigate discrepancies in the expense of treating similar diseases between tertiary and secondary institutions in terms of CPH and CEI; the policy did not motivate tertiary institutions to shorten the ALOS and TEI.

Strengths and limitations of this study

- ► Few studies have focused on the performance differences resulting from the effects of points counting payment based on a DRG policy from the perspective of institutional levels, making the findings of this study meaningful.
- ► This study also scrutinized the latent anticipatory effects after the local DRG policy reform had been initiated, announced, and publicized, given its important role in the healthcare field.
- ► To optimize the effect assessment of policy intervention, this study used a robust multiple interrupted time series analysis that allowed estimated lagged or anticipatory effects, and further improved internal validity, while reducing bias.

► Despite the inevitable impact of the global COVID-19 pandemic on hospital services, this study succeeded in revealing the anticipatory and positive effects of DRG policy on hospitals' performance.

Introduction

The healthcare systems in most countries are confronted with increasing pressure, especially regarding the growing demand for healthcare services, while struggling with growing medical insurance expenditures.[1,2] The diagnosis-related group (DRG) payment system, which originated in the United States in the 1970s, involves the adoption of a standard pricing framework that provides equity across healthcare providers in terms of payments for the same types of services.[3,4] Subsequently, many countries and regions have adopted the DRG-medical-insurance payment method in attempts to contain expenditure and increase the transparency, efficiency, and safety of their healthcare systems.[4-7] Maximizing healthcare performance in hospitals has been prioritized worldwide,[8] which also influenced the primary content of the healthcare system reform in China.[9] China consistently places great importance on the impact of DRG policy on healthcare performance, but the empirical evidence of the DRG impact is still inadequate. Before the effect of DRG on hospital service performance can be assessed, it is necessary to obtain a comprehensive understanding of the policy's role and influence, and to optimize the study design to achieve an optimal evaluation and manage policy makers expectations.

DRG-payment reform, as an important means of regulating medical service behaviors in China,[10] can revolutionize the healthcare system. It involves several complementary elements: a mechanism for allocating funding for hospital services; one for hospital management; and a provider payment mechanism within the broader healthcare financing system.[3] Such a response can have a wide-ranging effect,[11,12] especially, if it can affect hospitals' direct interests that motivate them to regulate clinical behaviors and transform management from passive to positive, which can, in turn, improve medical service performance. The time at which the DRG payment reform was initiated, announced, and publicized, could impact its outcomes.[13] China officially announced the implementation of DRG payment in 30 pilot national monitoring cities, and set the following arrangement that initial implementation work commences in 2019, simulated operation commences in 2020, and actual payment commences in 2021.[14,15] It sent a strong signal to those hospitals whose cities were not among the national pilots but then decided to implement the DRG policy. On one hand, it may have encouraged hospitals to change their management approaches to pursue a more positive adaption of the new payment method, such as by managing medical performance based on DRG' data,[16] improving information levels,[17] and strengthening the quality of diagnoses coding.[18] Such measures can provide an early foundation for the formal implementation of DRG' payment. On the other hand, it is also reasonable to assume that the obvious behavioral changes adopted by hospitals occurred after the official implementation, possibly due to certain problems concerning the requisite technology and operation of the DRG' payment approach.[19] Those hospitals with relatively weak basic conditions required more intensive policy adaptation, such as training personnel in DRG-associated tasks and improving familiarity with the diagnostic and operational codes for diseases. However, previous studies may have neglected to include the time until the effects of the policy manifested. This can reflect the reaction time and degree of importance across different hospitals to some extent.

To advance the reform of its healthcare system, China has devoted itself to the adoption of the DRG as one of the main payment methods and optimize the payment rules and relevant mechanisms to meet Chinese healthcare conditions.[7,20] The DRG scheme of China Healthcare Security were issued in 2019, which identified 26 Major Diagnosis Categories that divided into 376 Adjacent DRGs. The scheme of detailed DRG has been issued in 2020, further dividing the 376 Adjacent DRGs into 618 groups. China Healthcare Security encouraged cities to further expand the amount of their local DRGs based on these 618 groups, from where they can determine the detailed payment method of DRG for the inpatient service. Zhejiang Provincial Healthcare Security Administration furthered divide 376 Adjacent DRGs into 1006 DRGs; 11 cities of this province including Wenzhou, implemented this scheme. Zhejiang had localized the DRGs payment and issued the DRG Points Counting Payment Implementation Rules for Basic Medical Insurance Inpatient Expenses on November 12, 2019, announcing that they would be enforced across the entire province from January 2020.[21] This points counting payment (PCP) method applies the principle of work point theory with reference to the weighting criteria of each DRG, to establish the relative price relationship between medical costs and weights of different disease groups; and then convert the points of each DRG group, using these points to allocate the regional health insurance. The point value is based on the final amount of the medical insurance fund and the total points declared by the designated medical institutions in the present year. Whether the PCP method based on DRG can exert a strong motivation on hospitals to improve their service performance is important for China is still an open topic.

The Chinese health commission divides medical institutions into three levels according to their tasks and functions to form a hierarchical medical system: primary, secondary, and tertiary hospitals.[22] The ability of hospitals to improve their performance is likely to have varied across different institution levels, which were important factors to identify the effect of the DRG payment policy. Chinese tertiary hospitals are the main cost control objects, whose insurance expenditures account for the bulk of medical insurance funds. Previous studies have analyzed them as subgroups in healthcare, but few have done so regarding the assessment of the DRG on service performance across hospitals. Given that Chinese healthcare is provided almost exclusively by state-owned public general hospitals,[23] it is of great importance to explore whether there are differences across subgroups in terms of the effects.[24] Wenzhou, one of cities in Zhejiang Province, has a similar DRG policy implementation process and healthcare system structure to other Chinese regions; therefore, it can act as a representative for estimating the effects of the PCP on public general hospitals. Considering the above, this study is aimed at analyzing performance differences when responding to the DRG' PCP from the perspective of institution levels, and taking the reaction time into account.

Material and methods

Data source and sample size

In recent years, the comprehensive medical-service strength of Wenzhou has been continuously enhanced.[25] Following the instructions set for Zhejiang Province, Wenzhou performed upfront preparations for DRG payment method with effect from May 20, 2019. Since then, more than 21 related symposia have been presented among hospitals in the region. On December 13, 2019, the Wenzhou local government held a publicity meeting on the payment policy and issued a detailed draft of the rules, after which it began communicating with hospitals to optimize the policy. On December 27, 2019, the DRG Points Counting Payment Implementation Rules was issued, and scheduled for implementation in that region in January 2020.[26] The timeline for 2019 for the implementation of the DRG payment policy is shown in Supplemental Figure 1.

First, in this region, the hospitals established DRG work teams, which could negotiate with the Healthcare Security Administration to safeguard the hospitals' interests. Such teams were generally composed of the dean, who oversaw the DRG' performance, the medical insurance department, and medical record quality-control department managers. Second, interior DRG data-monitoring information systems were introduced to strengthen hospital-wide disease groups monitoring, including monitoring of profit and loss and key indicators. Third, quality control of medical records was strengthened, especially the quality of code uploads, which was one of the prerequisites for obtaining medical insurance funds through DRG' payment.[3,27] Fourth, many hospitals launched training for clinicians across departments and specialized guidelines, which were also vital for the reform of the healthcare payment system, and helped to standardize diagnoses.[19,28]

Above-primary-level public general hospitals in Wenzhou amounted to 22 institutions, which consisted of 8 tertiary hospitals and 14 secondary hospitals. In 2021, the number of discharged cases in above-primary-level public general hospitals in Wenzhou was 840,342, accounting for 98.79% of the 850,652 discharged cases. In this study, the payment policy was implemented officially in 2020, the time-series data of these 22 institutions collected spanned from 2018 to 2021. Before implementation: 2018–2019; after implementation: 2020–2021. The data were collected from the Zhejiang Hospital Quality Management and Performance Evaluation Platform, which was developed by the Zhejiang Provincial Health Commission. The platform launched in 2017 and, to ensure high-quality data, includes logic verification and a key quality control index. The platform allows administrative departments and medical institutions to query DRG' performance data. In 2021, it evaluated 424 above-primary-level hospitals in Zhejiang province. This high level of data means this region can be considered a suitable representative for assessing the DRG policy's effect on all public general hospitals.

Indicators

This study used five performance indicators calculated by the Zhejiang Hospital Quality Management and Performance Evaluation Platform that reflects the medical capacity and consumption of medical resources. The five performance indicators are case-mix index (CMI), cost per hospitalization (CPH), its cost efficiency index (CEI), average length of stay (ALOS), and its time efficiency index (TEI). Among them, the CMI, CEI, and TEI derives from the DRG system, which has been found to have excellent homogeneous comparability.[29-32] The following are the definitions and calculations of CMI, CEI, and TEI as determined by Zhejiang Provincial Health Commission.

1) CMI reflects the comprehensive ability of a medical institution to treat the disease and its seriousness or difficulty of the diseases. The calculation method is

$$\text{CMI} = \frac{\text{Total weight}}{\text{Number of cases in region}}$$

where $\text{Total weight} = \sum \text{Each DRG weight} \times \text{Number of cases in each DRG}$

2) CEI reflects the cost of treating similar diseases in this region. When the value of CEI was close to 1, it suggested that, for the region in question, the cost efficiency was close to the average level; values below 1 or above 1 suggested that the efficiency was above or below the average level, respectively.

$$\text{CEI} = \frac{\sum_j k_j^c n_j}{\sum_j n_j} \qquad \text{Cost ratio}(k^c) = \frac{\text{average cost of a DRG in hospital } (c_i)}{\text{average cost of a DRG in region } (\overline{C_i})}$$

where n_j represents the number of cases in DRG_j, CEI is the weighted average of k^c.

3) TEI reflects the time spent treating similar diseases in this region. When the value of TEI was close to 1, it suggested that, for the region in question, the time efficiency was close to the average level; values below 1 or above 1 suggested that the efficiency was above or below the average level, respectively.

$$\text{TEI} = \frac{\sum_j k_j^l n_j}{\sum_j n_j} \qquad \text{Average length of stay ratio } (k^l) = \frac{\text{average length of a DRG in hospital } (l_i)}{\text{average length of a DRG in region } (\overline{L_i})}$$

where n_j represents the number of cases in DRG_j, TEI is the weighted average of k^l.

To reflect the differences in performances from the perspective of different institutional levels, we employed the mean of every indicator in total institutions, tertiary institutions, and secondary institutions, respectively.

Statistical analysis

Interrupted time series has become a core study design for the evaluation of health policies.[13] This study used the robust multiple interrupted time series (R-MITS) model, which have several advantages over the traditional interrupted time series designs.[33-35] R-MITS affords the identification of inferences regarding the estimation of global change points across units, rather than removing some time points as decided by the study team. Next, decreasing outcome variability can contribute to less extreme values, while increasing auto-correlation can contribute to a more predictable outcome. Therefore, it is important to capture alongside mean differences in variability and auto-correlation. Compared with traditional designs, this R-MITS model can capture all the aforementioned time series' changes. Furthermore, it does not assume that the intervention is instantaneous and allows the effect of the intervention to be abrupt or gradual, thus allowing the presence of a potentially lagged (or anticipatory) effect. It also can obtain more accurate estimates,[33] which further improves internal and external validity, and reduces bias.[36] Given these advantages, this robust model has been widely adopted to evaluate: improvements in care quality at both the system and unit level[36]; the effectiveness of a quality-improvement initiative for reducing cardiac arrests among infants and children[37]; the impact of trauma center accreditation on mortality and complications[38]; and the effectiveness of chest-pain units on observation times, diagnostic agreement, and emergency department re-admission and hospitalizations.[39]

The mean function of outcome for unit i at time t described in Cruz et al.[35] is

$$\mu_{it} = \begin{cases} \beta_{i0}^{\tau} + \beta_{i1}^{\tau} t, & t < \tau \\ (\beta_{i0}^{\tau} + \delta_i^{\tau}) + (\beta_{i1}^{\tau} + \Delta_i^{\tau}) t, & t \geq \tau \end{cases},$$

where β_{i0}^{τ} denotes the intercept of the mean function prior to the change-point, β_{i1}^{τ} denotes the slope of the outcome prior to the change-point, $\beta_{i0}^{\tau} + \delta_i^{\tau}$ is the intercept of the post-intervention phase, $\beta_{i1}^{\tau} + \Delta_i^{\tau}$ is the slope of the post-intervention phase for the outcome in unit i, and τ denotes the global over-all-unit change-point of the response. In the case with only one unit, τ denotes the change-point for that one-time series. If $\delta_i^{\tau} + \Delta_i^{\tau} = 0$, then there is no change in the mean function of unit i before and after τ.

Results

Descriptive analysis

The yearly results for the CMI mean were as follows: 1) For the full sample of institutions, it decreased from 0.799 to 0.785 before 2020 and increased from 0.834 to 0.859 after 2020. 2) Focusing only on tertiary institutions, it decreased from 0.868 to 0.864 before 2020 and increased from 0.886 to 0.904 after 2020. 3) Focusing only on secondary institutions, it decreased from 0.759 to 0.740 before 2020 and increased from 0.805 to 0.833 after 2020.

The yearly results for the CPH mean were as follows: 1) For the full sample of institutions, it increased from 8721.05 to 9484.77 yearly before and after 2020. 2) Focusing only on tertiary institutions, it decreased from 12658.60 to 12569.73 before 2020 and decreased from 13192.78 to 12951.49 after 2020. 3) Focusing only on secondary institutions, it increased from 6471.02 to 7503.78 yearly before and after 2020.

The yearly results for the CEI mean were as follows: 1) For the full sample of institutions, it decreased from 0.884 to 0.870 before 2020 and increased from 0.908 to 0.929 after 2020. 2) Focusing only on tertiary institutions, it decreased from 1.124 to 1.095 before 2020 and decreased from 1.110 to 1.108 after 2020. 3) Focusing only on secondary institutions, it decreased from 0.747 to 0.741 before 2020 and increased from 0.792 to 0.826 after 2020.

The yearly results for the ALOS mean were as follows: 1) For the full sample of institutions, it decreased from 8.27 to 8.03 before 2020 and decreased from 8.19 to 7.77 after 2020. 2) Focusing only on tertiary institutions, it decreased from 8.52 to 7.05 yearly before and after 2020. 3) Focusing only on secondary institutions, it decreased from 8.13 to 8.12 before 2020 and decreased from 8.47 to 8.18 after 2020.

The yearly results for the TEI mean were as follows: 1) For the full sample of institutions, it increased from 0.966 to 0.968 before 2020 and increased from 1.000 to 1.012 after 2020. 2) Focusing only on tertiary institutions, it decreased from 1.029 to 0.988 yearly before and after 2020. 3) Focusing only on secondary institutions, it increased from 0.931 to 1.026 before and after 2020.

The yearly outcomes for every target indicator for the period of 2018–2021 (representing pre- and post-implementation) are shown in Table 1.

Table 1. Unit-specific yearly results before and after the policy implementation

Unit	Year	CMI mean	CPH mean	CEI mean	ALOS mean	TEI mean
Total institutions	2018	0.799	8721.05	0.884	8.27	0.966
	2019	0.785	8775.51	0.870	8.03	0.968
	2020	0.834	9418.26	0.908	8.19	1.000
	2021	0.859	9484.77	0.929	7.77	1.012
Tertiary institutions	2018	0.868	12658.60	1.124	8.52	1.029
	2019	0.864	12569.73	1.095	7.87	1.008
	2020	0.886	13192.78	1.110	7.69	0.993
	2021	0.904	12951.49	1.108	7.05	0.988
Secondary institutions	2018	0.759	6471.02	0.747	8.13	0.931
	2019	0.740	6607.38	0.741	8.12	0.946
	2020	0.805	7261.39	0.792	8.47	1.004
	2021	0.833	7503.78	0.826	8.18	1.026

ALOS: average length of stay; CEI: cost efficiency index; CMI: case-mix index; CPH: cost per hospitalization; TEI: time efficiency index; before implementation: from 2018 to 2019; after implementation: from 2020 to 2021

R-MITS analysis

R-MITS results for the CMI mean were as follows: 1) For the full sample of institutions, the effect change-point was estimated for March 2020 ($p<0.0001$) with the slope increased from −0.0001949 to 0.001267 ($p<0.0001$). 2) Focusing only on tertiary institutions, the effect change-point was estimated February 2020 ($p<0.05$) with the slope changing from −0.0001065 to 0.0008474 ($p<0.01$). 3) Focusing only on secondary institutions, the effect change-point was estimated for January 2020 ($p<0.0001$) with the slope changing from −0.001213 to 0.001924 ($p<0.0001$).

R-MITS results for the CPH mean were as follows: 1) For the full sample of institutions, the effect change-point was estimated for January 2020 ($p<0.001$) with the slope changed from −6.442 to −5.286 ($p>0.05$). 2) Focusing only on tertiary institutions, the effect change-point was estimated January 2020 ($p<0.01$) with the slope changing from −18.71 to −43.31 ($p>0.05$). 3) Focusing only on secondary institutions, the effect change-point was estimated for April 2020 ($p<0.001$) with the slope changing from 8.554 to −0.8039 ($p>0.05$).

R-MITS results for the CEI mean were as follows: 1) For the full sample of institutions, the effect change-point was estimated for June 2019 ($p<0.0001$) with the slope changing from 0.001073 to 0.002386 ($p<0.01$). 2) Focusing only on tertiary institutions, the effect change-point was estimated for June 2019 ($p<0.0001$) with the slope changing from 0.002008 to 0.0004813 ($p<0.05$). 3) Focusing only on secondary institutions, the effect change-point was estimated for April 2020 ($p<0.0001$) with the slope changing from 0.0002849 to 0.001513 ($p<0.05$).

R-MITS results for the ALOS mean were as follows: 1) For the full sample of institutions, the effect change-point was estimated for January 2020 ($p<0.0001$) with the slope changing from −0.02019 to −0.03918 ($p<0.05$). 2) Focusing only on tertiary institutions, the effect change-point was estimated for January 2020 ($p<0.05$) with the slope changing from −0.04884 to −0.06181 ($p>0.05$). 3) Focusing only on secondary institutions, the effect change-point was estimated for January 2020 ($p<0.0001$) with the slope changing from −0.004108 to −0.02610 ($p<0.01$).

R-MITS results for the TEI mean were as follows: 1) For the full sample of institutions, the effect change-point was estimated for April 2020 ($p<0.001$) with the slope changing from 0.0006056 to −0.0004526 ($p>0.05$). 2) Focusing only on tertiary institutions, the effect change-point was estimated for June 2020 ($p>0.05$) with the slope changing from −0.001041 to −0.0004987 ($p>0.05$). 3) Focusing only on secondary institutions, the effect change-point was estimated for April 2020 ($p<0.0001$) with the slope changing from 0.001591 to −0.000257 ($p<0.05$).

Unit-specific estimated changes using R-MITS are shown in Table 2. The fitted regression lines are shown in Figures 1 through 3.The estimates of the stochastic component parameters are shown in Supplemental Table 1. The residuals and auto-correlation functions are shown in Supplemental Figures 2 through 6.

Table 2. Unit-specific changes using R-MITS before and after the DRG policy implementation

Outcomes	Unit	Change-points		Changes in slope			Changes in intercept
		Estimate	Lag	Pre-change (β_{i1}^{τ})	Post-change ($\beta_{i1}^{\tau}+\Delta_{i}^{\tau}$)	Change (Δ_{i}^{τ})	(δ_{i}^{τ})
CMI mean	Total institutions	Mar-2020****	2 months after	−0.0001949	0.001267	0.001462****	0.009988
	Tertiary institutions	Feb-2020*	1 months after	−0.0001065	0.0008474	0.0009540**	0.001480
	Secondary institutions	Jan-2020****	-	−0.001213	0.001924	0.003137****	−0.009853
CPH mean	Total institutions	Jan-2020***	-	−6.442	−5.286	1.156	909.9**
	Tertiary institutions	Jan-2020**	-	−18.71	−43.31	−24.60	1957***
	Secondary institutions	Apr-2020***	3 months after	8.554	−0.8039	−9.358	1075***
CEI mean	Total institutions	Jun-2019****	6 months before	0.001073	0.002386	0.001313**	−0.05045**
	Tertiary institutions	Jun-2019****	6 months before	0.002008	0.0004813	−0.001527*	−0.02375
	Secondary institutions	Apr-2020****	3 months after	0.0002849	0.001513	0.001229*	0.01154
ALOS mean	Total institutions	Jan-2020****	-	−0.02019	−0.03918	−0.01899*	1.071***
	Tertiary institutions	Jan-2020*	-	−0.04884	−0.06181	−0.01297	0.9115*
	Secondary institutions	Jan-2020****	-	−0.004108	−0.02610	−0.02199**	1.154***
TEI mean	Total institutions	Apr-2020***	3 months after	0.0006056	−0.0004526	−0.001058	0.06593**
	Tertiary institutions	Jun-2020	5 months after	−0.001041	−0.0004987	0.0005421	−0.02309
	Secondary institutions	Apr-2020****	3 months after	0.001591	−0.0002571	−0.001848*	0.1102***

Significance level α=0.05; *, p<0.05, **, p<0.01; ***, p<0.001; ****, p<0.0001; ALOS: average length of stay; CEI: cost efficiency index; CMI: case-mix index; CPH: cost per hospitalization; TEI: time efficiency index

Discussion

Effects on the CMI after the DRG implementation

CMI can reflect the comprehensive ability of a medical institution. As for all the institutions, this study found that the CMI mean showed a downward curve before the change but increased after the change, from which it can be inferred that the DRG policy implementation likely motivated public general hospitals to devote themselves to improving CMI. According to the payment rules, improving CMI can contribute to securing more DRG points, thereby acquiring more medical insurance funds. Additionally, to encourage medical institutions to implement new technologies and enhance service capabilities, policymakers rewarded institutions that adopted new technologies with additional DRG points.[26] The study indicated that CMI change lagged two months after the DRG policy officially implemented, which may result from the difficulty in strengthening the comprehensive ability immediately even if they have stronger motivation than before.

First, focusing only on changes at institutional levels, after the policy was implemented, the results showed an obvious trend from downward to upward in both tertiary and secondary institutions. Second, as observed, the CMI mean in tertiary institutions remained higher than in secondary institutions, regardless of implementation. Because irrespective of improvements on the medical technical level, the proportion of high relative weight and the number of treating cases, tertiary institutions have innate advantages over secondary institutions. Third, this study found that the CMI mean in tertiary institutions changed later than in secondary institutions but also improved slower than in secondary institutions after change. As mentioned above, CMI represents the hospitals' average technical levels and the seriousness of cases treated in each discipline, which have already reached high levels in tertiary institutions. Thus, improving CMI would require more effort in secondary institutions. Another possible reason is that, due to Chinese citizens being at liberty to choose healthcare services at any institution, tertiary institutions with higher levels of comprehensive medical strength can attract more patients under the siphon effect.[40,41] With service accessibility improving, patients would rather visit senior hospitals, even though most of their illnesses are not severe.[42,43] As a result, it is common for a large number of low-related weight cases are treated at tertiary institutions that constrained improving CMI, and also affected the hierarchic healthcare treatment system performance for a long time. Significantly, under the DRG' payment rules, tertiary institutions tended to treat more serious diseases (that possess high relative weights) and were willing to transfer the more common or easy-to-treat cases to lower-grade institutions.[19] Subsequently, the proportion of the high relative weight of their treating diseases can be increased, which contributed to improving their CMI. In a sense, the implementation of the DRG policy can promote the hierarchic healthcare treatment system performing its intended role in China.

Effects on the CPH and CEI after the DRG implementation

This study found that the change in the slope of the CPH mean was not significant in all institutional levels. As for the CEI, its mean was still lower than 1 before and after change, which can suggest that the CEI of these institutions were stronger than the average in the region. Besides, the decrease of intercept can infer that the policy has a positive motivation for hospitals to strengthen cost control. Importantly, after the policy was implemented, the CEI means increased at a faster rate toward 1, from which it can be inferred that the policy can further motivate these institutions to narrow the differences across public general hospitals in terms of costs for treating similar diseases. Because, the basic billing rules of DRG' payment method in this region is that "excess expenditures will no longer be reimbursed and the balance will be retained."[28] Under this rule, institutions that implement higher medical costs for treating a disease group will lose money, and thus hospitals have no choice but to strengthen their cost control. Further, the policy can motivate hospitals to improve the correct coding rate and integrity of DRG' data, which can then align the cost for treating diseases with diseases' reality,[19] and subsequently bring the CEI close to 1. Then, in this study, the results showed that the CEI mean changed in June 2019, following the commencement of the policy reform in May 2019. As is known, from the postpaid system of Fee-for-Service to the prepaid system of DRGs payment, the level of hospitalization expenses was directly related to payment systems.[7] The introduction of the policy could therefore increase cost control consciousness and motivate associated management behaviors.[3,44] Therefore, under the policy influence, the response of these institutions to strengthening the CEI can be faster than other indicators.

Focusing only on changes at institutional levels, the CEI mean in tertiary institutions remained higher than 1, while remaining lower than 1 in secondary institutions before and after change. As is known, the CPH in tertiary institutions are higher than secondary institutions; similarly, the phenomenon of spending high costs on similar diseases was more common in tertiary institutions, which resulted in their high CEI. Because in tertiary institutions, medical equipment, technical levels, drugs, and consumables were more advanced, they attract higher expenses.[15] Significantly, under the payment rules of DRG, ultra-high-cost inpatient cases will cause financial losses for hospitals, and this can motivate hospitals to positively pursue cost-control behaviors, thereby causing the CEI to increase at slower rates, and even decrease. This study showed that after the policy was implemented, the uptrend of the CEI mean in tertiary institutions was slower than before, and the CEI mean in secondary institutions changed faster toward 1. It can be inferred that the difference in the tertiary and secondary institutions was mitigated by the DRG policy implementation. Then, this also found that tertiary institutions reacted faster than secondary institutions to the DRGs payment policy in terms of CEI. It is worth mentioning that, as in most parts of China, the policymakers also set a certain payment coefficient in the initial implementation of DRG, taking the actual conditions in the institutions and differences in treatment costs, into account.[26] In that context, some hospitals may believe that briefly implementing high costs can improve the historical base data for the future, despite leading to an immediate short-term loss.

Effects on the ALOS and TEI after the DRG implementation

A decrease in the length of stay can be either positive or negative for individual patients and for the health system as a whole.[45] For example, an outcome of shortening the length of stay may result from the early discharge.[19] It may also be the result in the DRG policymakers in this region strengthening monitoring, and introducing penalties related to the length of stay.[21,26] Certainly, shortening ALOS or strengthening TEI can be reasonable, but it requires the prior improvement of medical and care levels across public general hospitals. After the DRG policy was implemented, we found that the intercept of the ALOS mean in total institutions increased and the downward trend changed faster than before. Focusing only on changes at institutional levels, the downtrend of the ALOS mean and the TEI mean in tertiary institutions did not change significantly, while in secondary institutions it did. Compared with secondary institutions, tertiary institutions with stronger medical service capacity, whose inpatients with similar diseases can be treated better and therefor ALOS would be shorter. As a result, tertiary institutions require much more efforts to further decline the ALOS or TEI in a short time. As is mentioned above, after the policy was implemented, the CMI mean in tertiary institutions increased, from which it can be inferred that the percentage of serious or complex diseases treated in tertiary institutions have improved. Due to the high improvement space, it may be reasonable for secondary institutions to further shorten ALOS and strengthen TEI.

Limitations

This study has some limitations. First, as the sample of this study only focused on the public general hospitals, it is necessary to include other hospital types to form a comprehensive evaluation of the impact of the DRG' payment policy on inpatient service performance. Second, the COVID-19 pandemic had an unavoidable impact on the operation of hospitals worldwide. Importantly, although the epidemic reached Zhejiang Province at the end of January, the setback was temporary due to the implementation of strict prevention and control measures, and recovery gradually occurred after February 2020. According to the purpose of this study, the selected indicators reflect the changes of comprehensive medical strength and efficiency in hospitals not the operation indicators, thus, the possible disturbance resulting from the pandemic can be ignored. Despite the inevitable impact of the pandemic on hospital services, this study succeeded in revealing the anticipatory and positive effects of DRG policy on hospitals' performance, used a R-MITS analysis that further improved internal validity, while reducing bias. Third, although the assessment of healthcare policies is complex due to interacting and interdependent components, we found no other important changes in health policy or medical insurance policy in this region when the payment policy was implemented. It can be considered that the changes in the monitored indicators originated from the DRG' payment reform.

Conclusions

This study found that it was effective in motivating public general hospitals to improve medical capacity in terms of CMI, shorten the ALOS, and strengthen the CEI that suggested the policy can expedite the mitigation of discrepancies in the expense of treating the same types of diseases across institutions. It did not cause a change of CPH and TEI significantly in the full sample of public general hospitals. There were differences in the performance between tertiary and secondary hospitals when responding to DRG policy. Tertiary institutions react faster than secondary institutions in terms of cost control and improve their medical capacity slower than secondary institutions in terms of CMI. The gap between the cost utilization efficiency of secondary and tertiary hospitals was decreasing, which means that the cost of treating the similar diseases in hospitals of different tiers tended to be consistent. The policy did not motivate tertiary institutions to shorten the ALOS and TEI, while secondary hospitals had shown continuous improvement in time utilization efficiency.

Acknowledgments

The authors would like to thank the support of the Hospital Quality Management and Performance Evaluation Platform from Zhejiang Provincial Health Commission.

Contributors

TZ and CC were responsible for conceptualization, data collection, methodology, resources, and software. XZ was responsible for conceptualization, data collection, and methodology. QY was responsible for methodology and software. YH was responsible for data collection and descriptive analysis. RL was responsible for data collection and investigation. XZ and YD were responsible for conceptualization, funding acquisition, resources, supervision, reviewing, and editing of the manuscript. All authors contributed to the critical revision of the manuscript for important intellectual content, review, and approval of the final manuscript.

Funding

This research was sponsored by Zhejiang Provincial Medical and Health Science Technology Plan Project (2023KY889) from Zhejiang Provincial Health Commission in China.

Competing interests

None declared.

Patient and public involvement

Patients and/or the public were not involved in the design, conduct, reporting, or dissemination plans of this research

Patient consent for publication

Not required.

Informed Consent Statement

Not applicable.

Data availability statement

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

References

1. Stadhouders N, Kruse F, Tanke M, Koolman X, Jeurissen P. Effective healthcare cost-containment policies: a systematic review. *Health Policy*. Jan 2019;123(1):71–79. doi:10.1016/j.healthpol.2018.10.015
2. Kim YL, K.-H.; Choi, S.W. Multifaced evidence of hospital performance in Pennsylvania. *Healthcare*. 2021;9:670. doi:10.3390/healthcare9060670
3. Annear PL, Kwon S, Lorenzoni L, et al. Pathways to DRG-based hospital payment systems in Japan, Korea, and Thailand. *Health Policy*. Jul 2018;122(7):707–713. doi:10.1016/j.healthpol.2018.04.013

4. Busse R, Geissler A, Aaviksoo A, et al. Diagnosis related groups in Europe: moving towards transparency, efficiency, and quality in hospitals? *BMJ*. Jun 7 2013;346:f3197. doi:10.1136/bmj.f3197
5. Chok L, Bachli EB, Steiger P, et al. Effect of diagnosis related groups implementation on the intensive care unit of a Swiss tertiary hospital: a cohort study. *BMC Health Serv Res*. Feb 5 2018;18(1):84. doi:10.1186/s12913-018-2869-4
6. Zou K, Li HY, Zhou D, Liao ZJ. The effects of diagnosis-related groups payment on hospital healthcare in China: a systematic review. *BMC Health Serv Res*. Feb 12 2020;20(1):112. doi:10.1186/s12913-020-4957-5
7. Mathauer I, Wittenbecher F. Hospital payment systems based on diagnosis-related groups: experiences in low- and middle-income countries. *Bull World Health Organ*. Oct 1 2013;91(10):746–756A. doi:10.2471/BLT.12.115931
8. Bush RW. Reducing waste in US health care systems. *JAMA*. Feb 28 2007;297(8):871–4. doi:10.1001/jama.297.8.871
9. Cui Y, Wu Z, Lu Y, Jin W, Dai X, Bai J. Effects of the performance management information system in improving performance: an empirical study in Shanghai Ninth People's Hospital. *Springerplus*. 2016;5(1):1785. doi:10.1186/s40064-016-3436-2
10. Council CS. Guidance on further deepening the basic medical insurance payment reform http://www.gov.cn/gongbao/content/2017/content_5210497.htm (accessed 8 Feb 2023)
11. McClellan M. Reforming payments to healthcare providers: the key to slowing healthcare cost growth while improving quality? *J Econ Perspect*. Spring 2011;25(2):69–92. doi:10.1257/jep.25.2.69
12. Bystrov V, Staszewska-Bystrova A, Rutkowski D, Hermanowski T. Effects of DRG-based hospital payment in Poland on treatment of patients with stroke. *Health Policy*. Aug 2015;119(8):1119–25. doi:10.1016/j.healthpol.2015.04.017
13. Lopez Bernal J, Soumerai S, Gasparrini A. A methodological framework for model selection in interrupted time series studies. *J Clin Epidemiol*. Nov 2018;103:82–91. doi:10.1016/j.jclinepi.2018.05.026
14. Administration NHS. Notice on the lists of national pilot cities for diagnosis related grouping payment. http://www.gov.cn/zhengce/zhengceku/2019-10/12/content_5438769.htm (accessed 8 Feb 2023)
15. Bredenkamp C, Bales, S., & Kahur, K. *Transition to diagnosis-related group (DRG) payments for health lessons from case studies*. 2020.
16. Zhao S, Gu Y, Huang Z, Zeng W. Building a performance management system for hospitals based on diagnosis-related group (DRG) payment. *Journal of Sensors*. 2022;2022:1-10. doi:10.1155/2022/7001423
17. Jian W, Lu M, Han W, Hu M. Introducing diagnosis-related groups: is the information system ready? *Int J Health Plann Manage*. Jan-Mar 2016;31(1):E58-68. doi:10.1002/hpm.2270
18. Garcia Calderón V, Figueiras Huante IA, Carbajal Martinez M, et al. The impact of improving the quality of coding in the utilities of diagnosis related groups system in a private healthcare institution. 14-year experience. *Int J Med Inform*. Sep 2019;129:248-252. doi:10.1016/j.ijmedinf.2019.06.019
19. Barouni M, Ahmadian L, Anari HS, Mohsenbeigi E. Challenges and adverse outcomes of implementing reimbursement mechanisms based on the diagnosis-related group classification system: a systematic review. *Sultan Qaboos Univ Med J*. Aug 2020;20(3):e260–e270. doi:10.18295/squmj.2020.20.03.004
20. Jian W, Lu M, Liu G, Chan KY, Poon AN. Beijing's diagnosis-related group payment reform pilot: impact on quality of acute myocardial infarction care. *Soc Sci Med*. Dec 2019;243:112590. doi:10.1016/j.socscimed.2019.112590
21. Administration ZPHS. Notice on the issuance of the interim rules for payment of DRGs points for inpatient costs of basic medical insurance in Zhejiang province. http://ybj.zj.gov.cn/art/2019/11/27/art_1229113757_564142.html (accessed 8 Feb 2023)
22. Dong E, Liu S, Chen M, et al. Differences in regional distribution and inequality in health-resource allocation at hospital and primary health centre levels: a longitudinal study in Shanghai, China. *BMJ Open*. Jul 19 2020;10(7):e035635. doi:10.1136/bmjopen-2019-035635

23. Zhao X, Xiao J, Chen H, et al. Patient preferences and attitudes towards first choice medical services in Shenzhen, China: a cross-sectional study. *BMJ Open*. May 24 2022;12(5):e057280. doi:10.1136/bmjopen-2021-057280

24. Meng Z, Hui W, Cai Y, Liu J, Wu H. The effects of DRGs-based payment compared with cost-based payment on inpatient healthcare utilization: a systematic review and meta-analysis. *Health Policy*. Apr 2020;124(4):359-367. doi:10.1016/j.healthpol.2020.01.007

25. Government WMPs. Notice on the fourteenth five-year plan for Wenzhou healthcare development. http://www.wenzhou.gov.cn/art/2021/12/1/art_1229116916_1911427.html (accessed 8 Feb 2023)

26. Administration WHS. Notice on Wenzhou basic medical insurance total hospitalization cost budget and DRGs point payment implementation rules. http://ybj.wenzhou.gov.cn/art/2019/12/30/art_1229135972_623703.html (accessed 8 Feb 2023)

27. Medarevic AP. Describing Serbian hospital activity using Australian refined diagnosis related groups: a case study in Vojvodina province. *Zdr Varst*. Mar 2020;59(1):18-26. doi:10.2478/sjph-2020-0003

28. Yu L, Lang J. Diagnosis-related groups (DRG) pricing and payment policy in China: where are we? *Hepatobiliary Surg Nutr*. Dec 2020;9(6):771–773. doi:10.21037/hbsn-2020-8

29. Ding S, Jiang X, Zheng Y, et al. Improving the clinical ability and quality of endocrinology department with diagnosis-related groups tool. *Ann Transl Med*. Feb 2022;10(4):167. doi:10.21037/atm-22-147

30. Feng L, Tian Y, He M, et al. Impact of DRGs-based inpatient service management on the performance of regional inpatient services in Shanghai, China: an interrupted time series study, 2013–2019. *BMC Health Serv Res*. Oct 12 2020;20(1):942. doi:10.1186/s12913-020-05790-6

31. Jian W, Huang Y, Hu M, Zhang X. Performance evaluation of inpatient service in Beijing: a horizontal comparison with risk adjustment based on Diagnosis Related Groups. *BMC Health Serv Res*. Apr 30 2009;9:72. doi:10.1186/1472-6963-9-72

32. Liu X, Liu F, Wang L, Wu M, Yang L, Wei L. Performance evaluation of medical service for breast cancer patients based on diagnosis related groups. *BMC Health Serv Res*. May 24 2021;21(1):496. doi:10.1186/s12913-021-06439-8

33. Cruz M, Bender M, Ombao H. A robust interrupted time series model for analyzing complex health care intervention data. *Statistics in medicine*. Dec 20 2017;36(29):4660-4676. doi:10.1002/sim.7443

34. Cruz M, Gillen DL, Bender M, Ombao H. Assessing health care interventions via an interrupted time series model: Study power and design considerations. *Statistics in medicine*. May 10 2019;38(10):1734-1752. doi:10.1002/sim.8067

35. Cruz M, Pinto-Orellana MA, Gillen DL, Ombao HC. RITS: a toolbox for assessing complex interventions via interrupted time series models. *BMC Med Res Methodol*. Jul 8 2021;21(1):143. doi:10.1186/s12874-021-01322-w

36. Bender M, Murphy EA, Cruz M, Ombao H. System- and unit-level care quality outcome improvements after integrating clinical nurse leaders into frontline care delivery. *J Nurs Adm*. Jun 2019;49(6):315–322. doi:10.1097/NNA.0000000000000759

37. Ferguson LP, Thiru Y, Staffa SJ, Guillen Ortega M. Reducing cardiac arrests in the PICU: initiative to improve time to administration of prearrest bolus epinephrine in patients with cardiac disease. *Crit Care Med*. Jul 2020;48(7):e542–e549. doi:10.1097/CCM.0000000000004349

38. Batomen B, Moore L, Strumpf E, Champion H, Nandi A. Impact of trauma centre accreditation on mortality and complications in a Canadian trauma system: an interrupted time series analysis. *BMJ Qual Saf*. Nov 2021;30(11):853–866. doi:10.1136/bmjqs-2020-011271

39. Alzate-Gutiérrez N, Roldán-Toro H, Gallo-Villegas J, Aristizábal-Ocampo D. Efectividad de una unidad de dolor torácico en el primer nivel de atención en salud con telecardiología: un estudio cuasiexperimental de series de tiempo interrumpidas. *Revista Colombiana de Cardiología*. 2022;28(4)doi:10.24875/rccar.M21000062

40. Jiang Q, Pan J. The evolving hospital market in China after the 2009 healthcare reform. *Inquiry*. Jan-Dec 2020;57:46958020968783. doi:10.1177/0046958020968783

41. Yan K, Yang C, Zhang H, et al. Impact of the zero-mark-up drug policy on drug-related expenditures and use in public hospitals, 2016–2018: an interrupted time series study in Shaanxi. *BMJ Open*. Nov 26 2020;10(11):e037034. doi:10.1136/bmjopen-2020-037034

42. Chu Y, Tao H. Detecting undifferentiation of tertiary and county hospitals in China in adoption of DRG instrument. *Healthcare (Basel)*. Jul 21 2021;9(8)doi:10.3390/healthcare9080922
43. Yu W, Li M, Nong X, et al. Practices and attitudes of doctors and patients to downward referral in Shanghai, China. *BMJ Open*. Apr 3 2017;7(4):e012565. doi:10.1136/bmjopen-2016-012565
44. Choi JW, Kim SJ, Park HK, Jang SI, Kim TH, Park EC. Effects of a mandatory DRG payment system in South Korea: analysis of multi-year nationwide hospital claims data. *BMC Health Serv Res*. Oct 30 2019;19(1):776. doi:10.1186/s12913-019-4650-8
45. Kone I, Maria Zimmermann B, Nordstrom K, Simone Elger B, Wangmo T. A scoping review of empirical evidence on the impacts of the DRG introduction in Germany and Switzerland. *Int J Health Plann Manage*. Jan 2019;34(1):56–70. doi:10.1002/hpm.2669

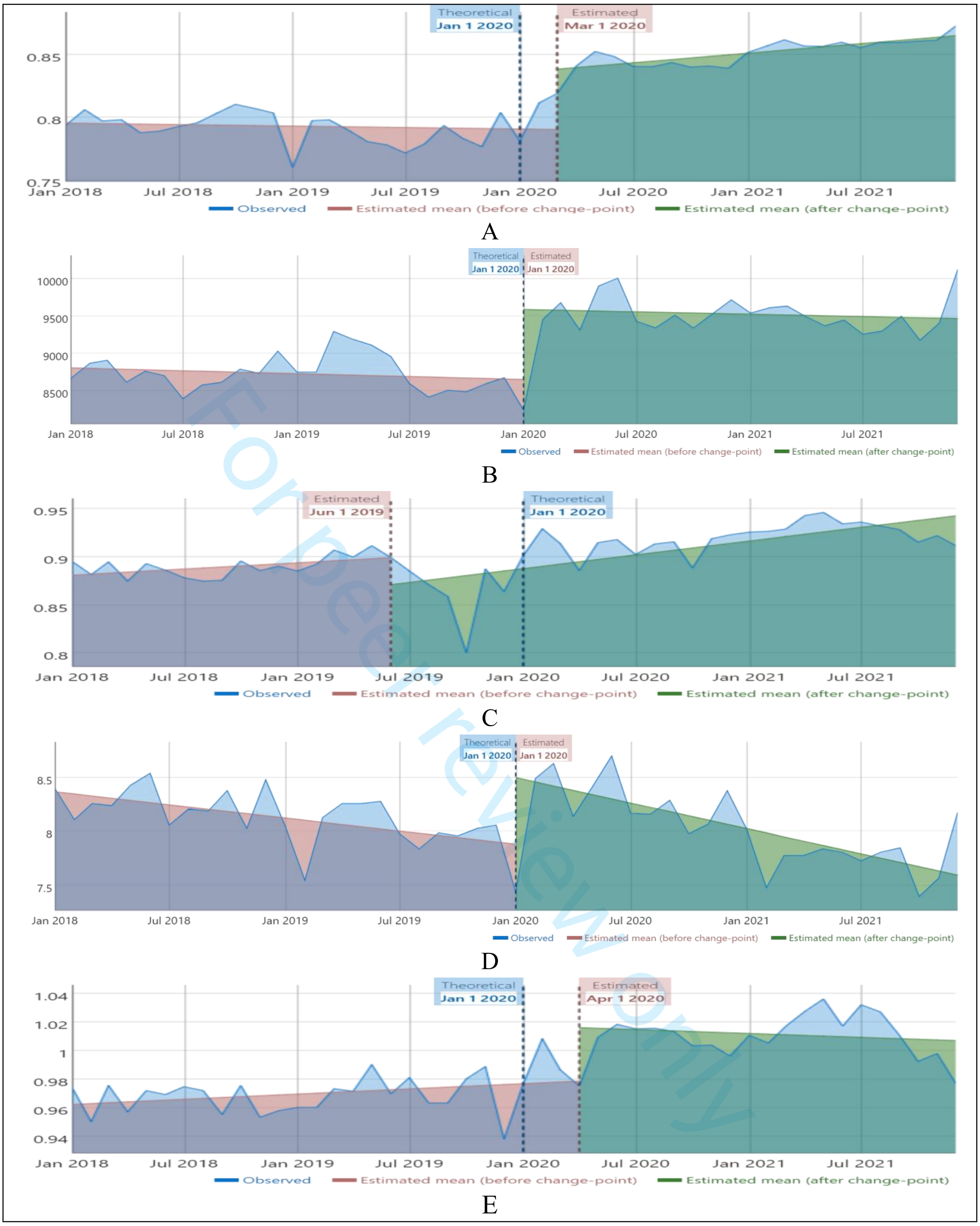


Figure 1. The fitted regression lines for total institutions before and after change-points.

(A) CMI mean; (B) CPH mean; (C) CEI mean; (D) ALOS mean; (E) TEI mean

ALOS: average length of stay; CEI: cost efficiency index; CMI: case-mix index; CPH: cost per hospitalization; TEI: time efficiency index

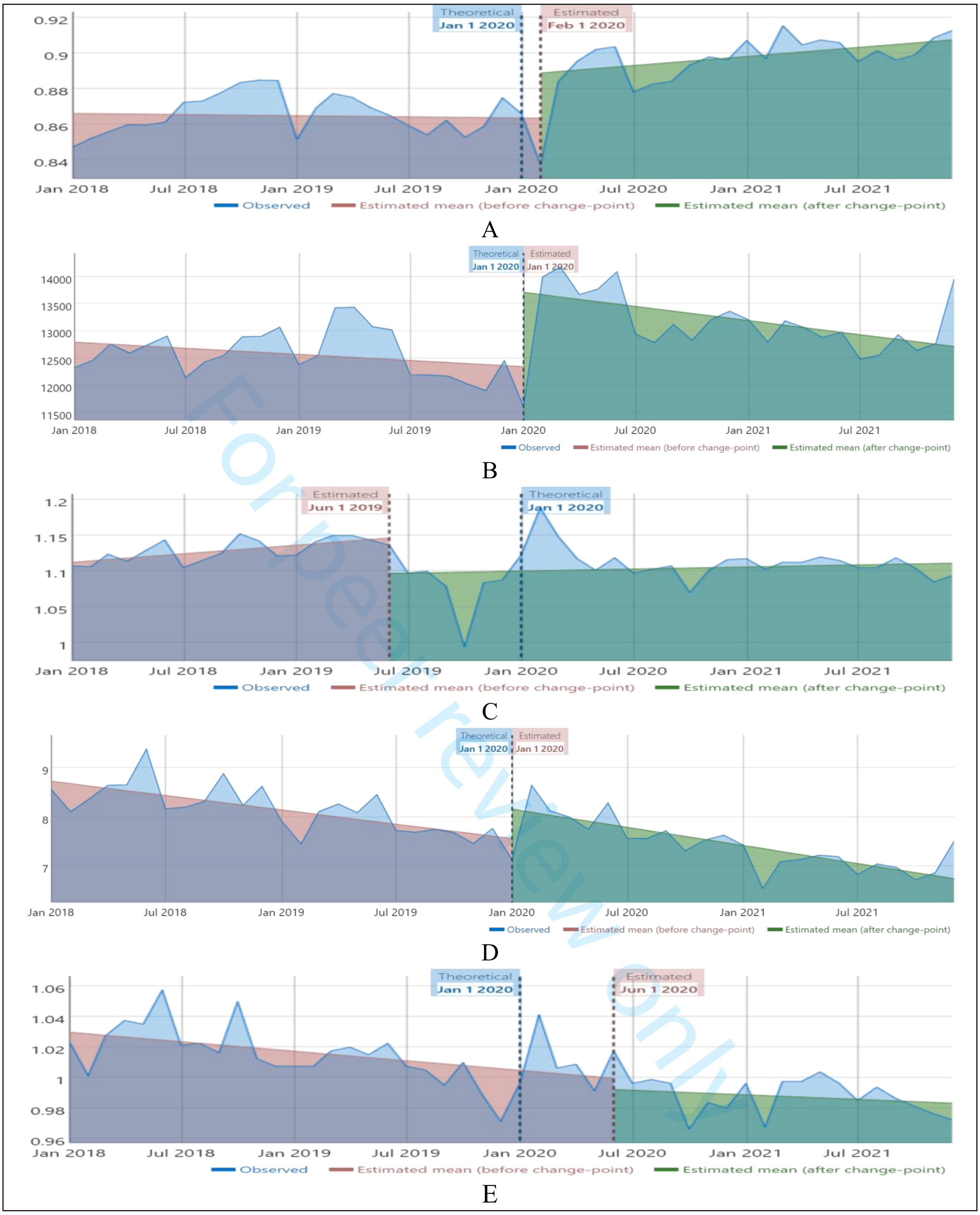


Figure 2. The fitted regression lines for tertiary institutions before and after change-points.

(A) CMI mean; (B) CPH mean; (C) CEI mean; (D) ALOS mean; (E) TEI mean

ALOS: average length of stay; CEI: cost efficiency index; CMI: case-mix index; CPH: cost per hospitalization; TEI: time efficiency index

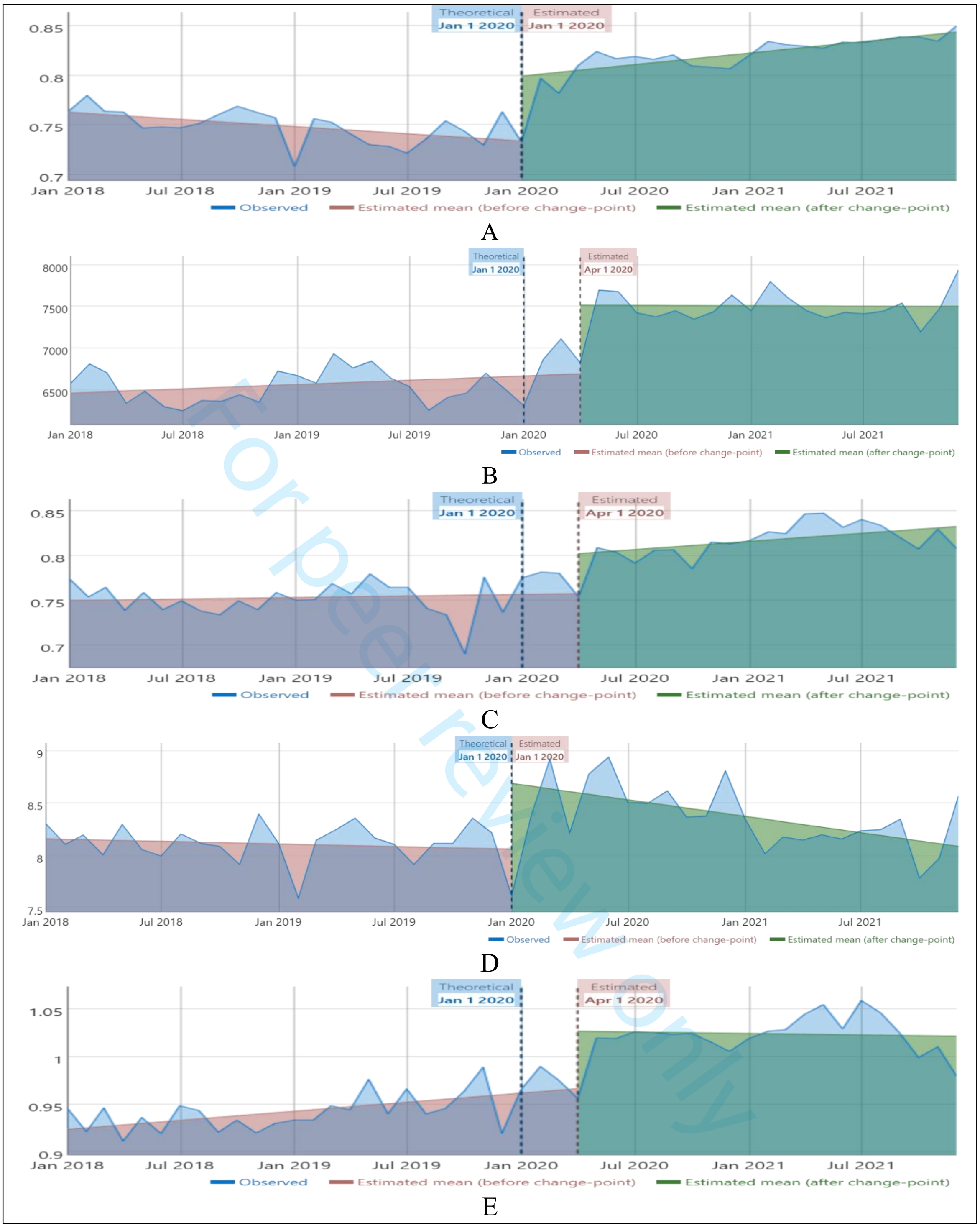


Figure 3. The fitted regression lines for secondary institutions before and after change-points.

(A) CMI mean; (B) CPH mean; (C) CEI mean; (D) ALOS mean; (E) TEI mean
ALOS: average length of stay; CEI: cost efficiency index; CMI: case-mix index; CPH: cost per hospitalization; TEI: time efficiency index

Research Checklist

Getting Organized	Due Date
1. Identify all important elements from the directions.	5 July 2022
2. Set a realistic timetable using this checklist, or some other method.	20 July 2022
Choosing a topic	
3. Find an appropriate topic.	5 Aug 2022
4. Pre-write and do some pre-research.	15 Aug 2022
5. Reconsider or modify your topic if necessary.	30 Aug 2022
6. Make a tentative outline.	15 Sep 2022
Writing a Thesis Statement	
7. Turn your topic into a guiding question you want to answer.	25 Sep 2022
8. Write a thesis statement that answers your guiding question.	5 Oct 2022
Starting Research	
9. Gather your sources.	5 Nov 2022
10. Start a running bibliography list.	15 Nov 2022
11. Take notes.	30 Nov 2022
12. Read and understand the section on plagiarism.	5 Dec 2022
13. Write a more formal outline.	15 Dec 2022
Start Writing the First Draft	
14. Write a strong introduction.	31 Dec 2022
15. Write the body paragraphs.	15 Jan 2023
16. Write a conclusion.	31 Jan 2023
Revise, Edit, and Proofread	
17. Re-read your paper and decide if it needs re-organizing.	20 Feb 2023
18. Re-write it in a final form.	5 Mar 2023
19. Complete a Works Cited or References page.	10 Mar 2023
20. Check for errors: grammar, punctuation, spelling, and capitalization.	15 Mar 2023

BMJ Open

Differences in inpatient performance of public general hospitals following implementation of a points-counting payment based on diagnosis-related groups: a robust multiple interrupted time series study in Wenzhou, China

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2023-073913.R1
Article Type:	Original research
Date Submitted by the Author:	02-Nov-2023
Complete List of Authors:	Zhu, Tingting; Wenzhou Medical University, School of Public Health and Management Chen, Chun; Wenzhou Medical University, School of Public Health and Management Zhang, Xinxin; Wenzhou Medical University, School of Public Health and Management Yang, Qingren; Wenzhou Medical University, School of Public Health and Management Hu, Yipao; Health Commission of Wenzhou, Health Information Center Liu, Ruoyun; Wenzhou Medical University, School of Public Health and Management Zhang, Xiangyang; Wenzhou Medical University First Affiliated Hospital Dong, Yin; Health Community Group of Yuhuan People’s Hospital
Primary Subject Heading:	Health policy
Secondary Subject Heading:	Medical management, Health services research
Keywords:	Health policy < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Organisation of health services < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Health Services, Public Hospitals < Hospitals, Public, Hospitalization

SCHOLARONE™
Manuscripts

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Differences in inpatient performance of public general hospitals following implementation of a points-counting payment based on diagnosis-related groups: a robust multiple interrupted time series study in Wenzhou, China

Abstract

Objectives This study measures the differences in inpatient performance after a points-counting payment policy based on diagnosis-related groups (DRGs) was implemented. The point value is dynamic; its change depends on the annual DRGs' cost settlements and points of the current year, which are calculated at the beginning of the following year.

Design A longitudinal study using a robust multiple interrupted time-series model to evaluate service performance following policy implementation.

Setting Twenty-two public general hospitals (eight tertiary institutions and 14 secondary institutions) in Wenzhou, China.

Intervention The intervention was implemented in January 2020.

Outcome measures Indicators were case-mix index (CMI), cost per hospitalization (CPH), average length of stay (ALOS), cost efficiency index (CEI), and time efficiency index (TEI). The outcomes employed the means of these indicators.

Results The impact of COVID-19, which reached Zhejiang province at the end of January 2020, was temporary, given rapid containment following strict control measures. After the intervention, except for ALOS mean, the change-points of other outcomes ($p<0.05$) in tertiary and secondary institutions were inconsistent. The CMI mean turned to uptrend in tertiary ($p<0.01$) and secondary institutions ($p<0.0001$) than before. Although the slope of CPH mean did not change ($p>0.05$), the uptrend of CEI mean in tertiary institutions alleviated ($p<0.05$) and furthered increase ($p<0.05$) in secondary institutions. The slopes of ALOS and TEI mean in secondary institutions changed ($p<0.05$) but did not change in tertiary institutions ($p>0.05$).

Conclusions This study showed a positive effect of the DRG policy in Wenzhou, even during COVID-19. The policy can motivate public general hospitals to improve comprehensive capacity and mitigate discrepancies the efficiency in the expense of treating similar diseases. Policymakers are interested in whether the reform successfully stimulates hospitals to strengthen the internal impetus to improve performance, which is supported by this study.

Keywords diagnosis-related group, points counting payment, performance evaluation, public general hospitals, robust multiple interrupted time series

Strengths and limitations of this study

- ► A robust multiple interrupted time series (R-MITS) model does not assume that the effect of intervention is instantaneous and allows the presence of a potentially lagged (or anticipatory) effect, which match the characteristics of the DRGs' policy.
- ► R-MITS estimates the global change-point, at which the effect of the intervention initiates for the entire health system, rather than removing some points as decided by the study team.
- ► R-MITS can capture mean differences in variability and auto-correlation of time-series changes, improve validity and reduces bias to obtain more accurate estimates.
- ► Due to the inner differences in different hospitals, the DRGs-related indicators selected and subgroups' analysis make the performance evaluation more homogeneous and comparable.
- ► Despite failing to use a difference-in-differences method, this study conducted R-MITS analysis to enhances the relationship between the evolution of results and the intervention.

INTRODUCTION

The healthcare systems in most countries are confronted with increasing pressure, especially regarding the growing demand for healthcare services, while struggling with growing medical insurance expenditures [1,2]. Before the DRGs' payment, the inpatient medical services in China mainly adopted Fee-For-Service, which easily causes the over-treatment and restrains the improvement of service performance. As a pre-payment system, the diagnosis-related groups' (DRGs) payment system, which originated in the United States in the 1970s, involves the adoption of a standard pricing framework that provides equity across healthcare providers in terms of payments for the same types of services [3,4]. Subsequently, many countries and regions have adopted the DRGs-medical-insurance payment in attempts to contain expenditure and increase the transparency, efficiency, and safety of their healthcare systems [4-7]. However, the classification rules of DRGs require high-quality medical record home page data and judgment from physicians. Therefore, the roll-out of DRGs has some challenges in middle-and-low-income countries with limited resources [8].

Maximizing hospitals' service performance in has been prioritized worldwide [9], which also influenced the primary content of the healthcare system reform in China [10]. Before 2020, no province in China implemented the DRGs' payment system in all public hospitals because the DRGs' payment system is costly and requires advanced information systems. To advance the reform of its healthcare system, China has devoted itself to the adoption of the DRGs as one of the main payment methods and optimize the payment rules and relevant mechanisms to meet Chinese healthcare conditions [7,11]. China officially announced the implementation of DRGs' payment in 30 pilot national monitoring cities, and set the following arrangement that initial implementation work commences in 2019, simulated operation commences in 2020, and actual payment commences in 2021 [12,13]. DRGs' payment reform, as an important means of regulating medical behaviors in China [14], can revolutionize the healthcare system. It involves several complementary elements: a mechanism for allocating funding for hospital services; one for hospital management; and a provider payment mechanism within the broader healthcare financing system [3]. Under the funding constraints of DRGs' total budget management, hospitals have to respond aggressively. The DRGs' payment implementation requires hospitals to be proactive regarding the quality of medical records (like more long coding time to ensure the quality of case grouping) and an improvement in the performance of their healthcare services, to avoid losses after the policy implementation. Under the given pricing standard, hospitals are responsible for their losses in cases where medical costs exceed the pricing standard. Such a response can have a wide-ranging effect [15,16], especially, if it can affect hospitals' direct interests that motivate them to regulate clinical behaviors and transform management from passive to positive, which can, in turn, improve service performance.

China places great importance on the impact of DRGs' payment on healthcare performance, because it can influence the overall arrangement and progress but the empirical evidence of its impact in China is still inadequate. Before the effect on service performance can be assessed, it is necessary to obtain a comprehensive understanding of the policy's influence, and to optimize the study design to achieve an optimal evaluation and manage policy makers expectations. Above all, the time at which the payment reform was initiated, announced, and publicized, could impact its outcomes [17]. It sent a strong signal to those hospitals whose cities were not among the national pilots but then decided to implement the DRGs' payment. On one hand, it may have encouraged hospitals to change their management approaches to pursue a more positive adaption of the new payment method, such as by managing medical performance based on DRGs' data [18], improving information levels [19], and strengthening the quality of diagnoses coding [20]. Such measures can provide an early foundation for the formal implementation of DRGs' payment. On another hand, it is also reasonable to assume that the obvious behavioral changes adopted by hospitals occurred after the official implementation, possibly due to certain problems concerning the requisite technology and operation of the DRGs' payment approach [21]. Those hospitals with relatively weak basic conditions required more intensive policy adaptation, such as training personnel in DRGs-associated tasks and improving familiarity with the diagnostic and operational codes for diseases. It can affect reaction time and degree of importance across different hospitals to some extent; however, previous studies may have neglected to include the time until the policy' effects manifested [22,23]. Subsequently, the ability of hospitals to improve their performance is likely to have varied across different institution levels, which is an important factor in identifying the effects of the payment policy. Chinese Health Commission divides medical institutions into three levels according to their tasks and functions to form a hierarchical medical system: primary, secondary, and tertiary hospitals [24]. Tertiary hospitals are the main cost control objects, whose insurance expenditures account for the bulk of medical insurance funds. Previous studies have analyzed them as subgroups in healthcare [25], but few have done so regarding the assessment of the DRGs' payment on service performance across hospitals. Given that

Chinese healthcare is provided almost by state-owned public general hospitals [26], it is important to explore whether there are differences across subgroups in terms of the effects [27].

On November 12, 2019, Zhejiang province issued interim measures on DRGs' points-counting payment (PCP) for hospitalization expenses in basic medical insurance that required all sub-provincial and prefecture-level cities to implement PCP based on DRGs under the total budget management from January 2020 [28]. The points of a given case are determined as follows:

$$\text{Points(case)} = \text{corresponding DRG benchmark points} \times \text{difference coefficient}$$

The health insurance department of every sub-provincial or prefecture-level city sets the calculation standards for benchmark points and difference coefficients. The health insurance department of the district or county-level city calculates the point value (PV). PV is dynamic and changes depending on the current annual DRGs' cost settlement and points, which are calculated at the beginning of the following year. PV is determined as follows:

$$\text{PV} = \frac{\text{Annual DRGs cost settlement in a district or county - level city}}{\text{Annual DRGs points in this district or county - level city}}$$

The revenue of DRGs' payment for a hospital is determined as follows:

$$\text{Revenue(Hospital)=Annual points(Hospital)} \times \text{PV}$$

This PCP based on DRGs applies the principle of work point theory with reference to the weighting criteria of each DRG, to establish the relative price relationship between medical costs and weights of different disease groups; and then convert the points of each DRG group, using these points to allocate the regional health insurance. Zhejiang detailed hospitalization cases into 1006 DRGs based on the nation 618 DRGs grouping scheme. Zhejiang is the first province to announce enforcement across the entire province; whether the PCP method based on DRGs can exert a strong motivation on hospitals to improve their service performance is important for China but it is still an open topic. Given that, this study takes Wenzhou in Zhejiang province as a sample city to evaluate hospital performance before and after the implementation of the DRGs' reform between 2018 and 2021 in public general hospitals from the perspective of institution levels.

METHODS

Data source and sample size

Wenzhou City lies in the southeast of Zhejiang province. In 2021, the resident population was 9.6450 million—the second most populous city in Zhejiang. It has a wealth of successful experience in reform and strong comprehensive medical service [29], a similar DRGs' policy implementation process and healthcare system structure to other cites, which act as a representative for estimating the effects of the PCP on public general hospitals. Wenzhou performed upfront preparations for DRGs' payment method with effect from May 20, 2019, earlier than the other cities. Since then, more than 21 related symposia have been presented among hospitals. In 2019, the Wenzhou Healthcare Security Administration conducted disease grouping and cost measurement for the past three years. The base data was collected from 161 medical institutions in the city that offer hospitalization. After grouping by DRGs, hospitalization cases were divided into 912 DRGs. Among them were 55 DRGs where the number of cases was ≤5; 857 DRGs where number of cases was>5 and their coefficient of variation was <1; reduction in variance was 75.04% and the cost differences across groups were all >20%, which reached the national technical specifications for DRGs grouping and payment. On December 13, 2019, the Wenzhou local government held a publicity meeting on the payment policy and issued a detailed draft of the rules, after which it began communicating with hospitals to optimize the policy. On December 27, 2019, the DRGs' Points Counting Payment Implementation Rules was issued, and scheduled for implementation in January 2020 [30]. Figure 1 shows the timeline for 2019 for the implementation of the DRGs' payment policy.

According to Chinese hospital classification and management standards, primary hospitals are hospitals that provide preventive, medical, healthcare, and rehabilitation services directly to a community of a certain population. Secondary hospitals are regional hospitals that provide comprehensive medical and health services to multiple communities and undertake certain teaching and research. Tertiary hospitals are hospitals above the regional level that provide high-level specialized medical services to several regions and perform higher medical education and research. Above-primary-level public general hospitals in Wenzhou amounted to 22 institutions, which consisted of eight tertiary hospitals and 14 secondary hospitals. In 2021, the number of discharged cases in above-primary-level public general hospitals in Wenzhou was 840,342, accounting for 98.79% of the 850,652 discharged cases. The intervention was implemented officially in 2020, the time-series data of these 22 institutions collected spanned from 2018 to 2021. Before implementation: 2018–2019; after implementation: 2020–2021. The data were collected from the Zhejiang Hospital Quality Management and Performance Evaluation Platform, which was developed by the Zhejiang Provincial Health Commission. The platform launched in 2017 and, reviewed the quality of the data uploaded by each hospital by logic verification and key quality control index monthly before evaluating the performance. The platform allows administrative departments and medical institutions to query DRGs' performance data. In 2021, it evaluated 424 above-primary-level hospitals in Zhejiang province.

Indicators

This study used five performance indicators calculated by the Zhejiang Hospital Quality Management and Performance Evaluation Platform that reflects the medical capacity and consumption of medical resources. The five performance indicators are case-mix index (CMI), cost per hospitalization (CPH), its cost efficiency index (CEI), average length of stay (ALOS), and its time efficiency index (TEI). Among them, the CMI, CEI, and TEI derives from the DRGs' system, which has been found to have excellent homogeneous comparability [31-34]. The following are the definitions and calculations of CMI, CEI, and TEI as determined by Zhejiang Provincial Health Commission.

1) CMI reflects the average comprehensive medical ability of a medical institution and its average seriousness or difficulty of the diseases. The calculation method is:

$$\text{CMI} = \frac{\text{Total weight}}{\text{Number of cases in region}}$$

$$\text{Total weight} = \sum \text{Each DRG weight} \times \text{Number of cases in each DRG}$$

2) CEI reflects the cost of treating similar diseases in this region. When the value of CEI was close to 1, it suggested that, for the region in question, the cost efficiency was close to the average level; values below 1 or above 1 suggested that the efficiency was above or below the average level, respectively.

$$\text{CEI} = \frac{\sum_j k_j^c n_j}{\sum_j n_j}$$

$$\text{Cost ratio}(k^c) = \frac{\text{average cost of a DRG in hospital } (c_i)}{\text{average cost of a DRG in region } (\overline{C}_i)}$$

where n_j represents the number of cases in DRG_j, CEI is the weighted average of k^c.

3) TEI reflects the time spent treating similar diseases in this region. When the value of TEI was close to 1, it suggested that, for the region in question, the time efficiency was close to the average level; values below 1 or above 1 suggested that the efficiency was above or below the average level, respectively.

$$\text{TEI} = \frac{\sum_j k_j^l n_j}{\sum_j n_j}$$

$$\text{Average length of stay ratio } (k^l) = \frac{\text{average length of a DRG in hospital } (l_i)}{\text{average length of a DRG in region } (\overline{L}_i)}$$

where n_j represents the number of cases in DRG_j, TEI is the weighted average of k^l.

To reflect the differences in performances from the perspective of institutional levels, this study employed the mean of indicator in total institutions, tertiary institutions, and secondary institutions, respectively.

Statistical analysis

Interrupted time series has become a core study design for the evaluation of health policies [17]. This study used a robust multiple interrupted time series (R-MITS) model, which have several advantages over the traditional ITS designs [35-37]. Above all, R-MITS borrows information from all micro-systems to estimate the global change point, that is, it determines the time point at which the effect of the intervention initiates for the entire health system, rather than removing some points as decided by the study team [38]. An intervention to increase the stability and predictability of an outcome may have decreased outcome variability and increased auto-correlation [37]. Given that decreasing outcome variability can contribute to less extreme values, while increasing auto-correlation can contribute to a more predictable outcome, it is important to capture alongside mean differences in variability and auto-correlation. Compared with traditional designs, this R-MITS model can capture all the aforementioned time series' changes. Furthermore, it does not assume that the intervention is instantaneous and allows the effect of the intervention to be abrupt or gradual, thus allowing the presence of a potentially lagged (or anticipatory) effect. It can obtain more accurate estimates, which further improves internal and external validity, and reduces bias [35,39]. This robust model has been widely adopted to evaluate: improvements in care quality at both the system and unit level [39]; the effectiveness of a quality-improvement initiative for reducing cardiac arrests among infants and children [40]; the impact of trauma center accreditation on mortality and complications [41]; and the effectiveness of chest-pain units on observation times, diagnostic agreement, and emergency department re-admission and hospitalizations [42]. The R-MITS model is shown in Appendix A.

Patient and public involvement

None.

RESULTS

Descriptive analysis

The yearly outcomes for every target indicator for the period of 2018–2021 (representing pre- and post-implementation) are shown in Table 1.

The yearly results for the CMI mean in tertiary institutions decreased from 0.8678 to 0.8641 before 2020 and increased from 0.8864 to 0.9041 after 2020; in secondary institutions, it decreased from 0.7590 to 0.7402 before 2020 and increased from 0.8047 to 0.8332 after 2020.

For the CPH mean in tertiary institutions, it decreased from 12658.6038 to 12569.7275 before 2020 and decreased from 13192.7838 to 12951.4875 after 2020; in secondary institutions, it increased from 6471.0164 to 7503.7829 yearly before and after 2020. For the CEI mean in tertiary institutions, it decreased from 1.1238 to 1.0950 before 2020 and decreased from 1.1100 to 1.1075 after 2020; in secondary institutions, it decreased from 0.7471 to 0.7407 before 2020 and increased from 0.7921 to 0.8264 after 2020.

For the ALOS mean in tertiary institutions, it decreased from 8.5188 to 7.0475 yearly before and after 2020; in secondary institutions, it decreased from 8.1279 to 8.1171 before 2020 and decreased from 8.4679 to 8.1793 after 2020. For the TEI mean in tertiary institutions, it decreased from 1.0288 to 0.9875 yearly before and after 2020; in secondary institutions, it increased from 0.9307 to 1.0264 yearly before and after 2020.

Table 1. Unit-specific yearly results pre- and post-policy implementation

Unit	Year	CMI mean	CPH mean	CEI mean	ALOS mean	TEI mean
Total institutions	2018	0.7986	8721.0482	0.8841	8.2700	0.9664
	2019	0.7852	8775.5059	0.8695	8.0264	0.9682
	2020	0.8344	9418.2600	0.9077	8.1859	1.0000
	2021	0.8590	9484.7664	0.9286	7.7677	1.0123
Tertiary institutions	2018	0.8678	12658.6038	1.1238	8.5188	1.0288
	2019	0.8641	12569.7275	1.0950	7.8675	1.0075
	2020	0.8864	13192.7838	1.1100	7.6925	0.9925
	2021	0.9041	12951.4875	1.1075	7.0475	0.9875
Secondary institutions	2018	0.7590	6471.0164	0.7471	8.1279	0.9307
	2019	0.7402	6607.3793	0.7407	8.1171	0.9457
	2020	0.8047	7261.3893	0.7921	8.4679	1.0043
	2021	0.8332	7503.7829	0.8264	8.1793	1.0264

Note: ALOS: average length of stay; CEI: cost efficiency index; CMI: case-mix index; CPH: cost per hospitalization; TEI: time efficiency index; pre-implementation: from 2018 to 2019; post-implementation: from 2020 to 2021.

R-MITS analysis

Estimated changes using R-MITS: for all samples and subgroups are shown in Table 2; the fitted regression lines are shown in Figure 2; the estimates of the stochastic component parameters are shown in Appendix B; the residuals of the estimated model and auto-correlation functions are as expected and are shown in Appendix 1 through 5, demonstrating that the curves for both the pre- and post-groups for the measures included are linear and independent.

The R-MITS results for the CMI mean in tertiary institutions, the effect change-point was estimated February 2020 ($p<0.05$) with the slope changing from −0.0001 to 0.0008 ($p<0.01$); the effect change-point in secondary institutions was estimated for January 2020 ($p<0.0001$) with the slope changing from −0.0012 to 0.0019 ($p<0.0001$).

For the CPH mean in tertiary institutions, the effect change-point was estimated January 2020 ($p<0.01$) with the slope changing from −18.7100 to −43.3100 ($p>0.05$); in secondary institutions, the effect change-point was estimated for April 2020 ($p<0.001$) with the slope changing from 8.5540 to −0.8039 ($p>0.05$). For the CEI mean in tertiary institutions, the effect change-point was estimated for June 2019 ($p<0.0001$) with the slope changing from 0.0020 to 0.0005 ($p<0.05$); in secondary institutions, the effect change-point was estimated for April 2020 ($p<0.0001$) with the slope changing from 0.0003 to 0.0015 ($p<0.05$).

For the ALOS mean in tertiary institutions, the effect change-point was estimated for January 2020 ($p<0.05$) with the slope changing from −0.0488 to −0.0618 ($p>0.05$); in secondary institutions, the effect change-point was estimated for January 2020 (p<0.0001) with the slope changing from −0.0041 to −0.0261 (p<0.01). For the TEI mean in tertiary institutions, the effect change-point was estimated for June 2020 (p>0.05) with the slope changing from −0.0010 to −0.0005 (p>0.05); in secondary institutions, the effect change-point was estimated for April 2020 (p<0.0001) with the slope changing from 0.0016 to −0.0002 (p<0.05).

Table 2. Unit-specific changes estimated by R-MITS pre- and post-intervention

Outcomes	Unit	Change-points	Changes in slope			Changes in intercept (δ_i^{τ})
			Pre-change (β_{i1}^{τ})	Post-change ($\beta_{i1}^{\tau}+\Delta_i^{\tau}$)	Change (Δ_i^{τ})	
CMI mean	Total institutions	Mar-2020****	−0.0002	0.0013	0.0015****	0.0100
	Tertiary institutions	Feb-2020*	−0.0001	0.0008	0.0009**	0.0015
	Secondary institutions	Jan-2020****	−0.0012	0.0019	0.0031****	−0.0099
CPH mean	Total institutions	Jan-2020***	−6.4420	−5.2860	1.1560	909.9000**
	Tertiary institutions	Jan-2020**	−18.7100	−43.3100	−24.6000	1957.0000***
	Secondary institutions	Apr-2020***	8.5540	−0.8039	−9.3580	1075.0000***
CEI mean	Total institutions	Jun-2019****	0.0011	0.0024	0.0013**	−0.0505**
	Tertiary institutions	Jun-2019****	0.0020	0.0005	−0.0015*	−0.0238
	Secondary institutions	Apr-2020****	0.0003	0.0015	0.0012*	0.0115
ALOS mean	Total institutions	Jan-2020****	−0.0202	−0.0392	−0.0190*	1.0710***
	Tertiary institutions	Jan-2020*	−0.0488	−0.0618	−0.0130	0.9115*
	Secondary institutions	Jan-2020****	−0.0041	−0.0261	−0.0220**	1.1540***
TEI mean	Total institutions	Apr-2020***	0.0006	−0.0005	−0.0011	0.0659**
	Tertiary institutions	Jun-2020	−0.0010	−0.0005	0.0005	−0.0231
	Secondary institutions	Apr-2020****	0.0016	−0.0002	−0.0018*	0.1102***

Note: Significance level α=0.05; *, p<0.05, **, p<0.01; ***, p<0.001; ****, p<0.0001; ALOS: average length of stay; CEI: cost efficiency index; CMI: case-mix index; CPH: cost per hospitalization; TEI: time efficiency index.

DISCUSSION

With limited and stretched healthcare resources, it is important for hospitals to further improve the performance of their medical services after the new policy implementation. In general, the results from all samples and subgroups seem positive from various perspectives.

As for all samples and subgroups, this study found that the CMI mean showed a downward curve before the change but increased after the change, from which it can be inferred that the policy implementation likely motivated these hospitals to devote themselves to improving CMI. According to the payment rules, improving CMI can contribute to securing more DRGs' points, thereby acquiring more medical insurance funds. The study indicated that CMI change lagged two months after the DRG policy officially implemented, which may result from the difficulty in strengthening the comprehensive ability immediately even if they have stronger motivation than before. Besides, this study also found that the CMI mean in tertiary institutions changed later than in secondary institutions and improved slower than in secondary institutions after change. CMI in tertiary institutions has already reached high levels and its improving would require more effort than secondary institutions. Due to Chinese citizens being at liberty to choose healthcare services at any institution, tertiary institutions with higher levels of comprehensive medical strength can attract more patients under the siphon effect [43,44]. With service accessibility improving, patients would rather visit senior hospitals, even though most of their illnesses are not severe [45,46]. As a result, it is common for a large number of low-related weight cases are treated at tertiary institutions that constrained improving CMI, and also affected the hierarchic healthcare treatment system performance for a long time. Significantly, under the influence of DRGs' payment rules, tertiary institutions tended to treat more serious diseases (that possess high relative weights) and were willing to transfer the more common or easy-to-treat cases to lower-grade institutions [21]. Subsequently, the proportion of the high relative weight of their treating diseases can be increased, which contributed to improving their CMI. In a sense, the implementation of the DRGs' policy can promote the hierarchic healthcare treatment system performing its intended role in China.

After the policy intervention, the CEI mean in all samples increased at a faster rate toward 1, from which it can be inferred that the policy can further motivate these institutions to narrow the differences across institutions in terms of costs for treating similar diseases. Interestingly, the results showed that the CEI mean changed in June 2019, following the commencement of the policy reform in May 2019. Different from Fee-for-Service, the rules of DRGs' payment is that "excess expenditures will no longer be reimbursed and the balance will be retained." [47], meaning ultra-high-cost inpatient cases in same group will cause financial losses for hospitals. The introduction of the policy could increase cost-control consciousness, and motivate associated management behaviors [3,48], like improving coding quality of DRGs' data [21], which made the response of these institutions to strengthening the CEI faster than other indicators. Focusing only on changes at subgroups, this study showed that tertiary institutions reacted faster than secondary institutions to the DRGs' policy in terms of CEI; the CEI mean in tertiary institutions remained higher than 1, while remaining lower than 1 in secondary institutions before and after change. As is known, the CPH in tertiary institutions are higher than secondary institutions; similarly, the phenomenon of spending high costs on similar diseases was more common in tertiary institutions, which resulted in their high CEI. Because medical equipment, technical levels, drugs, and consumables in tertiary institutions were more advanced, they attract higher expenses.[13] It was also found that the uptrend of the CEI mean in tertiary institutions was slower than before; the CEI mean in secondary institutions changed faster toward 1, which can be inferred that the difference in the tertiary and secondary institutions was mitigated by the intervention.

A decrease in the length of stay can be either positive or negative for individual patients and for the health system as a whole [49]. ALOS reflects the average time of healthcare resources taken up by each inpatient; the longer it takes, the more likely the hospital makes a loss. It may be the reason that the DRGs' policymakers are strengthening the readmission rate monitoring, and some penalties [28,30], because it is not guaranteed that all hospitals will not intentionally discharge inpatients early to minimize their losses or to increase revenues [21]. Certainly, decreasing TEI can be reasonable, but it requires the prior improvement of medical levels across hospitals. After the intervention, this study found that the TEI mean in tertiary institutions did not change significantly, while in secondary institutions it did. As is known, compared with secondary institutions, tertiary institutions with stronger medical service capacity, whose inpatients with similar diseases can be treated better. As a result, tertiary institutions require much more efforts to further decline the TEI in a short time. Besides, after the policy was implemented, the CMI mean both in tertiary and secondary institutions from decrease to increase can be inferred that the percentage of serious or complex diseases treated in these institutions have all improved, which mean hospitals require more time to treat inpatients. Under the CMI mean from decrease to increase, secondary institutions have a huge potential to improve and successfully change the TEI, compared to tertiary institutions.

The findings presented here should be generalized with caution. The impact of COVID-19 reached Zhejiang province (to where Wenzhou is located) at the end of January; fortunately, the setback was temporary due to the implementation of strict prevention and control measures, and recovery gradually occurred after February 2020. Nevertheless, it may have unavoidably resulted in decreasing patient visits and hospital revenues, while increasing operation costs and medical efficiency, which may hinder the improvement of these selected indicators. We could have minimized the disruption of the COVID-19 by removing data from the early years of the policy, but we did not do so because it would have been too subjective. The R-MITS model can identify inferences regarding the estimation of global change points across units, rather than removing some series as decided by the study team. In such an external environment, the implementation of the DRGs' payment would be timely and may make up for the negative impact of the epidemic, which could motivate cost-control consciousness and management behaviors [3,48]. Because under the motivation of DRGs' payment, hospitals have to make adjustments to strengthen service performance, otherwise the revenue loss would be huge. Theoretically, the policy can guide high-grade-level hospitals to focus on the treatment of more complex and severe diseases and the breakthrough of medical service level, while returning routine and common diseases to low-grade-level hospitals for treatment to improve the CMI indicator. In addition, motivating hospitals to improve medical levels and efficiency can promote other indicators. This study showed that the DRGs' policy had positive effects on hospital performance, even during COVID-19.

Regarding the analysis method, we were unable to apply the difference-in-differences, because the basic medical insurance designated medical institutions of Zhejiang province were all required to implement DRGs' payment, including Wenzhou. In addition, we could not obtain data from the public general hospitals outside Zhejiang province where the DRGs or other payment reforms (like Diagnosis-Intervention Packet payment [50]) were not implemented, under the process of the nationwide payment reform. Fortunately, R-MITS is an optimized version of ITS, whose strengths can enhance the relationship between the evolution of results and the DRGs' policy intervention. Besides, the subject of this study were public general hospitals; however, we collected data from secondary and tertiary public general hospitals, excluding primary hospitals that did not have time-series data of DRGs-related indicators pre- and post-intervention. DRGs' indicators can homogeneously compare the medical performance, but their calculations are more dependent on the quality of the homepage of inpatient medical records than traditional indicators, especially disease coding quality. Thus, previous studies that used the DRGs' indicators were almost conducted in above-primary-level hospitals, include this study. CMI reflects the average level of medical technology and complexity of the diseases attended to by each hospital; CEI or TEI reflect the relative efficiency of CPH or ALOS, which are calculated based on DRGs. This evaluation was conducted close to the intervention's time, which might not produce the same results as other evaluations carried out at a later date.

CONCLUSIONS

Despite the inevitable impact of the COVID-19 on hospitals, the design of this study succeeded in revealing some anticipated and positive effects of the DRGs' policy on hospitalization performance. Implementing the policy can motivate public general hospitals to improve the medical comprehensive capacity and, mitigate their efficiency discrepancies in the expense of treating similar diseases across the different grade institutions. Hospitals firstly focused on strengthening hospitalization efficiency, though there were differences between tertiary and secondary hospitals when responding to DRGs' policy. One of the original intentions of the reform is to stimulate hospitals to form the internal impetus to improve performance, changing the rude and less-effective development mode of the past. From this perspective, our results further supported the target of the reform in China.

Contributors

TTZ and CC were responsible for conceptualization, data collection, methodology, resources, and software. XXZ was responsible for conceptualization, data collection, and methodology. QRY was responsible for methodology and software. YPH was responsible for data collection and descriptive analysis. RYL was responsible for data collection and investigation. XYZ and YD were responsible for conceptualization, funding acquisition, resources, supervision, reviewing, and editing of the manuscript. All authors contributed to the critical revision of the manuscript for important intellectual content, review, and approval of the final manuscript.

Competing interests

None declared.

Funding

This research was sponsored by Zhejiang Provincial Medical and Health Science Technology Plan Project (2023KY889) from Zhejiang Provincial Health Commission in China.

Data availability statement

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval

The study does not involve human participants. Approvals regarding data access for the study were obtained from the Health Information Center, Health Commission of Wenzhou, China.

Acknowledgements

The authors would like to thank the support of the Hospital Quality Management and Performance Evaluation Platform from Zhejiang Provincial Health Commission.

References

1. Stadhouders N, Kruse F, Tanke M, Koolman X, Jeurissen P. Effective healthcare cost-containment policies: a systematic review. *Health Policy*. Jan 2019;123(1):71–79. doi:10.1016/j.healthpol.2018.10.015
2. Kim YL, K.-H.; Choi, S.W. Multifaced evidence of hospital performance in Pennsylvania. *Healthcare*. 2021;9:670. doi:10.3390/healthcare9060670
3. Annear PL, Kwon S, Lorenzoni L, et al. Pathways to DRG-based hospital payment systems in Japan, Korea, and Thailand. *Health Policy*. Jul 2018;122(7):707–713. doi:10.1016/j.healthpol.2018.04.013
4. Busse R, Geissler A, Aaviksoo A, et al. Diagnosis related groups in Europe: moving towards transparency, efficiency, and quality in hospitals? *BMJ*. Jun 7 2013;346:f3197. doi:10.1136/bmj.f3197
5. Chok L, Bachli EB, Steiger P, et al. Effect of diagnosis related groups implementation on the intensive care unit of a Swiss tertiary hospital: a cohort study. *BMC Health Serv Res*. Feb 5 2018;18(1):84. doi:10.1186/s12913-018-2869-4
6. Zou K, Li HY, Zhou D, Liao ZJ. The effects of diagnosis-related groups payment on hospital healthcare in China: a systematic review. *BMC Health Serv Res*. Feb 12 2020;20(1):112. doi:10.1186/s12913-020-4957-5
7. Mathauer I, Wittenbecher F. Hospital payment systems based on diagnosis-related groups: experiences in low- and middle-income countries. *Bull World Health Organ*. Oct 1 2013;91(10):746–756A. doi:10.2471/BLT.12.115931
8. Tang X, Zhang X, Chen Y, Yan J, Qian M, Ying X. Variations in the impact of the new case-based payment reform on medical costs, length of stay, and quality across different hospitals in China: an interrupted time series analysis. *BMC Health Serv Res*. Jun 2 2023;23(1):568. doi:10.1186/s12913-023-09553-x
9. Bush RW. Reducing waste in US health care systems. *JAMA*. Feb 28 2007;297(8):871–4. doi:10.1001/jama.297.8.871
10. Cui Y, Wu Z, Lu Y, Jin W, Dai X, Bai J. Effects of the performance management information system in improving performance: an empirical study in Shanghai Ninth People's Hospital. *Springerplus*. 2016;5(1):1785. doi:10.1186/s40064-016-3436-2
11. Jian W, Lu M, Liu G, Chan KY, Poon AN. Beijing's diagnosis-related group payment reform pilot: impact on quality of acute myocardial infarction care. *Soc Sci Med*. Dec 2019;243:112590. doi:10.1016/j.socscimed.2019.112590
12. Administration NHS. Notice on the lists of national pilot cities for diagnosis related grouping payment. http://www.gov.cn/zhengce/zhengceku/2019-10/12/content_5438769.htm

13. Bredenkamp C, Bales, S., & Kahur, K. *Transition to diagnosis-related group (DRG) payments for health lessons from case studies*. 2020.
14. Council CS. Guidance on further deepening the basic medical insurance payment reform http://www.gov.cn/gongbao/content/2017/content_5210497.htm
15. McClellan M. Reforming payments to healthcare providers: the key to slowing healthcare cost growth while improving quality? *J Econ Perspect*. Spring 2011;25(2):69–92. doi:10.1257/jep.25.2.69
16. Bystrov V, Staszewska-Bystrova A, Rutkowski D, Hermanowski T. Effects of DRG-based hospital payment in Poland on treatment of patients with stroke. *Health Policy*. Aug 2015;119(8):1119–25. doi:10.1016/j.healthpol.2015.04.017
17. Lopez Bernal J, Soumerai S, Gasparrini A. A methodological framework for model selection in interrupted time series studies. *J Clin Epidemiol*. Nov 2018;103:82–91. doi:10.1016/j.jclinepi.2018.05.026
18. Zhao S, Gu Y, Huang Z, Zeng W. Building a performance management system for hospitals based on diagnosis-related group (DRG) payment. *Journal of Sensors*. 2022;2022:1-10. doi:10.1155/2022/7001423
19. Jian W, Lu M, Han W, Hu M. Introducing diagnosis-related groups: is the information system ready? *Int J Health Plann Manage*. Jan-Mar 2016;31(1):E58-68. doi:10.1002/hpm.2270
20. Garcia Calderón V, Figueiras Huante IA, Carbajal Martinez M, et al. The impact of improving the quality of coding in the utilities of diagnosis related groups system in a private healthcare institution. 14-year experience. *Int J Med Inform*. Sep 2019;129:248-252. doi:10.1016/j.ijmedinf.2019.06.019
21. Barouni M, Ahmadian L, Anari HS, Mohsenbeigi E. Challenges and adverse outcomes of implementing reimbursement mechanisms based on the diagnosis-related group classification system: a systematic review. *Sultan Qaboos Univ Med J*. Aug 2020;20(3):e260–e270. doi:10.18295/squmj.2020.20.03.004
22. Jiang Q, Xu Z, Yu L, Zhou H, Zhang Z. Impact Analysis on the Expenses of C-DRG Pricing and Payment Reform on Medical Facilities in Sanming City. *Chinese Health Economics*. 2021;40(04):13-16.
23. Fang J, Liu L, Peng Y, Tao H. Analysis of the Impact of CHS-DRG Payment Reform on Hospital Operation CHS-DRG. *Health Economics Research*. 2022;39(05):67-71.
24. Dong E, Liu S, Chen M, et al. Differences in regional distribution and inequality in health-resource allocation at hospital and primary health centre levels: a longitudinal study in Shanghai, China. *BMJ Open*. Jul 19 2020;10(7):e035635. doi:10.1136/bmjopen-2019-035635
25. Gao L, Shi L, Meng Q, Kong X, Guo M, Lu F. Effect of healthcare system reforms on public hospitals' revenue structures: Evidence from Beijing, China. *Soc Sci Med*. Aug 2021;283:114210. doi:10.1016/j.socscimed.2021.114210
26. Zhao X, Xiao J, Chen H, et al. Patient preferences and attitudes towards first choice medical services in Shenzhen, China: a cross-sectional study. *BMJ Open*. May 24 2022;12(5):e057280. doi:10.1136/bmjopen-2021-057280
27. Meng Z, Hui W, Cai Y, Liu J, Wu H. The effects of DRGs-based payment compared with cost-based payment on inpatient healthcare utilization: a systematic review and meta-analysis. *Health Policy*. Apr 2020;124(4):359-367. doi:10.1016/j.healthpol.2020.01.007
28. Administration ZPHS. Notice on the issuance of the interim rules for payment of DRGs points for inpatient costs of basic medical insurance in Zhejiang province. http://ybj.zj.gov.cn/art/2019/11/27/art_1229113757_564142.html
29. Government WMPs. Notice on the fourteenth five-year plan for Wenzhou healthcare development. http://www.wenzhou.gov.cn/art/2021/12/1/art_1229116916_1911427.html
30. Administration WHS. Notice on Wenzhou basic medical insurance total hospitalization cost budget and DRGs point payment implementation rules. http://ybj.wenzhou.gov.cn/art/2019/12/30/art_1229135972_623703.html
31. Ding S, Jiang X, Zheng Y, et al. Improving the clinical ability and quality of endocrinology department with diagnosis-related groups tool. *Ann Transl Med*. Feb 2022;10(4):167. doi:10.21037/atm-22-147

32. Feng L, Tian Y, He M, et al. Impact of DRGs-based inpatient service management on the performance of regional inpatient services in Shanghai, China: an interrupted time series study, 2013–2019. *BMC Health Serv Res*. Oct 12 2020;20(1):942. doi:10.1186/s12913-020-05790-6

33. Jian W, Huang Y, Hu M, Zhang X. Performance evaluation of inpatient service in Beijing: a horizontal comparison with risk adjustment based on Diagnosis Related Groups. *BMC Health Serv Res*. Apr 30 2009;9:72. doi:10.1186/1472-6963-9-72

34. Liu X, Liu F, Wang L, Wu M, Yang L, Wei L. Performance evaluation of medical service for breast cancer patients based on diagnosis related groups. *BMC Health Serv Res*. May 24 2021;21(1):496. doi:10.1186/s12913-021-06439-8

35. Cruz M, Bender M, Ombao H. A robust interrupted time series model for analyzing complex health care intervention data. *Statistics in medicine*. Dec 20 2017;36(29):4660-4676. doi:10.1002/sim.7443

36. Cruz M, Gillen DL, Bender M, Ombao H. Assessing health care interventions via an interrupted time series model: Study power and design considerations. *Statistics in medicine*. May 10 2019;38(10):1734-1752. doi:10.1002/sim.8067

37. Cruz M, Pinto-Orellana MA, Gillen DL, Ombao HC. RITS: a toolbox for assessing complex interventions via interrupted time series models. *BMC Med Res Methodol*. Jul 8 2021;21(1):143. doi:10.1186/s12874-021-01322-w

38. Cruz M, Gillen DL, Bender M, Ombao H. Assessing health care interventions via an interrupted time series model: study power and design considerations. *Stat Med*. May 10 2019;38(10):1734–1752. doi:10.1002/sim.8067

39. Bender M, Murphy EA, Cruz M, Ombao H. System- and unit-level care quality outcome improvements after integrating clinical nurse leaders into frontline care delivery. *J Nurs Adm*. Jun 2019;49(6):315–322. doi:10.1097/NNA.0000000000000759

40. Ferguson LP, Thiru Y, Staffa SJ, Guillen Ortega M. Reducing cardiac arrests in the PICU: initiative to improve time to administration of prearrest bolus epinephrine in patients with cardiac disease. *Crit Care Med*. Jul 2020;48(7):e542–e549. doi:10.1097/CCM.0000000000004349

41. Batomen B, Moore L, Strumpf E, Champion H, Nandi A. Impact of trauma centre accreditation on mortality and complications in a Canadian trauma system: an interrupted time series analysis. *BMJ Qual Saf*. Nov 2021;30(11):853–866. doi:10.1136/bmjqs-2020-011271

42. Alzate-Gutiérrez N, Roldán-Toro H, Gallo-Villegas J, Aristizábal-Ocampo D. Efectividad de una unidad de dolor torácico en el primer nivel de atención en salud con telecardiología: un estudio cuasiexperimental de series de tiempo interrumpidas. *Revista Colombiana de Cardiología*. 2022;28(4)doi:10.24875/rccar.M21000062

43. Jiang Q, Pan J. The evolving hospital market in China after the 2009 healthcare reform. *Inquiry*. Jan-Dec 2020;57:46958020968783. doi:10.1177/0046958020968783

44. Yan K, Yang C, Zhang H, et al. Impact of the zero-mark-up drug policy on drug-related expenditures and use in public hospitals, 2016–2018: an interrupted time series study in Shaanxi. *BMJ Open*. Nov 26 2020;10(11):e037034. doi:10.1136/bmjopen-2020-037034

45. Chu Y, Tao H. Detecting undifferentiation of tertiary and county hospitals in China in adoption of DRG instrument. *Healthcare (Basel)*. Jul 21 2021;9(8)doi:10.3390/healthcare9080922

46. Yu W, Li M, Nong X, et al. Practices and attitudes of doctors and patients to downward referral in Shanghai, China. *BMJ Open*. Apr 3 2017;7(4):e012565. doi:10.1136/bmjopen-2016-012565

47. Yu L, Lang J. Diagnosis-related groups (DRG) pricing and payment policy in China: where are we? *Hepatobiliary Surg Nutr*. Dec 2020;9(6):771–773. doi:10.21037/hbsn-2020-8

48. Choi JW, Kim SJ, Park HK, Jang SI, Kim TH, Park EC. Effects of a mandatory DRG payment system in South Korea: analysis of multi-year nationwide hospital claims data. *BMC Health Serv Res*. Oct 30 2019;19(1):776. doi:10.1186/s12913-019-4650-8

49. Kone I, Maria Zimmermann B, Nordstrom K, Simone Elger B, Wangmo T. A scoping review of empirical evidence on the impacts of the DRG introduction in Germany and Switzerland. *Int J Health Plann Manage*. Jan 2019;34(1):56–70. doi:10.1002/hpm.2669

50. Ding Y, Yin J, Zheng C, Dixon S, Sun Q. The impacts of diagnosis-intervention packet payment on the providers' behavior of inpatient care-evidence from a national pilot city in China. *Front Public Health*. 2023;11:1069131. doi:10.3389/fpubh.2023.1069131

Figure 1. Implementation process for the diagnosis-related groups (DRGs) payment policy for 2019 in Wenzhou

Figure 2. Fitted regression lines for institutions before and after change-points

Note: Pictures A_i, B_i, C_i, D_i, E_i are the fitted regression lines of outcomes (CMI mean, CPH mean, CEI mean, ALOS mean, TEI mean) in total institutions(i=1), tertiary institutions(i=2), secondary institutions(i=3) respectively. ALOS: average length of stay; CEI: cost efficiency index; CMI: case-mix index; CPH: cost per hospitalization; TEI: time efficiency index.

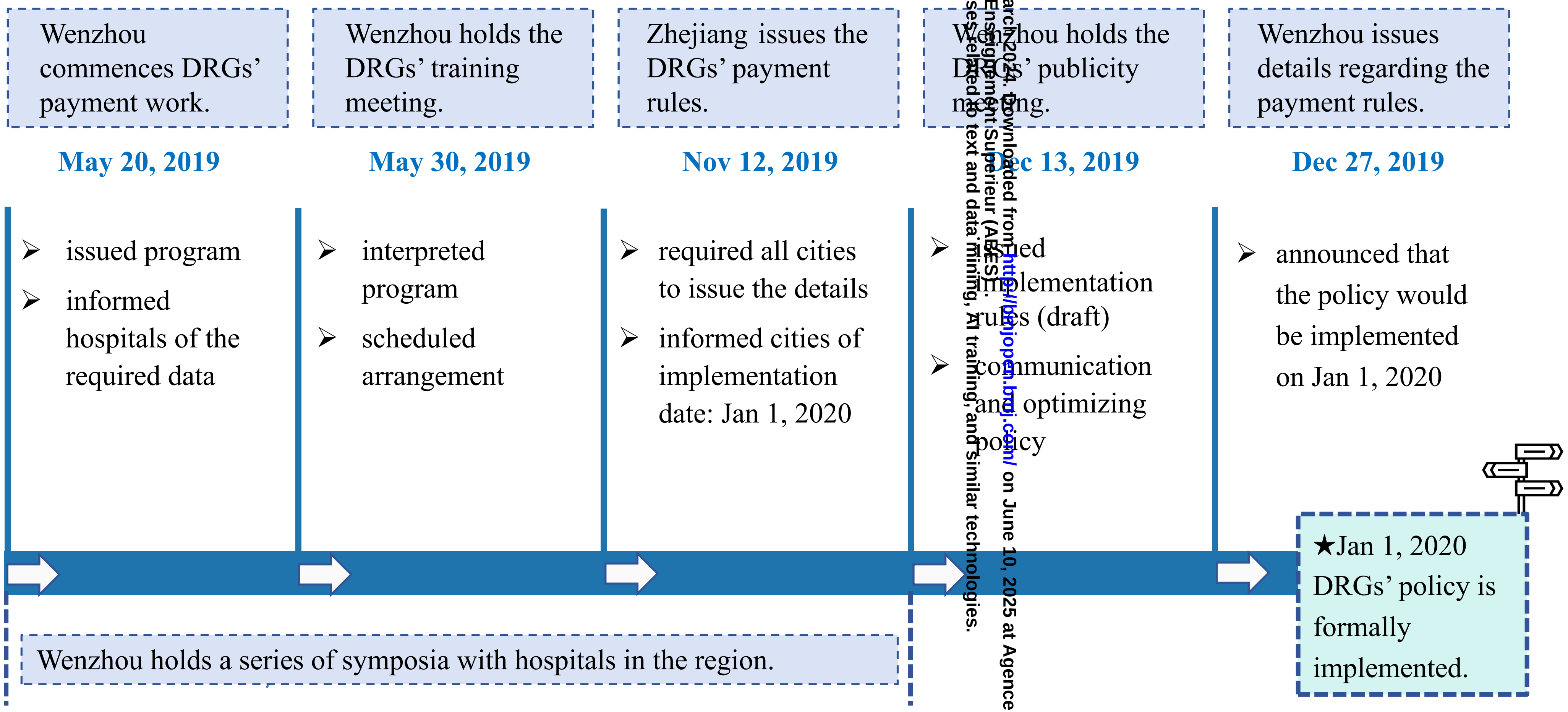


Figure 1. The implementation process for the diagnosis-related groups' (DRGs) policy for 2019 in Wenzhou.

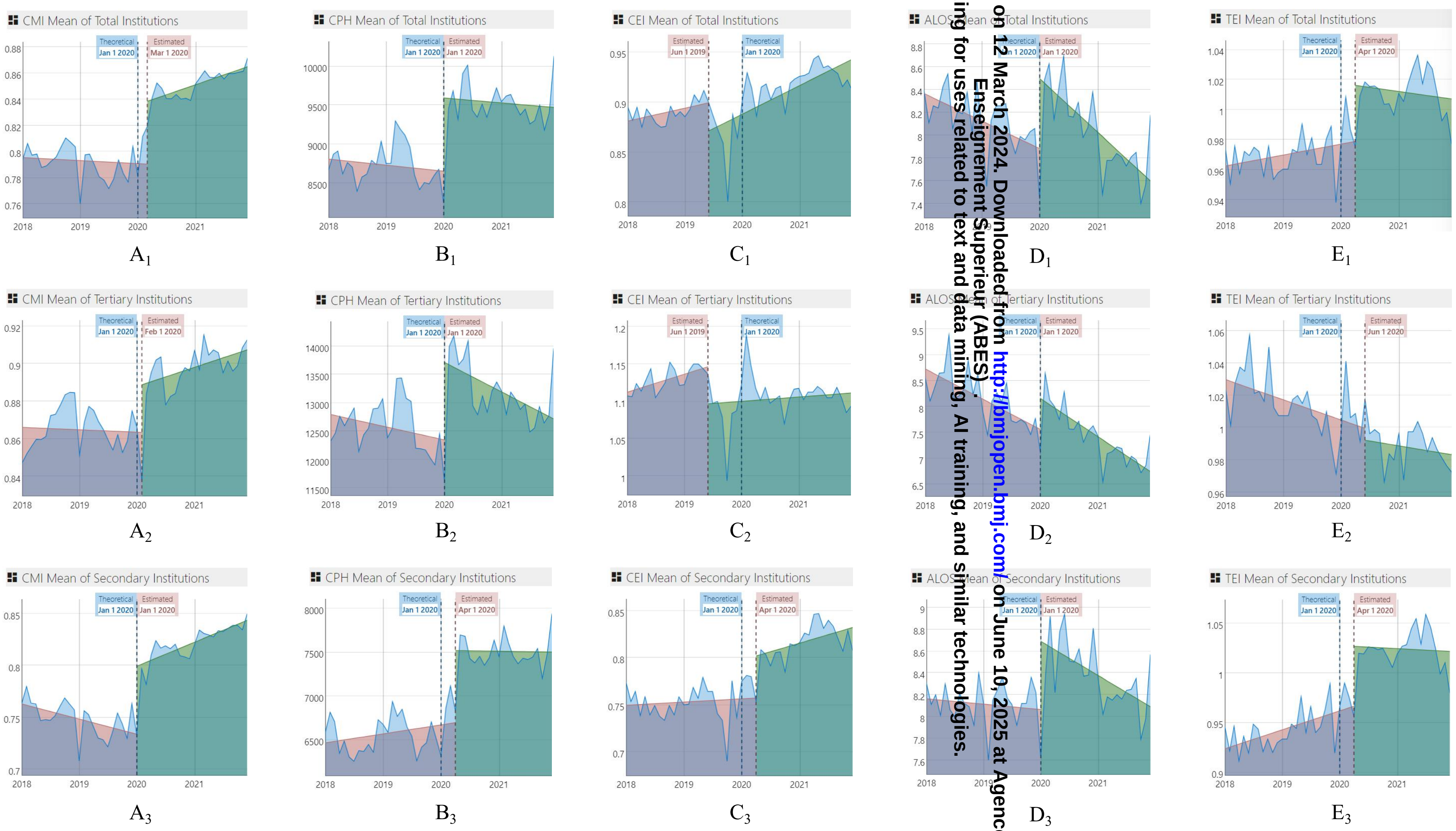


Figure 2. The fitted regression lines for institutions before and after change-points.

Note: Pictures A_i, B_i, C_i, D_i, E_i are the fitted regression lines of outcomes (CMI mean, CPH mean, CEI mean, ALOS mean, TEI mean) in total institutions(i=1), tertiary institutions(i=2), secondary institutions(i=3) respectively. ALOS: average length of stay; CEI: cost efficiency index; CMI: case-mix index; CPH: cost per hospitalization; TEI: time efficiency index.

Appendix A. The R-MITS model

Let y_{it} denote the outcome of interest for unit i at time t, with $i \in \{1,..., N\}$, $t \in \{1,...,n_i\}$, and n_i denoting the time series length for unit i. Then the general regression is defined as

$$y_{it} = \mu_{it} + \varepsilon_{it}$$

where μ_{it} is the mean function and ε_{it} is the stochastic process that model fluctuations around the mean functions and auto-correlation within the time series. The mean function of outcome is

$$\mu_{it} = \begin{cases} \beta^{\tau}_{i0} + \beta^{\tau}_{i1} t, & t < \tau \\ (\beta^{\tau}_{i0} + \delta^{\tau}_{i}) + (\beta^{\tau}_{i1} + \Delta^{\tau}_{i}) t, & t \geq \tau \end{cases}$$

where β^{τ}_{i0} denotes the intercept of the mean function prior to the change-point, β^{τ}_{i1} denotes the slope of the outcome prior to the change-point, $\beta^{\tau}_{i0} + \delta^{\tau}_{i}$ is the intercept of the post-intervention phase, $\beta^{\tau}_{i1} + \Delta^{\tau}_{i}$ is the slope of the post-intervention phase for the outcome in unit i, and τ denotes the global over-all-unit change-point of the response. In the case with only one unit, τ denotes the change-point for that one-time series. If $\delta^{\tau}_{i} + \Delta^{\tau}_{i} = 0$, then there is no change in the mean function of unit i before and after τ. Further model details about the estimation procedure of R-MITS are described in Cruz et al (2021) [37].

Appendix B. The stochastic component parameters of outcomes

Robust-ITS provides changes in the adjacent correlation (correlation between two consecutive time points) and variance of the response. Unit-specific estimates of the stochastic component parameters for every outcome are shown in Appendix Table 1.

As for all institutions, the adjacent correlation of CMI mean is larger and the standard deviation is smaller than before the policy was implemented; the same with the CEI mean performance. The adjacent correlation and standard deviation of the rest of the outcomes did not perform consistently well (either the adjacent correlation was small or the standard deviation was larger). It can be inferred that the change in CMI mean of the total public general hospitals may be more predictable and stable post-intervention, which are a positive result of the DRGs' policy intervention.

As can be seen in the subgroups, only the adjacent correlation of CMI mean in the secondary institutions is larger, and the standard deviation is smaller than before; only the adjacent correlation of CEI mean in the tertiary institutions is larger, and the standard deviation is smaller than before. The adjacent correlation and standard deviation of the rest of the outcomes did not perform well. It can be inferred that the change in CMI mean in secondary public general hospitals and the change in CEI mean of tertiary public general hospitals may be more predictable and stable post-intervention, which are positive results of the DRGs' policy intervention.

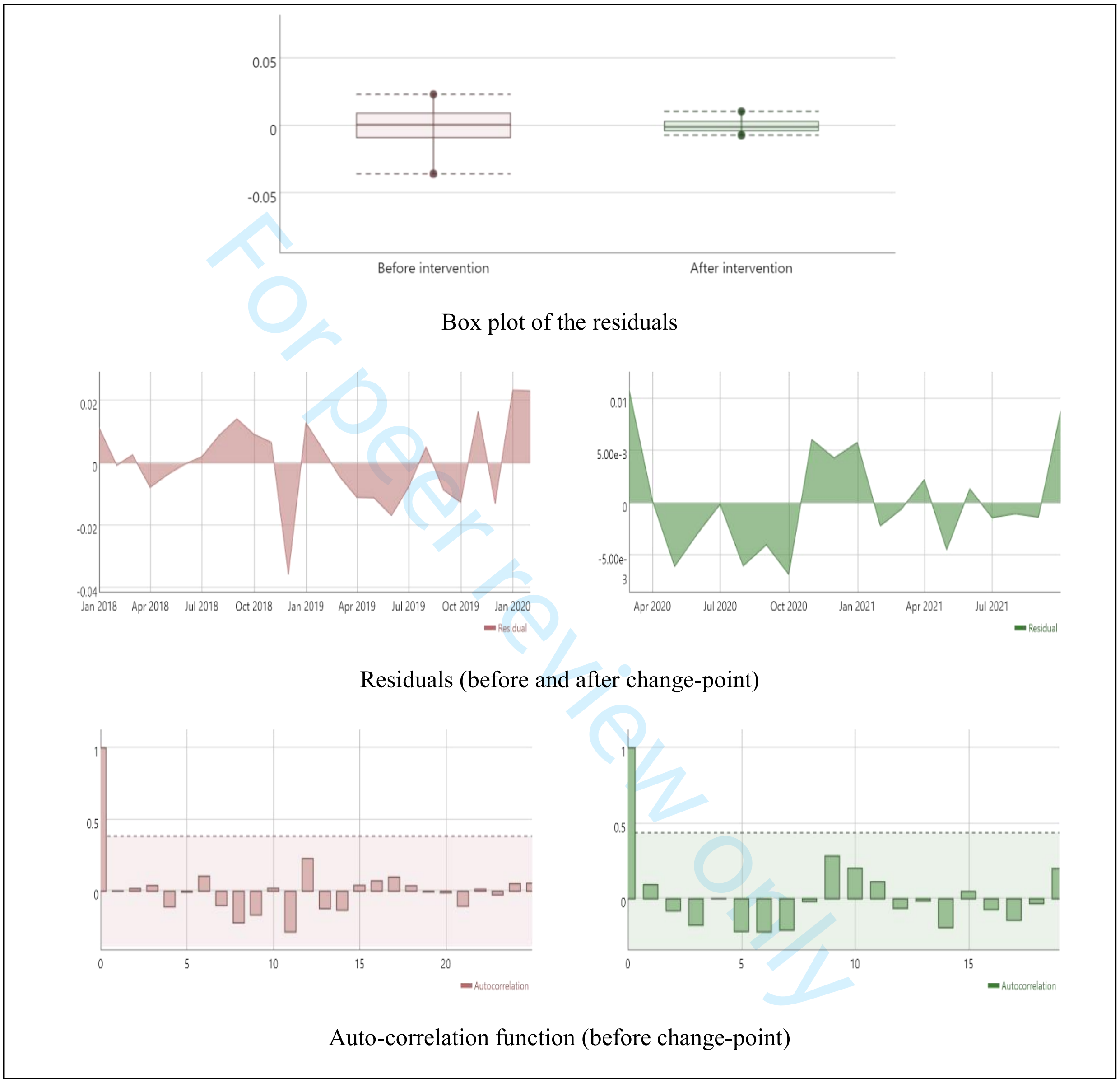
Appendix Table 1. Unit-specific estimates of the stochastic component parameters

Outcomes	Unit	Adjacent correlation			Standard deviation		
		Pre-change	Post-change	Change	Pre-change	Post-change	Change
CMI mean	Total institutions	0.2560	0.5109	↑	0.2139	0.2104	↓
	Tertiary institutions	0.5274	0.2934	↓	0.1859	0.2135	↑
	Secondary institutions	0.0718	0.3359	↑	0.2080	0.2071	↓
CPH mean	Total institutions	0.5988	0.0720	↓	0.1811	0.2700	↑
	Tertiary institutions	0.5246	0.3415	↓	0.1867	0.2683	↑
	Secondary institutions	0.4715	0.1647	↓	0.1734	0.2794	↑
CEI mean	Total institutions	0.0828	0.3093	↑	0.2302	0.1859	↓
	Tertiary institutions	0.0822	0.4342	↑	0.2534	0.1715	↓
	Secondary institutions	0.0986	0.4750	↑	0.1893	0.2293	↑
ALOS mean	Total institutions	−0.0453	−0.0369	↑	0.2302	0.2587	↑
	Tertiary institutions	0.0473	0.0535	↑	0.2140	0.2377	↑
	Secondary institutions	−0.1052	−0.2084	↓	0.2317	0.2296	↓
TEI mean	Total institutions	−0.1202	0.7081	↑	0.1920	0.2118	↑
	Tertiary institutions	0.1382	0.1259	↓	0.1911	0.2539	↑
	Secondary institutions	−0.2559	0.6688	↑	0.1840	0.2265	↑

Note: ALOS: average length of stay; CEI: cost efficiency index; CMI: case-mix index; CPH: cost per hospitalization; TEI: time efficiency index.

Appendix 1 through 5: Residuals and auto-correlation functions for outcomes

Appendix 1: Case-mix index (CMI) mean



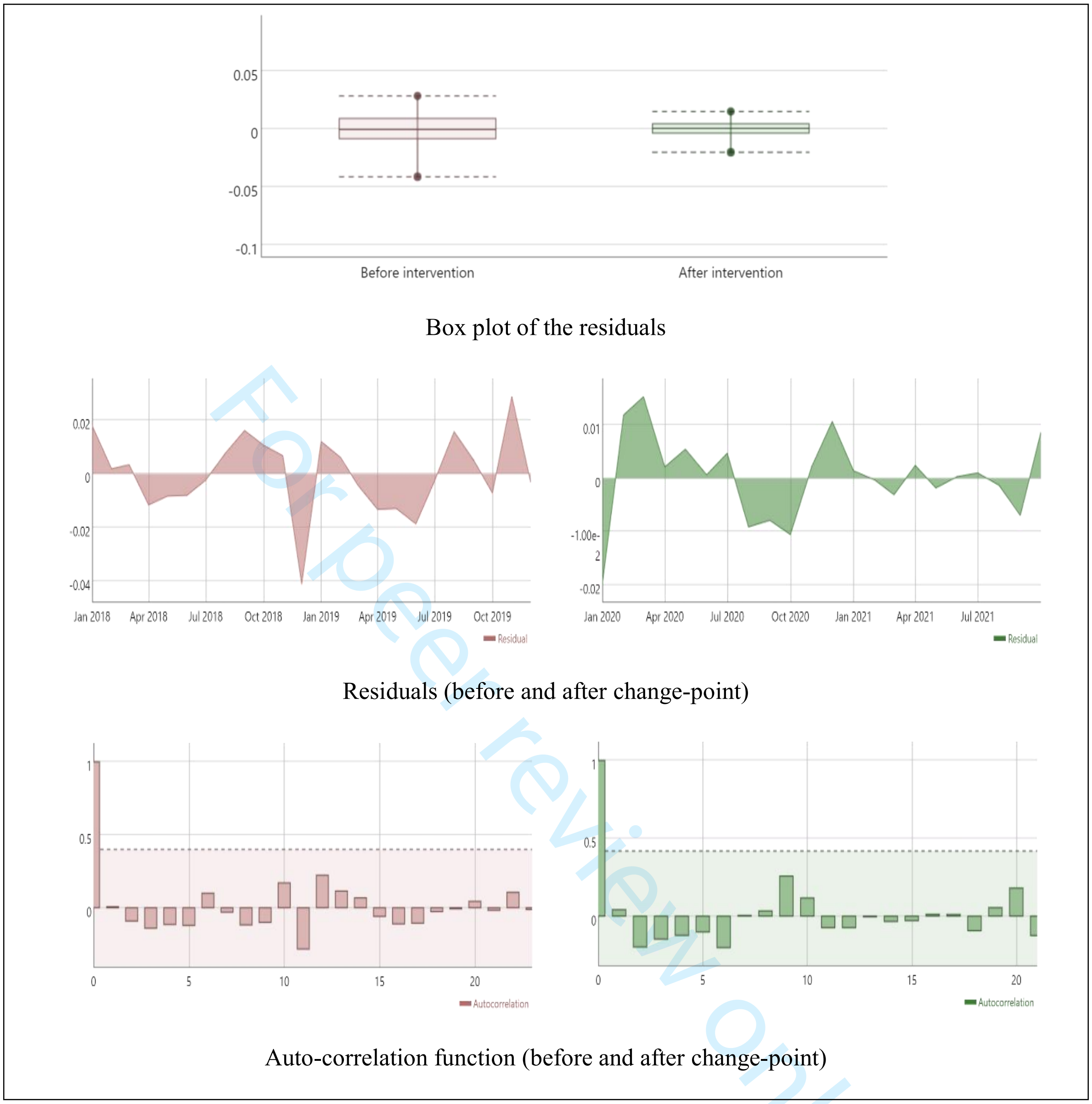
Box plot of the residuals

Residuals (before and after change-point)

Auto-correlation function (before change-point)

Plots 1: CMI mean in total institutions

Plots 2: CMI mean in tertiary institutions



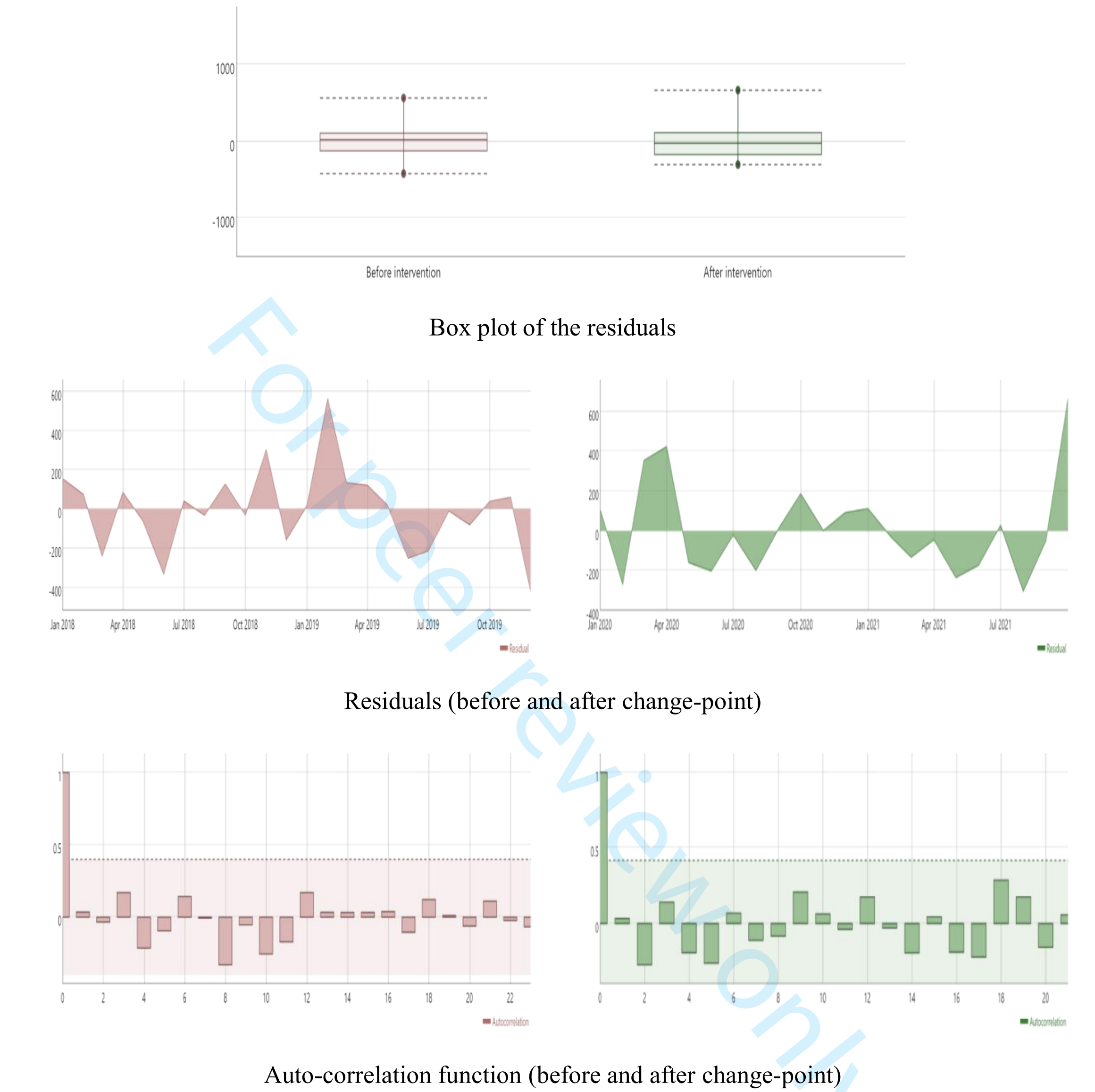
Box plot of the residuals

Residuals (before and after change-point)

Auto-correlation function (before and after change-point)

Plots 3: CMI mean in secondary institutions

Appendix 2: Cost per hospitalization (CPH) mean

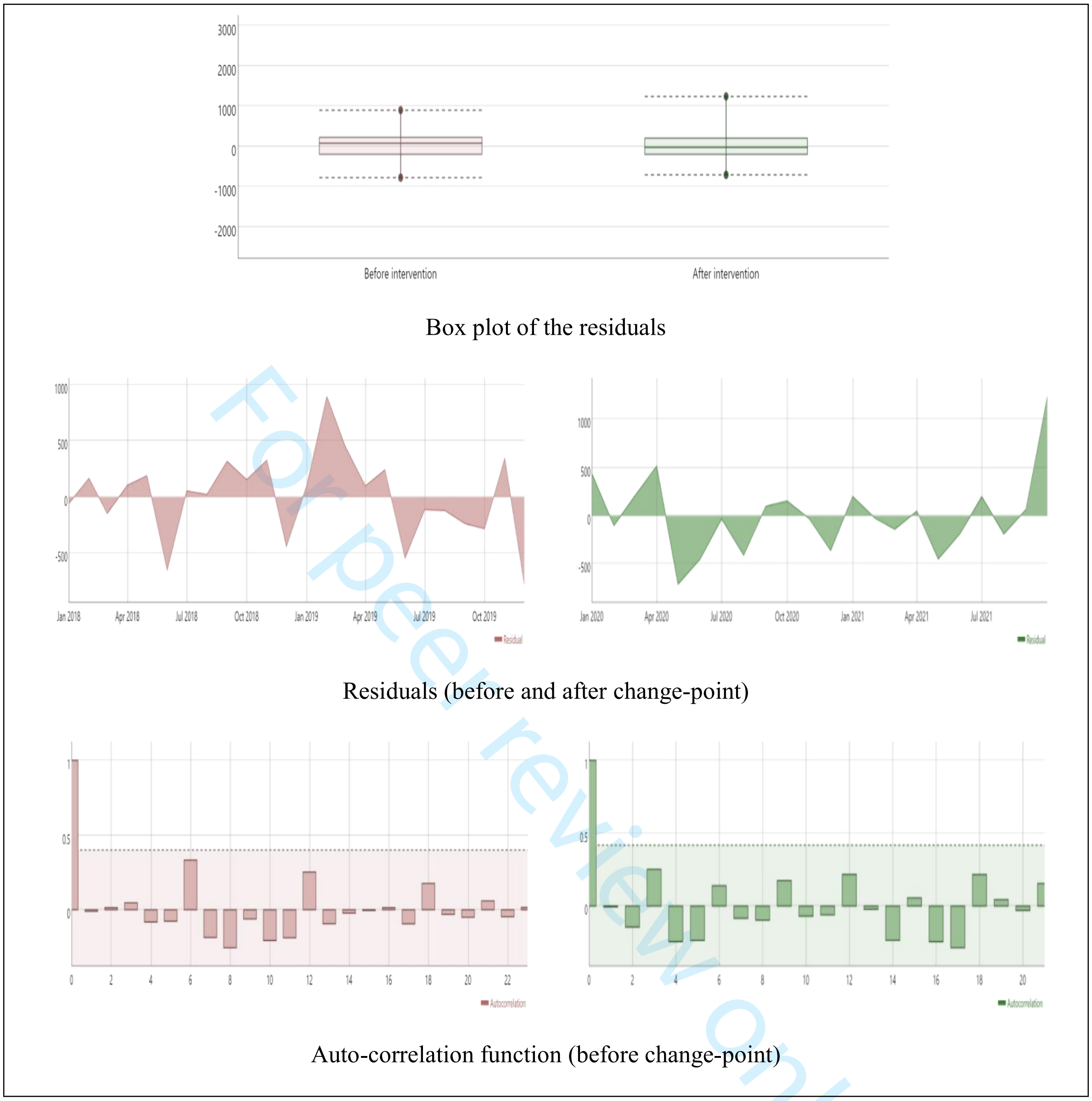


Box plot of the residuals

Residuals (before and after change-point)

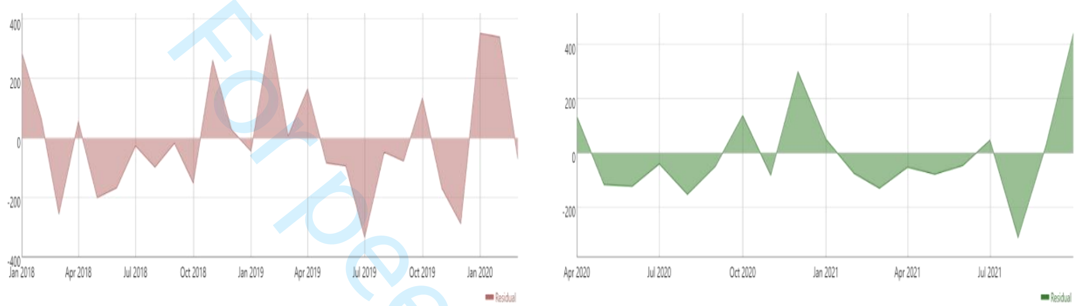
Auto-correlation function (before and after change-point)

Plots 1: CPH mean in total institutions

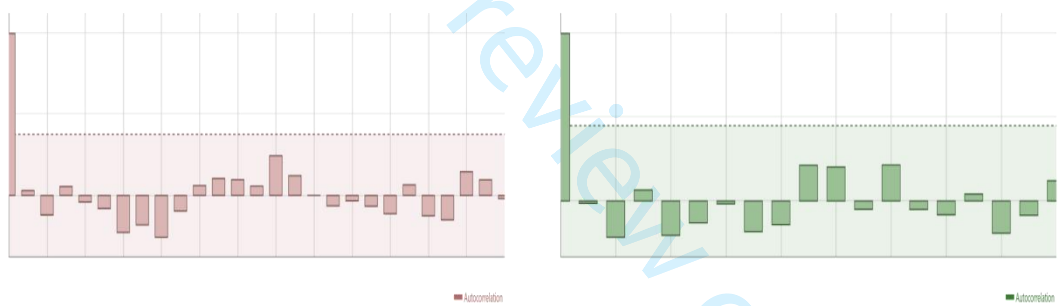


Plots 2: CPH mean in tertiary institutions

Box plot of the residuals



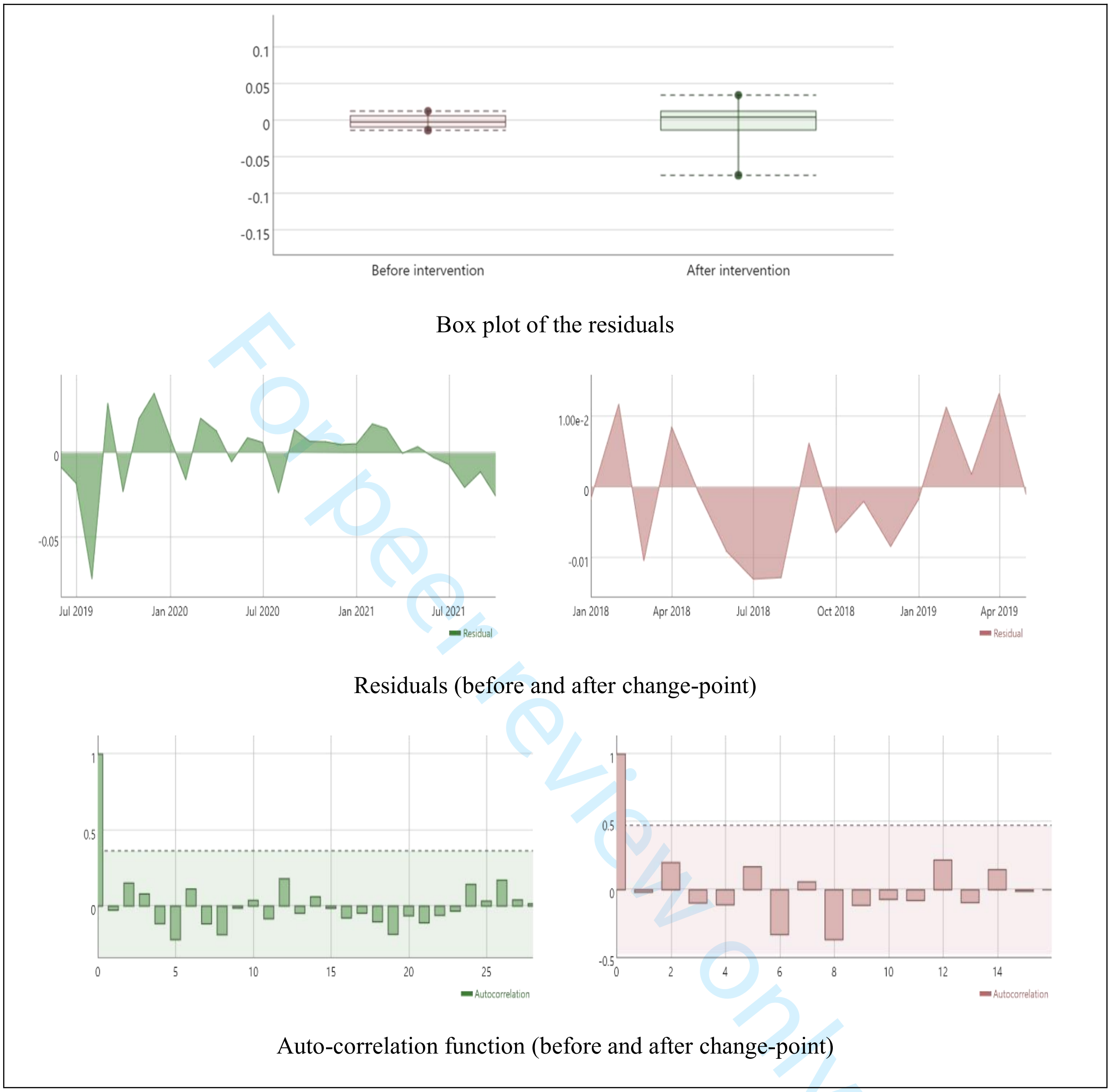
Residuals (before and after change-point)



Auto-correlation function (before and after change-point)

Plots 3: CPH mean in secondary institutions

Appendix 3: Cost efficiency index (CEI) mean

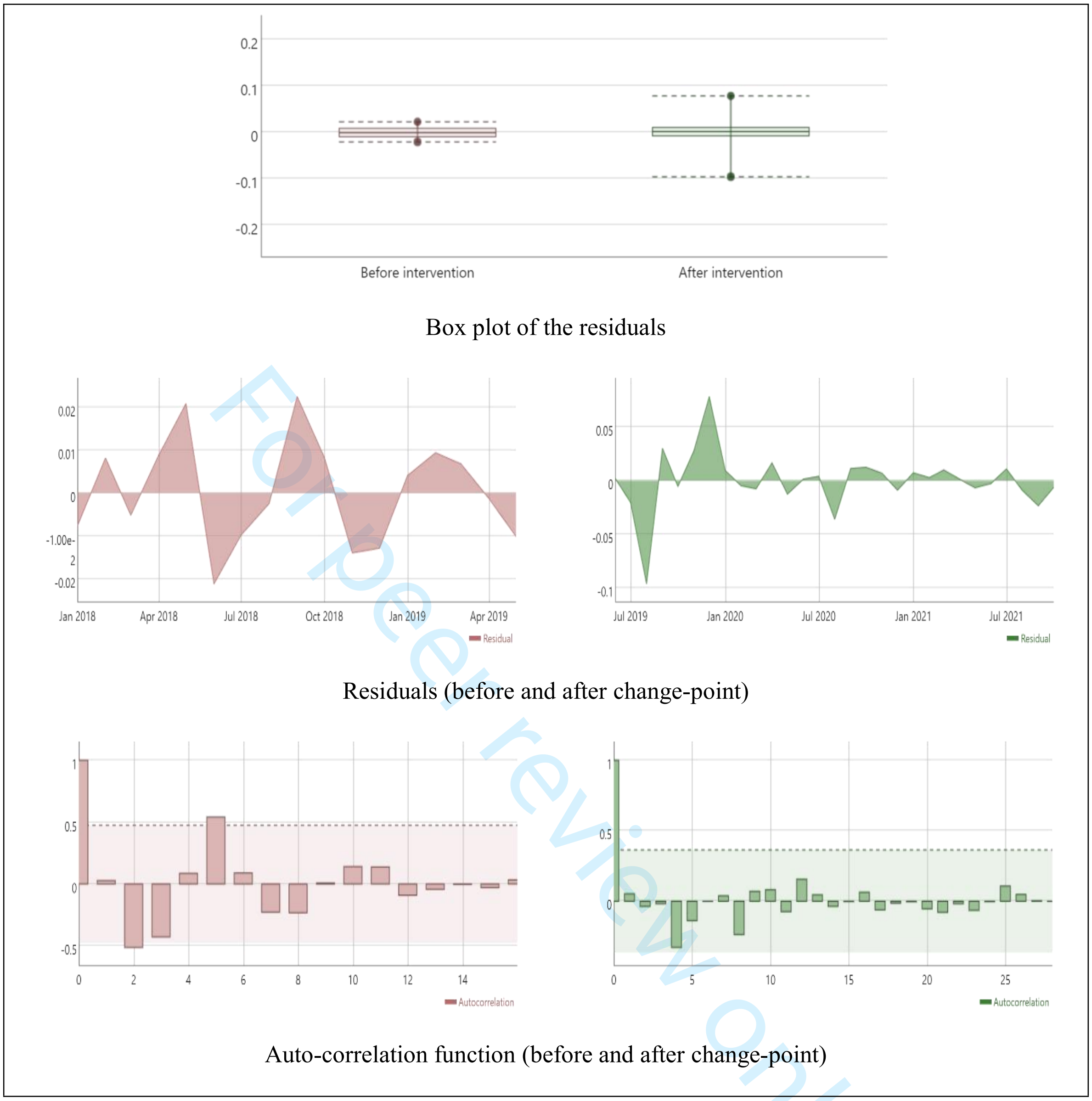


Box plot of the residuals

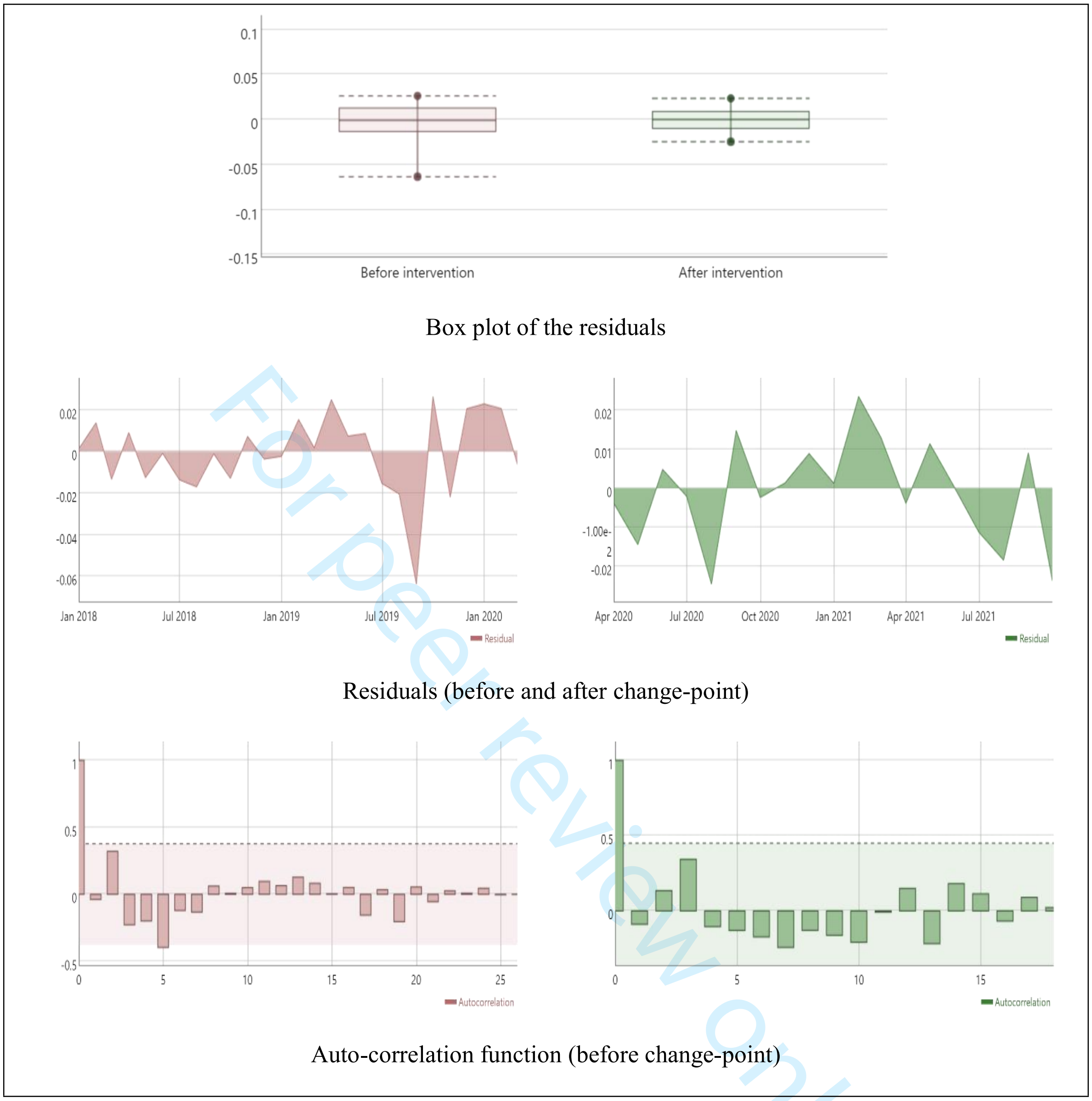
Residuals (before and after change-point)

Auto-correlation function (before and after change-point)

Plots 1: CEI mean in total institutions

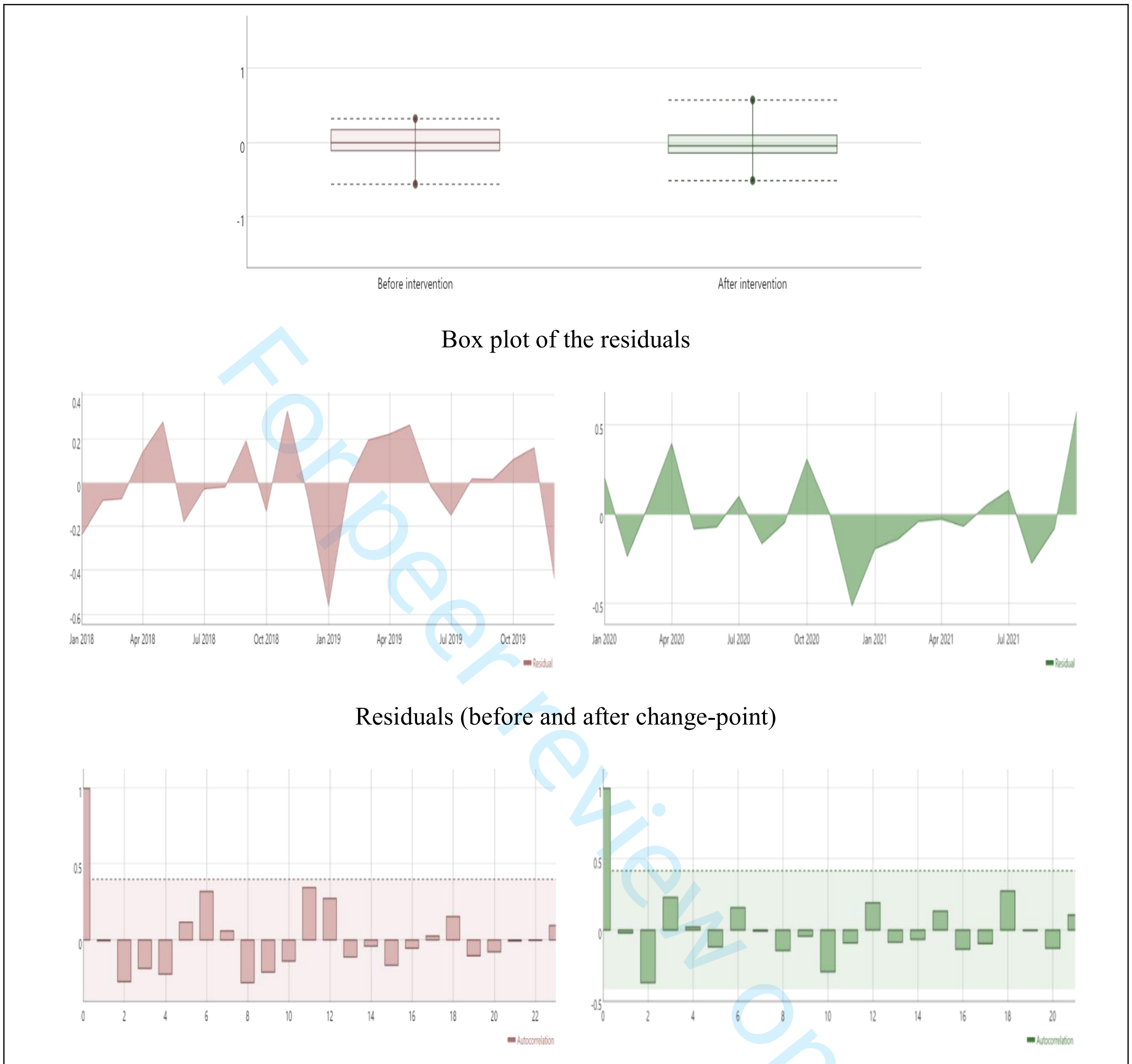


Plots 2: CEI mean in tertiary institutions



Plots 3: CEI mean in secondary institutions

Appendix 4: Average length of stay (ALOS) mean



Box plot of the residuals

Residuals (before and after change-point)

Auto-correlation function (before and after change-point)

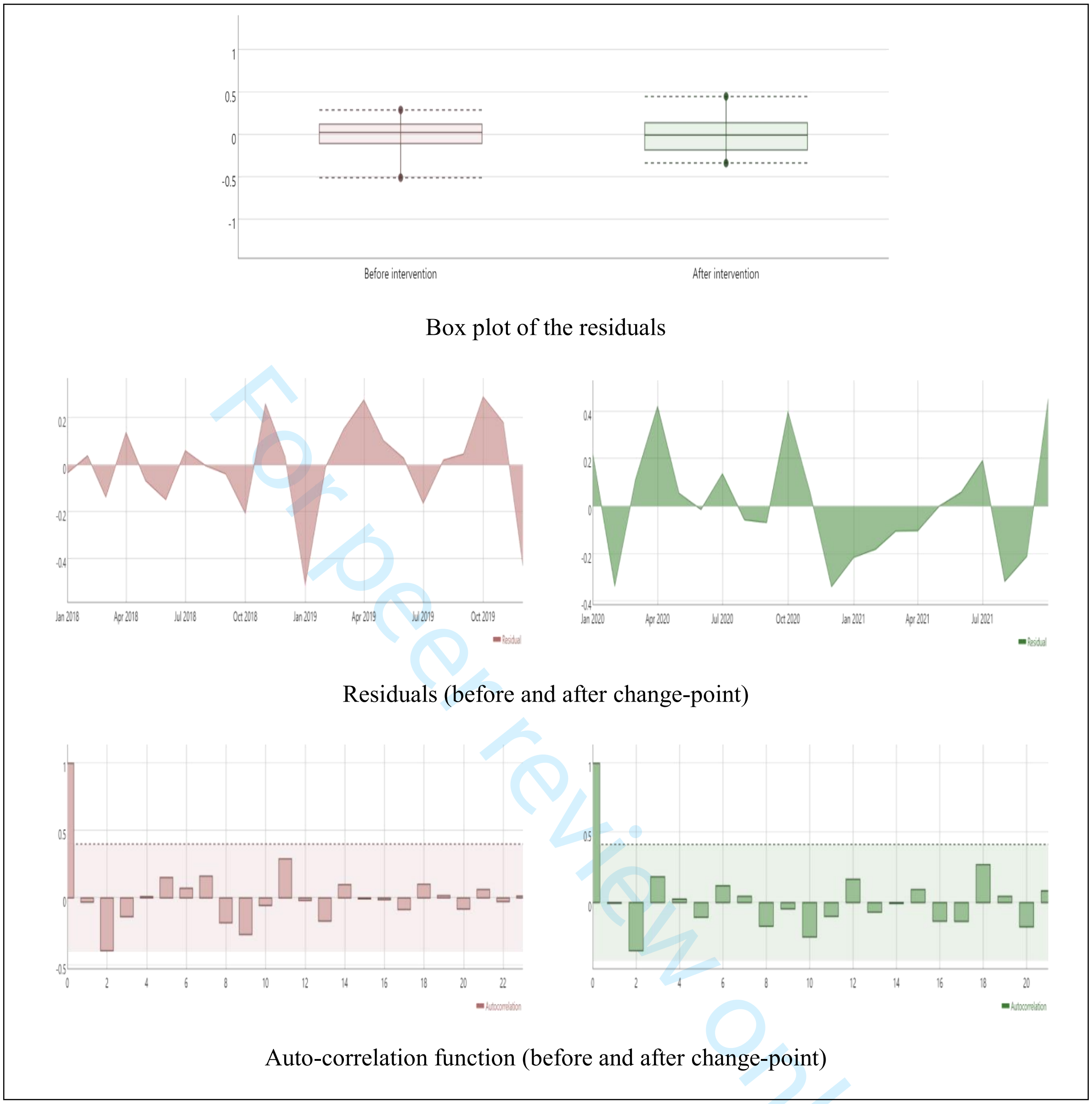
Plots 1: ALOS mean in total institutions

Box plot of the residuals

Residuals (before and after change-point)

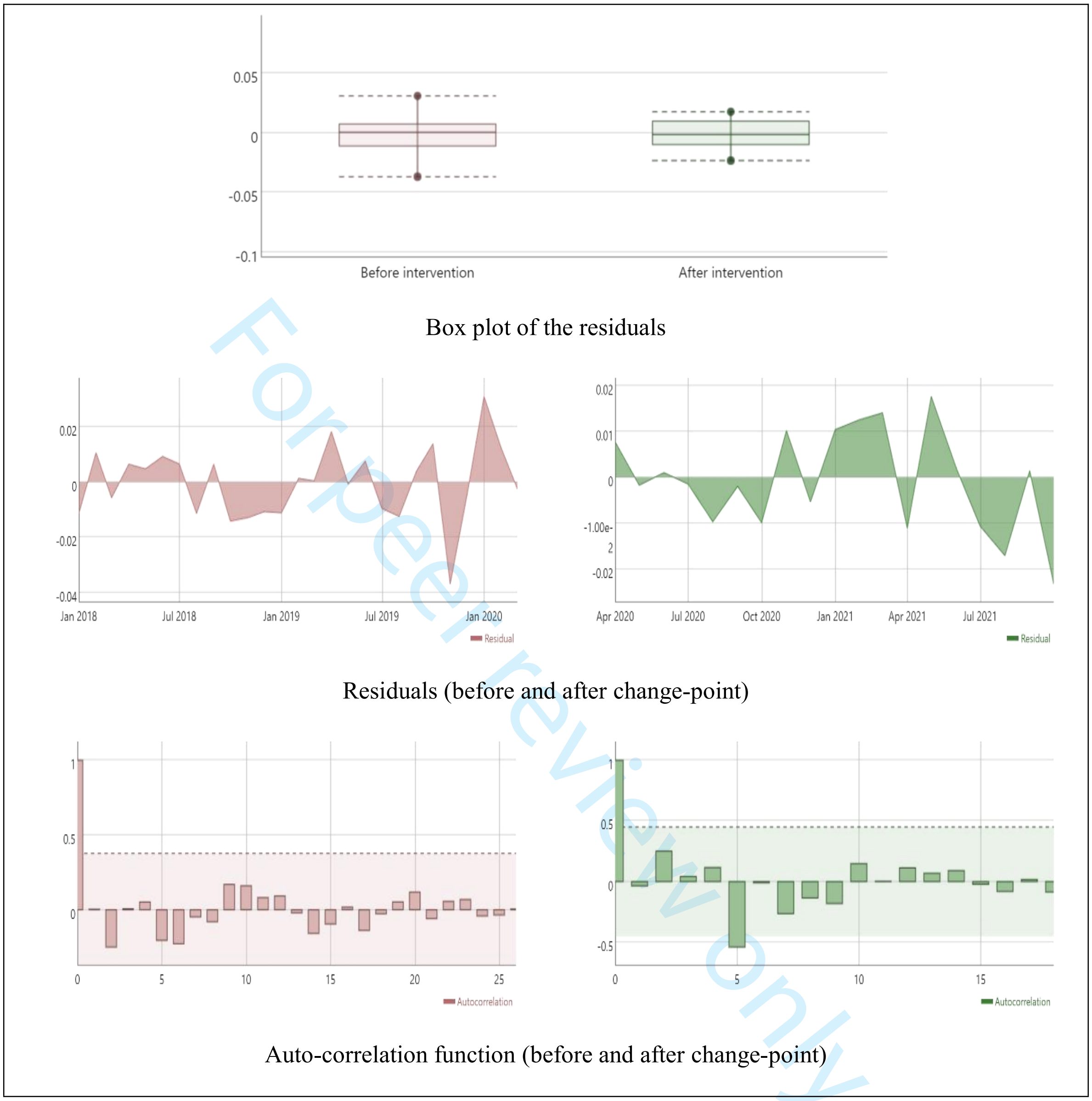
Auto-correlation function (before and after change-point)

Plots 2: ALOS mean in tertiary institutions

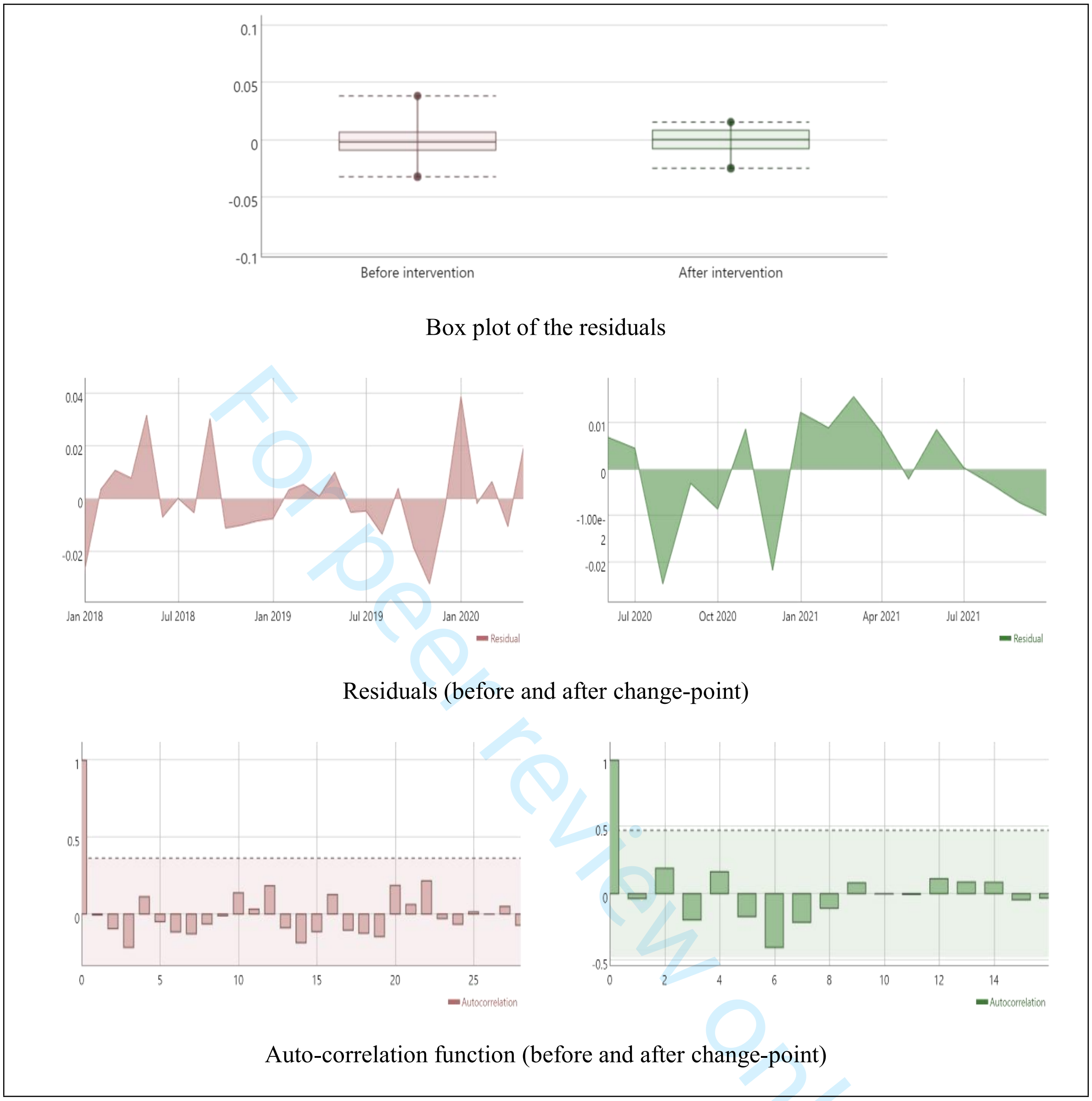


Plots 3: ALOS mean in secondary institutions

Appendix 5: Time efficiency index (TEI) mean



Plots 1: TEI mean in total institutions

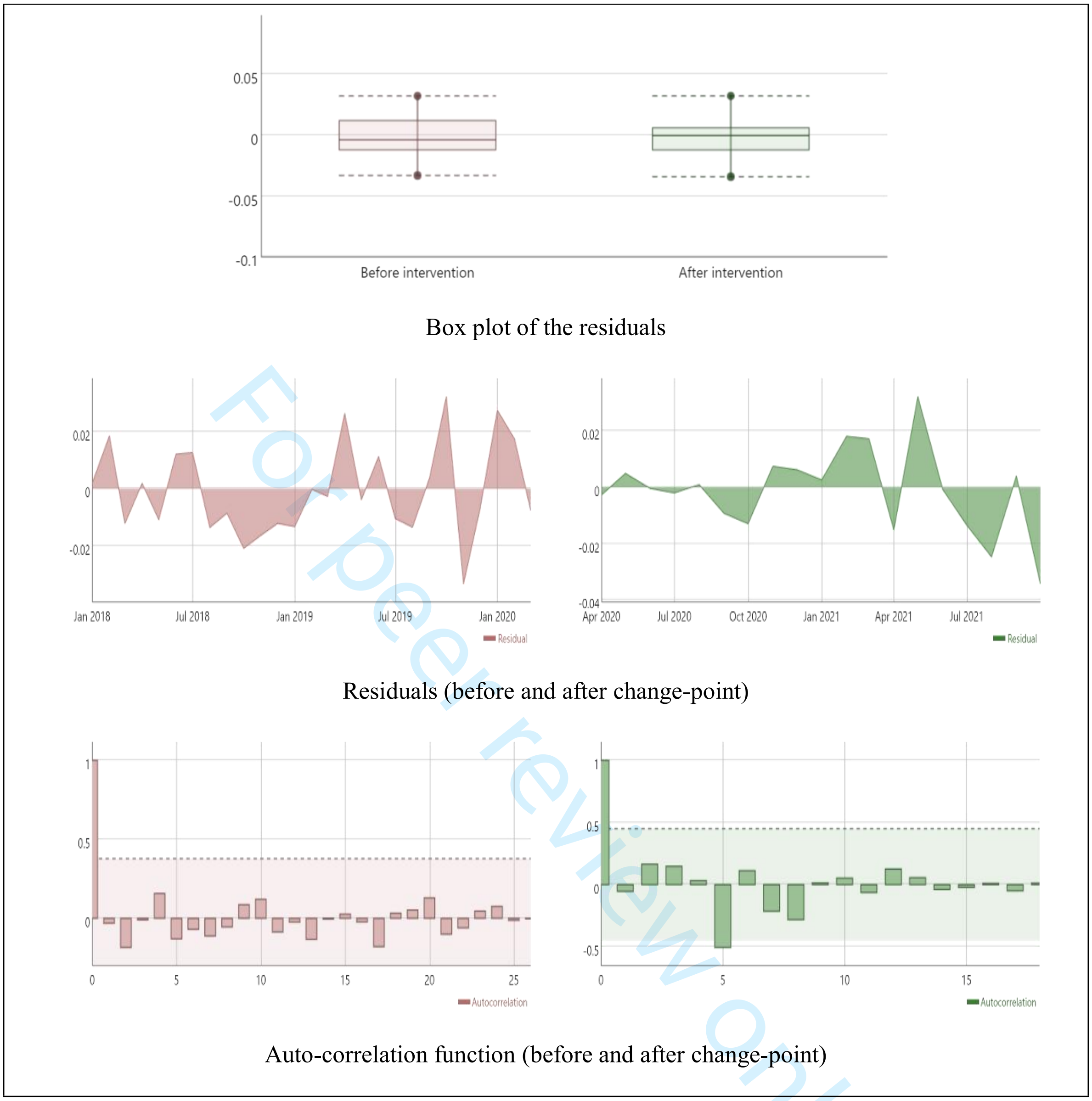


Box plot of the residuals

Residuals (before and after change-point)

Auto-correlation function (before and after change-point)

Plots 2: TEI mean in tertiary institutions



Plots 3: TEI mean in secondary institutions

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No.	Recommendation	Page No.	Relevant text from manuscript
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1	**Title:** Differences in inpatient performance of public general hospitals following implementation of a points-counting payment based on diagnosis-related groups: A robust multiple interrupted time series study in Wenzhou, China **Design**: A longitudinal study using a robust multiple interrupted time-series model to evaluate service performance following policy implementation.
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	1	This study showed a positive effect of the DRGs' policy in Wenzhou, even during COVID-19. It can motivate public general hospitals to improve the comprehensive capacity and mitigate discrepancies in the efficiency in the expense of treating similar diseases. Policymakers are interested in whether the reform successfully stimulates hospitals to strengthen the internal impetus to improve performance, which is supported by this study.
Introduction				
Background/ rationale	2	Explain the scientific background and rationale for the investigation being reported	2-4	Many countries and regions have adopted the DRGs-medical-insurance payment in attempts to contain expenditure and increase the transparency, efficiency, and safety of their healthcare systems. However, the classification rules of DRGs require high-quality medical record home page data and judgment from physicians. Therefore, the roll-out of DRGs has some challenges in middle-and-low-income countries with limited resources. China places great importance on the impact of DRGs' payment on healthcare performance, because it can influence the overall arrangement and progress but the empirical evidence of its impact in China is still inadequate. Zhejiang is the first province to announce enforcement across the entire province; whether the PCP method based on DRGs can exert a strong motivation on hospitals to improve their service performance is important for China but it is still an open topic.

Objectives	3	State specific objectives, including any prespecified hypotheses	3	Above all, the time at which the payment reform was initiated, announced, and publicized, could impact its outcomes. On one hand, it may have encouraged hospitals to change their management approaches to pursue a more positive adaption of the new payment method. On another hand, it is also reasonable to assume that the obvious behavioral changes adopted by hospitals occurred after the official implementation, possibly due to certain problems concerning the requisite technology and operation of the DRGs' payment approach. Those hospitals with relatively weak basic conditions required more intensive policy adaptation, such as training personnel in DRGs-associated tasks and improving familiarity with the diagnostic and operational codes for diseases. It can affect reaction time and degree of importance across different hospitals to some extent; however, previous studies may have neglected to include the time until the policy' effects manifested. Chinese Health Commission divides medical institutions into three levels according to their tasks and functions to form a hierarchical medical system: primary, secondary, and tertiary hospitals.
Methods				
Study design	4	Present key elements of study design early in the paper	4,6	Given that, this study takes Wenzhou in Zhejiang province as a sample city to evaluate hospital performance before and after the implementation of the DRGs' reform between 2018 and 2021 in public general hospitals from the perspective of institution levels. This study used a robust multiple interrupted time series (R-MITS) model, which have several advantages over the traditional ITS designs.
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	4,5	Wenzhou City lies in the southeast of Zhejiang province. The intervention was implemented officially in 2020, the time-series data of these 22 institutions collected spanned from 2018 to 2021. Before implementation: 2018–2019; after implementation: 2020–2021. The data were collected from the Zhejiang Hospital Quality Management and Performance Evaluation Platform, which was developed by the Zhejiang Provincial Health Commission.

Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	inapplicable	Patient and public involvement None.
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	inapplicable	Patient and public involvement None.
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5	This study used five performance indicators calculated by the Zhejiang Hospital Quality Management and Performance Evaluation Platform that reflects the medical capacity and consumption of medical resources. The five performance indicators are case-mix index (CMI), cost per hospitalization (CPH), its cost efficiency index (CEI), average length of stay (ALOS), and its time efficiency index (TEI).
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	5,6	Among them, the CMI, CEI, and TEI derives from the DRG system, which has been found to have excellent homogeneous comparability.31-34 The following are the definitions and calculations of CMI, CEI, and TEI as determined by Zhejiang Provincial Health Commission. 1) CMI reflects the average comprehensive medical ability of a medical institution and its average seriousness or difficulty of the diseases. The calculation method is $\text{CMI} = \dfrac{\text{Total weight}}{\text{Number of cases in region}}$ $\text{Total weight} = \sum \text{Each DRG weight} \times \text{Number of cases in each DRG}$ 2) CEI reflects the cost of treating similar diseases in this region. When the value of

				CEI was close to 1, it suggested that, for the region in question, the cost efficiency was close to the average level; values below 1 or above 1 suggested that the efficiency was above or below the average level, respectively. $$\mathrm{CEI} = \frac{\sum_j k_j^c n_j}{\sum_j n_j}$$ $$\text{Cost ratio}(k^c) = \frac{\text{average cost of a DRG in hospital } (c_i)}{\text{average cost of a DRG in region } (\overline{C}_i)}$$ where n_j represents the number of cases in DRG_j, CEI is the weighted average of k^c. 3) TEI reflects the time spent treating similar diseases in this region. When the value of TEI was close to 1, it suggested that, for the region in question, the time efficiency was close to the average level; values below 1 or above 1 suggested that the efficiency was above or below the average level, respectively. $$\mathrm{TEI} = \frac{\sum_j k_j^l n_j}{\sum_j n_j}$$ $$\text{Average length of stay ratio}(k^l) = \frac{\text{average length of a DRG in hospital } (l_i)}{\text{average length of a DRG in region } (\overline{L}_i)}$$ where n_j represents the number of cases in DRG_j, TEI is the weighted average of k^l. To reflect the differences in performances from the perspective of institutional levels, this study employed the mean of indicator in total institutions, tertiary institutions, and secondary institutions, respectively.
Bias	9	Describe any efforts to address potential sources of bias	7	Furthermore, it does not assume that the intervention is instantaneous and allows the effect of the intervention to be abrupt or gradual, thus allowing the presence of a potentially lagged (or anticipatory) effect. It can obtain more accurate estimates, which further improves internal and external validity, and reduces bias.
Study size	10	Explain how the study size was arrived at	5	Above-primary-level public general hospitals in Wenzhou amounted to 22 institutions, which consisted of 8 tertiary hospitals and 14 secondary hospitals. In 2021, the number of discharged cases in above-primary-level public general hospitals in Wenzhou was 840,342, accounting for 98.79% of the 850,652 discharged cases.

Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	6	To reflect the differences in performances from the perspective of institutional levels, this study employed the mean of indicator in total institutions, tertiary institutions, and secondary institutions, respectively.
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	6	Interrupted time series has become a core study design for the evaluation of health policies. This study used a robust multiple interrupted time series (R-MITS) model, which have several advantages over the traditional ITS designs.
		(*b*) Describe any methods used to examine subgroups and interactions	6	To reflect the differences in performances from the perspective of institutional levels, this study employed the mean of indicator in total institutions, tertiary institutions, and secondary institutions, respectively.
		(*c*) Explain how missing data were addressed	inapplicable	
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	inapplicable	
		(*e*) Describe any sensitivity analyses	7	R-MITS borrows information from all micro-systems to estimate the global change point, that is, it determines the time point at which the effect of the intervention initiates for the entire health system, rather than removing some points as decided by the study team. Given that decreasing outcome variability can contribute to less extreme values, while increasing auto-correlation can contribute to a more predictable outcome, it is important to capture alongside mean differences in variability and auto-correlation. Compared with traditional designs, this R-MITS model can capture all the aforementioned time series' changes.
Results				

Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	inapplicable	
		(b) Give reasons for non-participation at each stage	inapplicable	
		(c) Consider use of a flow diagram	inapplicable	
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	inapplicable	
		(b) Indicate number of participants with missing data for each variable of interest	inapplicable	
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	inapplicable	
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	inapplicable	
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	inapplicable	
		Cross-sectional study—Report numbers of outcome events or summary measures	inapplicable	
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	9	Table 2. Unit-specific changes estimated by R-MITS pre- and post-intervention
		(*b*) Report category boundaries when continuous variables were categorized	inapplicable	

		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	inapplicable	
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	8	Estimated changes using R-MITS for all samples and subgroups are shown in Table 2; the fitted regression lines are shown in Figure 2; the estimates of the stochastic component parameters are shown in Appendix B; the residuals of the estimated model and auto-correlation functions are as expected and are shown in Appendix 1 through 5, demonstrating that the curves for both the pre- and post-groups for the measures included are both linear and independent.
Discussion				
Key results	18	Summarise key results with reference to study objectives	9	With limited and stretched healthcare resources, it is important for hospitals to further improve the performance of their medical services after the new policy implementation. In general, the results from all samples and subgroups seem positive from various perspectives.
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	11	The findings presented here should be generalized with caution. The impact of COVID-19 reached Zhejiang province (to where Wenzhou is located) at the end of January; fortunately, the setback was temporary due to the implementation of strict prevention and control measures, and recovery gradually occurred after February 2020. Nevertheless, it may have unavoidably resulted in decreasing patient visits and hospital revenues, while increasing operation costs and medical efficiency, which may hinder the improvement of these selected indicators. We could have minimized the disruption of the COVID-19 by removing data from the early years of the policy, but we did not do so because it would have been too subjective. The R-MITS model can identify inferences regarding the estimation of global change points across units, rather than removing some series as decided by the study team. In such an external environment, the implementation of the DRGs' payment would be timely and may make up for the negative impact of the epidemic, which could motivate cost-control consciousness and management behaviors. Because under the motivation of DRGs' payment, hospitals have to make adjustments to strengthen service performance, otherwise the revenue loss would be large. Theoretically, the policy can guide high-grade-level hospitals to focus on the treatment of more complex and severe diseases

				and the breakthrough of medical service level, while returning routine and common diseases to low-grade-level hospitals for treatment to improve the CMI indicator. In addition, motivating hospitals to improve medical levels and efficiency can promote other indicators. This study showed that the DRG policy had positive effects on hospital performance, even during COVID-19.
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	11,12	Regarding the analysis method, we were unable to apply the difference-in-differences, because the basic medical insurance designated medical institutions of Zhejiang province were all required to implement DRGs' payment, including Wenzhou. In addition, we could not obtain data from the public general hospitals outside Zhejiang province where the DRGs or other payment reforms (like Diagnosis-Intervention Packet payment) were not implemented, under the process of the nationwide payment reform. Fortunately, R-MITS is an optimized version of ITS, whose strengths can enhance the relationship between the evolution of results and the DRG policy intervention. Besides, the subject of this study were public general hospitals; however, we collected data from secondary and tertiary public general hospitals, excluding primary hospitals that did not have time-series data of DRG-related indicators pre- and post-intervention. DRGs' indicators can homogeneously compare the medical performance, but their calculations are more dependent on the quality of the homepage of inpatient medical records than traditional indicators, especially disease coding quality. Thus, previous studies that used the DRGs' indicators were almost conducted in above-primary-level hospitals, include this study. CMI reflects the average level of medical technology and complexity of the diseases attended to by each hospital; CEI or TEI reflect the relative efficiency of CPH or ALOS, which are calculated based on DRGs. This evaluation was conducted close to the intervention's time, which might not produce the same results as other evaluations carried out at a later date.
Generalisability	21	Discuss the generalisability (external validity) of the study results	12	Despite the inevitable impact of the COVID-19 on hospitals, the design of this study succeeded in revealing some anticipated and positive effects of the DRGs' policy on hospitalization performance. Implementing the policy can motivate public general hospitals to improve the medical comprehensive capacity and, mitigate their efficiency discrepancies in the expense of treating similar diseases across the different grade institutions. Hospitals firstly focused on strengthening hospitalization

				efficiency, though there were differences between tertiary and secondary hospitals when responding to DRGs' policy. One of the original intentions of the reform is to stimulate hospitals to form the internal impetus to improve performance, changing the rude and less-effective development mode of the past. From this perspective, our results further supported the target of the reform in China.
Other information				
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	12	This research was sponsored by Zhejiang Provincial Medical and Health Science Technology Plan Project (2023KY889) from Zhejiang Provincial Health Commission in China.

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.