Social Science & Medicine 289 (2021) 114415

Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

The pilot of a new patient classification-based payment system in China: The impact on costs, length of stay and quality

Mengcen Qian [a,b,1], Xinyu Zhang [a,1], Yajing Chen [a], Su Xu [c], Xiaohua Ying [a,b,*]

[a] School of Public Health, Fudan University, Shanghai, China
[b] Key Laboratory of Health Technology Assessment (Fudan University), Ministry of Health, Shanghai, China
[c] Shanghai Medical and Health Development Foundation, Shanghai, China

ARTICLE INFO

Keywords:
Hospital financing
Patient classification-based payment
Inpatient costs
Quality of care

ABSTRACT

With the urgent need to regulate provider behaviors, China developed a novel patient classification with global budget payment system, expecting to achieve both easy implementation and cost containment. The new system, called "diagnosis-intervention packet (DIP)" payment, is based on a deterministic patient classification approach, which groups patients according to the combination of principal diagnosis ICD-10 (International Classification of Diseases, 10th Revision) codes and procedure ICD-9-CM3 (International Classification of Diseases, 9th Revision, Clinical Modification) codes and links each group to relative historical costs market-wide. This study investigated the impact of the DIP-based payment on inpatient costs, length of stay, and quality of care in the largest DIP pilot city of China. In 2018, the city changed from the "fixed rate per admission with a cap on annual total compensation" policy to DIP with global budget for all insured inpatients. A difference-in-differences approach was employed to identify changes in outcome variables before and after the DIP policy among insured relative to uninsured patients. We found an average of 8.5% ($p = 0.000$) increase in inpatient costs per case (as intended), trivial changes in length of stay, and a 3.6% ($p = 0.046$) reduction in postoperative complication rate in response to DIP adoption among patients with high severity. Our findings suggested that the DIP-based payment helped regulate provider behaviors when treating high-risk patients. And the new payment has the potential for rapid rollout in resource-limited areas where lack a uniform coding practice or high-quality historical data.

1. Introduction

Financing hospitals using a patient classification-based system has been widely recognized as a means of incentivizing efficiency improvement and containing cost from the healthcare provider side (Mihailovic et al., 2016). In such payment systems, fixed prices are defined for patient groups that are classified based on similar clinical characteristics and resource utilization intensity. The diagnosis-related groups (DRGs)-based payment, developed and first implemented in the 1980s (Fetter et al., 1980), is the most popular practice of its kind. The assignment of a DRG group is mainly based on diagnosis, while taking procedures and patient demographics into consideration. Early practices of the DRG-based payment are mainly from developed countries with rich resources (Busse et al., 2013). In recent years, with an urgent need to control soaring medical costs, middle-and-low income countries with limited resources also show great interest in a similar payment policy (Mathanuer and Wittenbecher, 2013).

Since 2009, the Chinese government has shown a firm intention of exploring alternative payment approaches to fee-for-service (FFS). Supply-side cost control policies were developed and adopted to combat rising medical spending in various settings. Examples of these policies include "flat-rate per admission" and "fee-limit per episode", which were recently criticized for the unintended consequences of premature discharges and decreased medical activities (Chan and Zeng, 2018). Meanwhile, China was one of the earliest developing countries to introduce the DRG-based payment (Jian et al., 2015). Early attempts at the use of DRG payment were applied on a limited scope by assigning flat fees to simple diseases that generally do not cause complications, such as non-purulent appendicitis, gallstones, and caesarean section. In 2012, the capital city, Beijing, launched the first broad adoption of DRG

* Corresponding author. Fudan University, School of Public Health; Key Laboratory of Health Technology Assessment (Fudan University), Ministry of Health, China, 130 Dong'an Road, Shanghai 200032, China.
E-mail address: xhying@fudan.edu.cn (X. Ying).

[1] Mengcen Qian and Xinyu Zhang contributed equally to this work.

https://doi.org/10.1016/j.socscimed.2021.114415
Received 18 March 2021; Received in revised form 3 September 2021; Accepted 17 September 2021
Available online 20 September 2021
0277-9536/

payment with a global budget (Jian et al., 2019).

Although in China the DRG-based payment scheme has proved effective in reducing medical expenditures (Jian et al., 2015), its rollout and implementation has been slow due to several challenges in patient classification. First, the DRG classifies patients into a number of manageable groups based on a coarse-to-fine procedure (see Table 1 for a summary on the DRG patient classification rules), usually consisting of hundreds of classes. However, variations in hospital coding practice across areas and incomplete discharge records in underserved regions hindered the ability of creating broad categories of clinically similar patients. Besides, changes in treatment norms and improvements in data quality can be a slow process. Second, the formation of the DRGs requires the judgment of a physician. The lack of consensus on clinical pathways across areas has made the classification resource-consuming and inflexible for future changes (Ying, 2021). These challenges are also commonly and widely shared among middle-and-low-income countries with limited resources.

In the spirit of regulating medical costs and provider behaviors with easy implementation and rapid rollout, a new patient classification-based payment system was developed and piloted concurrently in China. The system is called "Diagnosis-Intervention Packet" (DIP)-based global budget payment. It relies on a different and deterministic approach to classify patients according to direct combinations of principal diagnosis codes (first four digits) and procedure codes, resulting in a more finely defined classification of more than 10,000 groups. Based on historical cost data, relative weights are assigned to patient groups to reflect market-wide (typically, at the prefecture-city level) resource utilization relative to different groups. Relative weights are then converted to payments according to the global insurance budget of the market. Fig. 1 sketches the design of the DIP payment system. A detailed introduction of the policy was provided in section 2.1.

The roll-out of the DIP payment system has been rapid in China, partly attributable to its unique feature in patient classification (Ying, 2021). Table 1 presents a comparison between the DRG and DIP classification. With lower data quality requirements and less dependency on physician judgement, the DIP classification circumvents the aforementioned major challenges in the DRG implementation. The DIP was proposed several years later after the first adoption of DRG, but there were more than twice as many cities piloting the DIP than the DRG payment system across China by the end of 2020 (Fig. S1 in supplementary materials).

Table 1
A comparison on patient classification between the diagnosis-related group (DRG) and diagnosis-intervention packet (DIP).

Dimensions	DRG	DIP
(1)	(2)	(3)
Patient classification rules	First, dividing 26 Major Diagnostic Categories (MDC) based on organ system or etiology and clinical specialty. Then, dividing each MDC into medical and surgical categories based on principal diagnosis or surgical procedure codes. Finally, taking into account patient characteristics, complications and comorbidities and etc. to define DRGs	Combinations of principal diagnosis ICD-10 codes and procedure ICD-9-CM codes.
Physician judgement	Involved	Almost none; data driven
Number of groups	Hundreds, less than 1000	More than 10,000
Number of principal diagnosis in a group	Multiple	1 for most groups
Number of principal procedure codes in a group	Multiple	1 for most groups
Requirements on data quality	Higher than DIP	Lower than DRG
Separate groups for new and advanced technologies	No	Yes

Notes: ICD-10 stands for International Classification of Diseases, 10th Revision; ICD-9-CM3 stands for International Classification of Diseases, 9th Revision, Clinical Modification.

Although designed to address the unintended consequences of previous supply-side cost control policies, no studies have empirically investigated the impact of the DIP-based payment on costs, length of stay, and quality of care. Previous literature that examined DRG-based payment obtained mixed findings on costs, partially due to various institutional backgrounds before the policy change: studies from the United States and China documented a decline in inpatient costs associated with the shift from FFS to DRG-based payment (Jian et al., 2015; Liu et al., 2021; Mikkola et al., 2002; Magnussen et al., 2006; Yuan et al., 2019; Zhang et al., 2016); in contrast, studies in European countries reported an increase in inpatient costs in response to the change from global budgets to DRG-based payment (Farrar et al., 2009; Mateus, 2011). However, findings from the literature are more consistent regarding other measures: they generally report a negative impact on length of stay (Choi et al., 2019; Epstein et al., 1991; Tan and Melendez-Torres, 2018) and a negligible (Busse et al., 2006) or negative effect on quality of care (Kutz et al., 2019).

This study aims to understand the impact of the DIP-based global budget payment system on inpatient costs, length of stay, and quality measures in China. We chose the earliest and largest city that has broadly adopted the DIP system at the whole city level as our study site. In 2018, the study city moved from the "fixed rate per admission with a cap on annual total compensation" policy to the DIP-based global budget payment for all socially insured inpatients. We used a difference-in-differences approach to identify changes in outcomes before and after the policy change among insured relative to uninsured patients.

2. Methods

2.1. Settings and institutional background

The study city is one of the largest (2018 population of 14.90 million) and most economically developed cities in Southeast China. There are two pillars of the social insurance system in the study city: the urban employee basic medical insurance scheme (UEBMI), and the urban and rural resident basic medical insurance scheme (established since 2015 by integrating the original urban resident basic medical insurance scheme (URBMI) with the new cooperative medical scheme (NCMS)).

In 2015, the study city adopted a strict cost control policy that assigned a fixed rate for each admission and a ceiling of annual compensation for insured patients to each hospital that accept social insurance. The rates and ceilings were hospital-specific. Although the approach managed to contain medical spending, it also has raised concerns about selection of healthier patients, premature discharges, and insufficient medical activities.

On January 1, 2018, the city moved from the "fixed-rate per admission with a cap on annual compensation" policy to the DIP with global budget payment. Using over 8 million discharges over the past three years, patient classification was developed and the relative resource utilization intensity of each group was calculated. Based on combinations of the first four digits of the principal diagnosis ICD-10 (International Classification of Diseases, 10th Revision) code and procedure ICD-9-CM-3 (International Classification of Diseases, 9th Revision, Clinical Modification) codes, a total of 12,030 DIP patient groups were created (Xu et al., 2020). To establish a relationship between a DIP group and the medical resources that it consumes at the market level, relative weights (RW) for each group were calculated to represent average costs per admission of all hospitals relative to a pre-specified

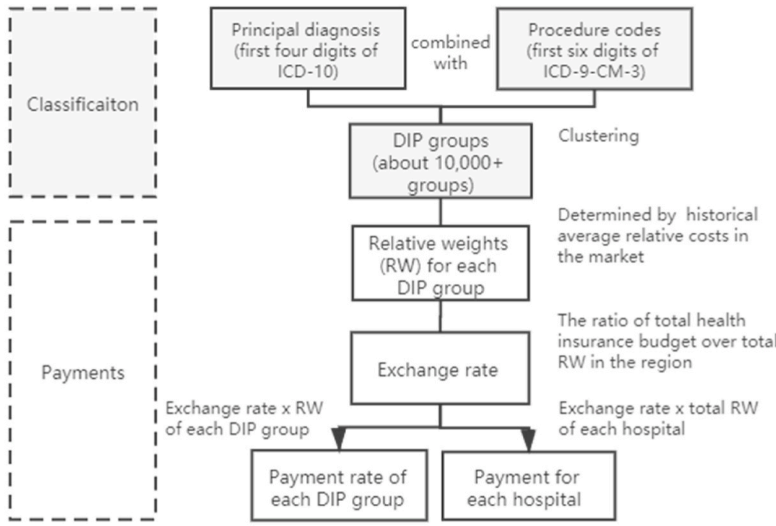

Fig. 1. The design of the diagnosis-intervention packet (DIP) based payment with global budget system. Notes: ICD-10 stands for International Classification of Diseases, 10th Revision; ICD-9-CM3 stands for International Classification of Diseases, 9th Revision, Clinical Modification; RW stands for relative weights, that reflects the relative medical resource consumption market-wide of each group.

reference group (i.e. acute appendicitis with laparoscopic appendicectomy) (Zhang et al., 2020). The total health insurance budget of the city was then used to determine the exchange rate of RW for payments. In such a system, treating clinically similar patients are rewarded with similar payments; treating patients who consume more medical resources receives higher rewards, which is a distinguishing feature from previous policy.

There were two other cost-relevant policies that coincided with the introduction of DIP-based payment in the study city. On July 15, 2017, the zero mark-up drug policy was adopted, which removed mark-ups on drugs in all public hospitals. On December 29, 2018, the zero mark-up on medical consumable supplies was further adopted. Since both polices led to a direct effect on inpatient costs, we used a different study window in our following analysis when inpatient costs were used as the outcome.

2.2. Data

Our data source was de-identified patient-level discharges from January 2016 to December 2019, collected from all hospitals in the study city. Each record provided us information about the hospital name, diagnosis and procedure codes, medical costs, admission and discharge dates, and patient's characteristics (age at admission, gender, and insurance type). We obtained hospital locations (districts) from the Health Insurance Bureau.

2.3. Measures

We used the log form of inpatient costs per case and the level form of length of stay as the outcome variables. We also investigated a set of quality measures: for patients who received surgical procedures, we generated indicators for operation associated infection (occurrence of T81.4 and O86.0 in the first four digits of any secondary diagnoses) and postoperative complication (occurrence of T81.0, T81.1, T81.3, T81.7, T81.8, T81.9, O70, and O71 in the first four digits of any secondary diagnoses) based on ICD-10 codes. Though in-hospital mortality is a widely discussed quality measure, we deliberately did not select it as an outcome. Because the measure captures extreme cases and thus is less sensitive to variations in quality when all diseases are included in the analysis.

We defined treatment a dummy equal to one if the patient is covered by social health insurances, and zero otherwise. Other control variables included indicators for age groups (in 10-year intervals), gender (male, female), dummies for the first three digits of the principal diagnosis code, dummies for the Charlson Comorbidity Index (CCI) (0,1,2, and 3), and interactions between the above variables.

2.4. Statistical analysis

We employed a difference-in-differences approach to identify the impact of the DIP payment introduction on costs, length of stay, and quality of care. For patient *i* admitted to hospital *h* in time *t*, we consider

$$Y_{iht} = \beta Treatment^{*}Post + \gamma X_{iht} + H_h + \tau_t + \varepsilon_{iht}, \tag{1}$$

where Y_{iht} is the outcome variables; *Post* is a dummy equal to one for discharges after January 1, 2018; X_{iht} is a rich set of individual characteristics including age, gender, insurance types, principal diagnosis, CCI, interaction of age and gender, and interaction of gender and CCI; H_h is the hospital fixed effects; τ_t is the interaction between hospital location and year-by-month fixed effects, capturing area-specific differential time effects; and ε_{iht} is the error term. Since we included indicators for insurance types as a more extensive form of the treatment dummy, the term *Treatment* is not present in equation (1). Standard errors were clustered at the hospital-year-month level. For each outcome, an identity link function was used to obtain comparable interpretation with the literature (Bertoli and Grembi, 2017; Cook and Averett, 2020; Jian et al., 2019). The difference-in-differences parameter, β, is the coefficient of interest. It captures the impact of the DIP-based payment adoption on outcomes among insured patients.

To further explore heterogenous effects of the DIP payment reform among patients with different severity, we performed three sets of additional analysis. First, we divided the sample into a group with less severity (CCI equal to 0) and a group with more severity (CCI equal to 1 or above). We then fit our model to the subsamples. Second, we examined variations in the impact of the DIP policy by social insurance types. We replaced *Treatment* in equation (1) with indicators for employee insurance and resident insurance. Finally, we replaced the *Post* dummy in equation (1) with dummies for the first and second year after reform to capture temporal dynamic effects of the reform.

We expected that the signs of the difference-in-differences coefficients could be either positive or negative, depending on the scenarios in question. One major change brought by the new policy is that the hospital-specific fixed rate per admission was replaced by the DIP group-specific fixed rate. On one hand, for mild patients who need less medical investment, the DIP group-specific price was more likely to be lower than the original hospital-specific fixed rate. In such a case, we expected lower medical costs per case and shorter length of stay after the DIP policy adoption. However, if the financial incentives from the early policy has been strong enough, there would be no much room left for providers to further decrease costs. Then, the impact of the policy change among mild patients could be trivial.

On the other hand, for severe patients who need more intensive care, the DIP group-specific price was more likely to be higher than the original hospital-specific fixed rate. With unintended consequences (such as reduced medical services and premature discharge) under the early policy as documented in the literature (Chan and Zeng, 2018), we expected higher medical costs and length of stay after the DIP policy adoption and the impact could be sizable.

Since we expected flipped impacts of the DIP policy adoption among mild and severe patients, the overall direction of the impact on costs and length of stay for the whole sample is unclear, depending on which subpopulation was driving the results. Accordingly, for quality of care measures, we expected quality improvement after the DIP policy adoption among severe patients due to anticipated increases in medical investment; in contrast, the quality impact for mild patients and the whole sample remained uncertain.

2.5. Validity tests and robustness checks

One key assumption of the difference-in-differences approach is that the control and treatment groups would have experienced a common trend in outcomes in the absence of the reform. Considering that the zero-mark-up policy introduced in late 2017 may differentially affect the hospitalization costs of patients with various insurance status, we restricted our estimation sample to discharges after October 2017 when the costs measure was used as the outcome variable. We checked differential trends between the control and treatment groups before the DIP policy reform by replacing post dummies in the interaction term in equation (1) with a series of year-month dummies. Any significant coefficient of the interaction term indicates a violation of the key assumption.

In addition, we also performed three sets of robustness checks. First, to further ensure the comparability between the treatment and control groups, we combined the matching and difference-in-differences methods. Following Blundell and Dias (2009), we matched treated before treatment, nontreated before treatment, and nontreated after treatment to the treated after treatment, separately. We then produced difference-in-differences estimates based on the common support region. Restricted by the computing capacity, the analysis was conducted on a 5% random sample of our main estimation dataset. We performed the "diff" command in Stata. We used kernel propensity matching with a bandwidth of 0.06.

Second, to the extent that the DIP policy also drove the selection of patients for surgical procedures, a potential selection bias would be inherent in the quality outcomes. We examined the robustness of our results by re-defining the operation associated infection and postoperative complication measures among all the patients rather just those received surgeries and reproducing our results. Finally, we reproduced our results for costs and length of stay using generalized linear models with gamma or Poisson distribution and a log link.

All analyses were performed by Stata 16.0 for Windows. We used 5% as the significance level.

2.6. Ethics

This study was approved by the institutional review board of the School of Public Health, Fudan University (IRB#2020-TYSQ-03-20).

2.7. Role of the funding sources

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

3. Results

3.1. Main results

The final sample consisted of 8.16 million discharges from 309 hospitals. Table 2 presents sample statistics before and after the DIP payment reform for insured and uninsured patients, respectively. Inpatient volumes of insured patients were twice those of uninsured patients. We observed an increase in inpatient costs among insured patients and a slight decrease among the uninsured. Other outcome measures decreased after the reform in both groups.

Table 3 presents difference-in-differences estimates of the impact of the DIP payment reform, first for the whole sample (column (2)), then for subgroups by severity levels of the patients (columns (3) and (4)). We also produced unadjusted estimates for the whole sample in column (1)

Table 2
Descriptive statistics before and after the diagnosis-intervention packet (DIP) payment policy among insured and uninsured patients.

	Before policy change (2016–2017)		After the policy change (2018–2019)	
	Insured	Uninsured	Insured	Uninsured
	(1)	(2)	(3)	(4)
Panel A: outcome variables				
Costs per case (RMB)	14005.28	13328.96	15033.04	13235.88
	(17995.25)	(18776.30)	(18587.84)	(18421.51)
Length of stay (days)	9.02	7.81	8.61	6.97
	(11.20)	(11.01)	(11.19)	(10.25)
Operation associated infection (%)	0.10	0.10	0.08	0.08
	(3.15)	(3.21)	(2.91)	(2.79)
Postoperative complication (%)	3.35	6.48	2.74	6.32
	(17.98)	(24.61)	(16.34)	(24.33)
Panel B: patient characteristics				
Age (years)	53.82	40.44	53.90	38.52
	(21.55)	(20.91)	(21.86)	(21.63)
Male (%)	45.78	44.19	46.23	43.62
	(49.82)	(49.66)	(49.86)	(49.59)
Charlson Comorbidity Index (%)				
0	62.11	77.58	58.46	79.25
	(48.51)	(41.71)	(49.28)	(40.55)
1	19.41	10.74	18.95	9.36
	(39.55)	(30.97)	(39.19)	(29.13)
2	11.74	6.83	13.83	6.61
	(32.19)	(25.22)	(34.52)	(24.85)
≥3	6.74	4.85	8.76	4.78
	(25.08)	(21.48)	(28.27)	(21.32)
Sample size	2,621,806	1,221,782	3,138,220	1,177,154

Notes: Means and standard deviations (in parentheses) are reported for insured and uninsured patients both before and after the adoption of the DIP-based payment.

Table 3
Impacts of the diagnosis-intervention packet (DIP) payment adoption: Difference-in-differences estimates.

	Whole sample		Low severity	High severity
	Un-adjusted	Adjusted	Adjusted	Adjusted
	(1)	(2)	(3)	(4)
Panel A: ln (costs per case)				
Coefficient	0.065	0.043	0.031	0.085
Standard error	(0.034)	(0.012)	(0.009)	(0.025)
P-value	0.061	**0.000**	**0.001**	**0.001**
Sample size	4,867,273	4,867,273	3,099,841	1,767,432
Panel B: length of stay				
Coefficient	0.495	0.080	0.088	0.252
Standard error	(0.106)	(0.050)	(0.049)	(0.073)
P-value	**0.000**	0.106	0.076	**0.001**
Sample size	8,158,962	8,158,962	5,301,877	2,857,085
Panel C: operation associated infection				
Coefficient (100* β)	0.009	−0.003	−0.000	−0.014
Standard error	(0.009)	(0.007)	(0.008)	(0.008)
P-value	0.371	0.719	0.975	0.406
Sample size	4,002,641	4,002,641	3,002,014	1,000,627
Panel D: postoperative complication				
Coefficient	−0.008	−0.002	−0.001	−0.001
Standard error	(0.003)	(0.001)	(0.001)	(0.000)
P-value	**0.003**	0.053	0.136	**0.046**
Sample size	4,002,641	4,002,641	3,002,014	1,000,627

Notes: Each panel of each column is a separate regression. The estimated difference-in-difference coefficients, standard errors and *P*-values are reported for the whole sample (columns (1) and (2)), patients with high severity (column (3)), and patients with low severity (column (4)). Column (1) reports the unadjusted results, where no control variables are included except for the post dummy, the treatment indicator, and the interaction term between treatment and post. In columns (2)–(4), all specifications include the full set of control variables (i.e., indicators for age groups, gender, insurance types, principal diagnosis, and CCI, interactions between age group and gender dummies, interactions between gender and CCI dummies, and hospital fixed effects) and hospital location indicators interacted with year-by-month fixed effects. Standard errors are clustered at the hospital-year-month level. When costs per case is used as the outcome variable, the estimation sample is restricted to patients discharged after October 2017. When operation associated infection and postoperative complication are used as outcome variables, the estimation sample is restricted to patients that received surgical procedures.

for reference. We found that the DIP payment reform was associated with a 4.3% ($P = 0.000$) increase in inpatient costs per case among insured patients. The estimated effect on length of stay was small, compared to the sample mean of this measure. Changes in the operation associated infection rate in response to the policy reform were insignificant and negligible in size, consistent with the results in sample statistics. The postoperative complication rate decreased by 0.2 percentage points after the introduction of the DIP-based payment, corresponding to a decrease of 5.9% (0.2/3.35). But the estimate was not statistically significant.

Regarding the subsample results presented in columns (3) and (4) of Table 3, we found that the DIP payment reform was associated with an 8.5% ($P = 0.001$) increase in costs per case among severe patients. The effect was about 2.7 times of that among patients with low severity (3.1%, $P = 0.001$). Changes in length of stay were also more sizable among severe patients. However, both estimates were not economically meaningful. We observed no variations in the impacts of DIP payment reform on the operation associated infection rate by patient severity levels. In contrast, a significant and negative impact on postoperative complication rate was only observed among severe patients (3.6%, $P = 0.046$).

3.2. Additional analysis

Table S1 in supplementary materials presents difference-in-differences estimates of the DIP policy impact among employee insurance participants and resident insurance participants, separately. Except for postoperative complication, the large P-value of the t-tests on the equality of the coefficients suggested no variations in the policy impact by social insurance types. For postoperative complication, we observed a significant and negative impact only among resident insurance participants, however, the size of the estimates was negligible.

Variations in the effects of the DIP payment reform over years are reported in Table 4. The results showed that across all outcome variables and patient severity levels, effects in the second year of the reform were generally stronger than those in the reform year. Previous findings largely remained except for length of stay: we estimated a half-day increase ($P = 0.000$) in length of stay in the second year of the policy reform among severe patients.

3.3. Validity tests and robustness checks

Using an event-study specification, we observed no pre-existing differential change in outcomes between the insured and uninsured before the policy reform. All changes in differences between the two groups were not statistically significant relative to the first month of the study period (Fig. S2 in the supplementary materials), suggesting that our findings were valid. Consistent with the main results for the whole sample, the estimated coefficients remained insignificant and fluctuated

Table 4
Variations in the impacts of the diagnosis-intervention packet (DIP) payment adoption over years.

	Low severity	High severity
	(1)	(2)
Panel A: ln (costs per case)		
DID coefficients for the first year of adoption	0.022	0.046
Standard error	(0.010)	(0.026)
P-value	**0.022**	0.071
DID coefficients for the second year of adoption	0.042	0.131
Standard error	(0.010)	(0.028)
P-value	**0.000**	**0.000**
Sample size	3,099,841	1,767,432
Panel B: length of stay		
DID coefficients for the first year of adoption	0.059	0.008
Standard error	(0.065)	(0.086)
P-value	0.367	0.930
DID coefficients for the second year of adoption	0.121	0.541
Standard error	(0.063)	(0.094)
P-value	0.053	**0.000**
Sample size	5,301,877	2,857,085
Panel C: operation associated infection		
DID coefficients for the first year of adoption (100* β)	−0.012	−0.025
Standard error	(0.009)	(0.020)
P-value	0.219	0.211
DID coefficients for the second year of adoption (100* β)	0.012	−0.003
Standard error	(0.009)	(0.018)
P-value	0.197	0.887
Sample size	3,002,014	1,000,627
Panel D: postoperative complication		
DID coefficients for the first year of adoption	−0.002	−0.001
Standard error	(0.001)	(0.000)
P-value	0.138	0.612
DID coefficients for the second year of adoption	–0.001	−0.001
Standard error	(0.001)	(0.000)
P-value	0.309	**0.009**
Sample size	3,002,014	1,000,627

Notes: Each panel of each column is a separate regression. Difference-in-differences estimates for the first year and second year of adoption, standard errors, and *P*-values are reported. All specifications include the full set of control variables (i.e., indicators for age groups, gender, principal diagnosis, and CCI, interactions between age group and gender dummies, interactions between gender and CCI dummies, and hospital fixed effects) and hospital location indicators interacted with year-by-month fixed effects.

around zero for the length of stay and quality measures. Regarding the costs measure, we noted that estimated effects were insignificant for most months post-treatment due to insufficient statistical power, since only observations in a given month rather all observations in the post period contributed to the identification of the dynamic effects. However, the average of the estimated effects in the post-treatment period was positive with a size consistent with that reported in Table 3 (column (2), panel A). The estimated coefficients also became more sizable in the second year of the policy, which echoed with the findings in Table 4.

Table S2 in supplementary materials presents results from the combined approach of matching and difference-in-differences. Although the estimates were less significant due to smaller sample sizes, the findings were largely unchanged: inpatient costs increased more after the DIP policy adoption among patients with high severity. We also found that our results were not affected by alternative definitions of the quality measures (Table S3 in supplementary materials) or alternative assumptions of the error term distributions of inpatient costs and length of stay (Table S4).

4. Discussion

The study city shifted from the “fixed-rate per admission with a cap on annual compensation” policy to the DIP payment with global budget for all insured patients in 2018. The difference-in-differences estimate showed an average of 8.5% ($p = 0.001$) increase in inpatient costs, trivial changes in length of stay, and a 3.6% ($P = 0.046$) decrease in the postoperative complication rate associated with the reform among patients with high severity. We also found larger changes in the second year of the policy implementation.

The results of a more sizable impact among patient with high severity were consistent with our expectation. Considering that patients with high severity need more intensive care compared to their counterparts, they served as a subpopulation that were disproportionately affected by the unintended consequences of the early policy in the study city. Accordingly, they were expected to be more responsive to the adoption of DIP-based payment. The more apparent increases in inpatient costs per case and length of stay among more severe patients observed in this study reflected more intensive medical investments in sicker patients. These findings were corroborated by interview data. During July and August 2020, we conducted a total of nine individual in-depths interviews with the directors and heads of health insurance office at different types of hospitals in the study city, to understand the experience of the policy implementation and observed behavior changes among providers. Answers from the participants also suggested increased medical activities among sicker patients in response to the DIP adoption:

> “Previously, we sometimes had to transfer patients to other hospitals before their treatment was fully completed. Otherwise, we would exceed the fixed rate (under the early policy) for the admission.”

> “Previously, performing surgeries with level 3 and level 4 complexities was somewhat discouraged as these surgeries were usually associated with higher costs. We now (under the DIP-based payment) have less hesitation in doing that (performing advanced surgeries), since we know that treating patients with higher severity will be reimbursed more.”

In contrast, our results did not show a decrease in inpatient costs among patients with less severity. One possible reason is that the CCI might not be an ideal measure to identify mild patients, since the majority of the patients (58.46%–77.58%) in our sample were recognized as less severe with a CCI level of zero. Another possible reason is that the early direct cost control policy might have set the hospital-specific fixed price at a very low level, leading to increased medical activities for most patients, but in different degrees, under the new and more relaxed policy.

At the market level, based on our calculation, the annual total medical spending among insured patients in the study city increased from 153.03 billion in 2016 to 194.34 billion in 2019. After the payment policy change, the annual growth rate slowed down from 12.61% to 6.52%. The findings suggested that under a global budget, changes in provider behaviors in response to the DIP-based payment were not associated with an increase in total costs market-wide. In this sense, we found no apparent evidence that the DIP payment policy was less effective in containing costs than a direct cost control policy with fixed rates and fee-limits.

The observed lingering effects were partially attributable to the implementation process of the DIP payment policy in the study city. Although the effective dates of the policy started in January of 2018, providers did not know the relative reward intensity of each DIP group until the end of that year. The DIP relative weights were developed and issued in November of 2018. Before then, hospitals only knew that constraints set by the previous policy were lifted and cases with higher severity would be reimbursed more in the new policy. Our findings suggest that the provider side was not very responsive until they received specific and clear information.

Findings from previous studies regarding DRG-based payment are not readily comparable in size and previous payment policies and policy implementation details varied across different institutional backgrounds. However, previous studies generally reported a decline in length of stay and quality of care after the introduction of a DRG-based payment (Epstein et al., 1991; Choi et al., 2019; Kutz et al., 2019). One of the reasons is that the financial incentive to minimize costs contributed to a shift from inpatient to outpatient settings (Farrar et al., 2009). Another widely documented mechanism is the reduction in the use of unnecessary and low-valued procedures due to financial incentives (Tan and Melendez-Torres, 2018). In contrast, our findings indicate that increases in resource utilization were accompanied with quality improvements after the adoption of the DIP-based payment. Although our findings contradict the literature, we note that our results cannot be interpreted as free of similar unintended effects of the DIP policy on provider behaviors. The discrepancy in findings is more related to the harsh constraints exerted by the previous policy in the study city, rather than differences in the nature between payment systems based on DIP and DRG.

Compared to a DRG-based payment system, the clearest distinguishing feature of the DIP payment lies in the patient classification approach using data driven techniques, producing ten times more finely defined patient groups. Nonetheless, as a member of the patient classification-based payment family, the DIP policy may raise similar concerns of unintended consequences, such as reduction in the use of advanced technology and upcoding. However, variations in the classification process may bring about differences in the size of these effects compared to the DRG payment. It has been debated that the DRG payment may discourage the use of advanced and new technology (Babic et al., 2015; Sommersguter-Reichmann, 2000). One concern is that patients receiving these procedures are more likely to incur higher costs than others in the same DRG group. The issue might be more significant in developing countries where hospitals have limited resources to adopt innovations or invest in new infrastructure (Fässler et al., 2015). With the DIP payment system, we expected such unintended impacts to be smaller, since patients receiving different procedures were generally classified into different groups. Moreover, shifting patients to groups with larger increases in price has been extensively reported in previous adoptions of the DRG payment in various settings (Hochuli, 2020; Barros and Braun, 2017; Melberg et al., 2016). We noted that the DIP payment may also provide upcoding incentives to providers and that careful monitoring was needed since the straightforward patient classification rules of DIP might lead to ease of manipulation on patient groups.

Our study contributes to the literature on hospital payment in the following dimensions. First, we introduced the practice of a new patient classification-based payment system, that was developed purely by data driven techniques. Second, we presented the first empirical evidence on the impact of the DIP payment system on provider behaviors. We showed that compared to a direct cost control policy with fixed rates and fee limits, the DIP-based global budget policy helped regulate provider behaviors and achieved better quality when treating high severity patients. Third, we used a large sample, enabling us to gain good statistical power and estimation precision. Fourth, our results yield important implications for middle-and -ow income countries. We presented evidence that the DIP payment with the global budget supported easy implementation, and avoided insufficient medical investments while achieved global cost control, and thus would be valuable for resource-limited areas who shared common challenges in payment implementation.

Our study has several caveats. First, the ideal condition for applying a difference-in-differences approach requires that the changes in outcomes of the treatment group should be the same as that of the control group in the absence of the policy. Although all our model specifications passed the parallel tests, other health insurance policies or changes of hospital management policies that occurred during our study window may have exerted differential impacts on the insured and uninsured population. We were not able to further disentangle the effects of these concurrent policies. Second, we ignored possible patient selection when splitting the sample by CCI levels. Third, upcoding is a general issue facing patient classification-based payment systems. Although we observed changes in provider behaviors, we were not able to tell whether such changes were appropriate or partially due to upcoding behaviors. Fourth, responses to the DIP policies may vary across different types of hospitals and diseases, which could be the focus of future research. Fourth, we examined an immediate impact of the DIP payment policy. Since hospitals may fundamentally change their way of providing care and develop their advantaged domains in terms of efficiency, the intermediate and long-run impacts of the policy are still unknown.

Declaration of competing interest

All authors report no competing interests.

Acknowledgements

We are grateful to Shenglan Tang from Duke University for his valuable comments. We thank Mengyun Sui, Hongzheng Li, and Long Xue from the School of Public Health, Fudan University for assistance in data cleaning. We are also thankful to Jiaqi Yan and Xue Tang for participating in group discussions.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.socscimed.2021.114415.

Credit author statements

Contributors: X.Y. and S.X. conceptualized the study design, coordinated resources and acquired funding to support the study. M.Q. and X.Z analyzed data and interpreted results. M.Q., X.Z., and Y.C. wrote the manuscript. M.Q. and X.Y. reviewed and edited the manuscript. M.Q. and X.Z. has verified the underlying data. MQ and XY contributed equally to the work.

Funding

X.Y. acknowledges financial support from the National Natural Science Foundation of China (No. 72074501) and National Healthcare Security Research Center (Capital Medical University) (No. YB2020B01).

References

Babic, U., Soldatovic, I., Vukovic, D., Santric-Milicevic, M., Vukotic, V., 2015. Comparative analysis of current payment system for hospital services in Serbia and projected payments under diagnostic related groups system in urology. Vojnosanit. Pregl. 72 (3), 251–257.

Barros, P., Braun, G., 2017. Upcoding in a national health service: the evidence from Portugal. Health Econ. 26 (5), 600–618.

Bertoli, P., Grembi, V., 2017. The political economy of diagnosis-related groups. Soc. Sci. Med. 190, 38–47.

Blundell, R., Dias, M.C., 2009. Alternative approaches to evaluation in empirical microeconomics. J. Hum. Resour. 44 (3), 565–640.

Busse, R., Cots, F., Kobel, C., O'Reilly, J., Tan, S.S., 2013. Diagnosis related groups in Europe: moving towards transparency, efficiency, and quality in hospitals? Br. Med. J. 346, f3197.

Busse, R., Schreygg, J., Smith, P.C., 2006. Hospital case payment system in Europe. Health Care Manag. Sci. 9 (3), 211–213.

Chan, M.K., Zeng, G., 2018. Unintended consequences of supply-side cost control? Evidence from China's new cooperative medical scheme. J. Health Econ. 61, 27–46.

Choi, J.W., Kim, S.J., Park, H.K., Jang, S.I., Park, E.C., 2019. Effects of a mandatory DRG payment system in South Korea: analysis of multi-year nationwide hospital claims data. BMC Health Serv. Res. 19 (1), 776.

Cook, A., Averett, S., 2020. Do hospitals respond to changing incentive structures? Evidence from Medicare's 2007 DRG restructuring. J. Health Econ. 73, 102319.

Epstein, A.M., Bogen, J., Dreyer, P., Thorpe, K.E., 1991. Trends in length of stay and rates of readmission in Massachusetts: implications for monitoring quality of care. Inquiry (28), 19–28.

Farrar, S., Yi, D., Sutton, M., Chalkley, M., Sussex, J., Scott, A., 2009. Has payment by results affected the way that English hospitals provide care? Difference-in-differences analysis. Br. Med. J. 339, b3047.

Fässler, M., Wild, V., Clarinval, C., Tschopp, A., Faehnrich, J.A., Biller-Andorno, N., 2015. Impact of the DRG-based reimbursement system on patient care and professional practice: perspectives of Swiss hospital physicians. Swiss Med. Wkly. 145, w14080.

Fetter, J., Shin, Y., Freeman, J.L., Averill, R.F., Thompson, J.D., 1980. Case mix definition by diagnosis-related groups. Med. Care 18 (Suppl. 2), 1–53.

Hochuli, P., 2020. Losing body weight for money: how provider-side financial incentives cause weight loss in Swiss low-birth-weight newborns. Health Econ. 29 (4), 406–418.

Jian, W., Lu, M., Chan, K.Y., Poon, A.N., Han, W., Hu, M., Yip, W., 2015. Payment reform pilot in Beijing hospitals reduced expenditures and out-of-pocket payments per admission. Health Aff. 34 (10), 1745–1752.

Jian, W., Lu, M., Liu, G., Chan, K.T., Poon, A.N., 2019. Beijing's diagnosis-related group payment reform pilot: impact on quality of acute myocardial infarction care. Soc. Sci. Med. 243, 112590.

Kutz, A., Gut, L., Ebrahimi, F., Wagner, U., Schuetz, P., Mueller, B., 2019. Association of the Swiss diagnosis-related group reimbursement system with length of stay, mortality, and readmission rates in hospitalized adult patients. JAMA Network Open 2 (2), e188332.

Liu, J., Liu, J., Wu, L., 2021. Effect evaluation on cost control of operative appendicitis patients under DRG-based comprehensive management. China. Health Econ. 40 (4), 82–84.

Magnussen, J., Hagen, T.P., Kaarbe, O.M., 2006. Centralized or decentralized? A case study of Norwegian hospital reform. Soc. Sci. Med. 64 (10), 2129–2137.

Mateus, C., 2011. Portugal: results of 25 years of experience with DRGs. In: Busse, R., Geissler, A., Quentin, W., Wiley, M. (Eds.), Diagnosis Related Groups in Europe: Moving towards Transparency, Efficiency and Quality in Hospitals. Open University Press, pp. 381–400.

Mathanuer, I., Wittenbecher, F., 2013. Hospital payment systems based on diagnosis-related groups: experiences in low- and middle-income countries. Bull. World Health Organ. 91 (10), 746-56A.

Melberg, H.O., Olsen, C.B., Pederson, K., 2016. Did hospitals respond to changes in weights of Diagnosis Related Groups in Norway between 2006 and 2013? Health Pol. 120 (9), 992–1000.

Mihailovic, N., Kocic, S., Jakovljevic, M., 2016. Review of diagnosis-related group-based financing of hospital care. Health Service. Res. Manag. Epidemiol. 3, 1–8.

Mikkola, H., Keskimaki, I., Hakkinen, U., 2002. DRG-related prices applied in a public health care system–can Finland learn from Norway and Sweden? Health Pol. 59 (1), 37–51.

Sommersguter-Reichmann, M., 2000. The impact of the Austrian hospital financing reform on hospital productivity: empirical evidence on efficiency and technology changes using a non-parametric input-based Malmquist approach. Health Care Manag. Sci. (3), 309–321.

Tan, S.Y., Melendez-Torres, G.J., 2018. Do prospective payment systems (PPSs) lead to desirable providers' incentives and patients' outcomes? a systematic review of evidence from developing countries. Health Pol. Plann. 33 (1), 137–153.

Xu, S., Xie, H., Cui, X., et al., 2020. Principle and method of the big data diagnosis-intervention packet (DIP) payment system. China Health Insurance (9), 23–28 (in Chinese).

Ying, Y., 2021. DIP & DRG: similarities and differences. China Health Insurance (1), 39–42 (*in Chinese*).

Yuan, S., Liu, W., Wei, F., et al., 2019. Impacts of hospital payment based on diagnosis related groups (DRGs) with global budget on resource use and quality of care: a case study in China. Iran J. Pub. Health 48 (2), 238–246.
Zhang, Y., Le, X., Zeng, Q., 2020. Implementation path and effect of the big data diagnosis-intervention packet (DIP) payment in guangzhou. China Health Insurance (9), 47–51 (in Chinese).
Zhang, Y., He, G., Liu, J., 2016. Comparison of medical costs and care of appendectomy patients between fee-for-service and set fee for diagnosis-related group systems in 20 Chinese hospitals. Southeast Asian J. Trop. Med. Publ. Health 47 (5), 1055–1061.