

Diagnostic related groups-based reimbursement in Kazakhstan: Have we learned enough?

Journal of Public Health Research
2025, Vol. 14(4), 1–13
© The Author(s) 2025
DOI: 10.1177/22799036251365567
journals.sagepub.com/home/phj



Daliya Kaskirbayeva^{1,2} , Silviya Nikolova³, George Ellison⁴ and Tim Ensor¹ 

Abstract

Background: Kazakhstan implemented a diagnosis-related group (DRG)-based payment system in 2012, transitioning from a fee-for-service remuneration model for public hospitals. The reform aimed to enhance hospital efficiency and cost control. A national rollout was preceded by a pilot phase involving 21 hospitals in 2011. While DRG systems are increasingly adopted and adapted globally, evidence on their effects remains limited in low- and middle-income countries.

Objective: To assess whether the introduction of the DRG-based system is associated with changes in hospital activity and the quality of care.

Design and methods: A controlled interrupted time series design was employed, using matched control hospitals to estimate the causal impact of the payment reform. The analysis was based on weekly hospital discharge data from the 2011/2012 to 2012/2013 fiscal years, obtained from the Ministry of Health.

Result: Findings suggest that the impact of the reform on hospital activity varied by hospital levels, with pilot hospitals showing a quicker response to nationwide implementation. No significant effects were observed on quality of care, as measured by standardised in-hospital mortality rates.

Conclusions: The introduction of the DRG-based payment system was associated with heterogenous effects on hospital activity but did not yield measurable short-term improvements in care quality. These findings underscore the importance of implementation context and highlight the need for further research to assess long-term and system-wide effects.

JEL classification: H51, I1, P36

Keywords

healthcare financing, provider payment methods, hospital reimbursement, health sector reform

Received: 13 March 2025; accepted: 17 July 2025

Introduction

Healthcare provider payment reforms are essential for enhancing the efficiency of the health sector, optimising resource allocation strategies, and achieving broader health system goals. A growing body of evidence synthesising global experiences with hospital payment reforms reports that many countries have implemented new payment systems and are actively seeking further innovative approaches to hospital reimbursement.^{1,2} For instance, systematic literature reviews report shifts between different payment systems, with one of the most prominent reforms over the past several decades being the transition from fee-for-service (FFS) models to case-based payment systems, particularly the adoption of diagnostic-related groups (DRG).^{2–5} Recent studies synthesising evidence

from different contexts – Western Pacific, Southeast Asia, Eastern Mediterranean, and Central and Eastern Europe – highlight the prevalence of technical challenges faced by many countries during the DRG implementation.^{3,6}

¹University of Leeds, Leeds, UK

²Maqsut Narikbayev University, Astana, Kazakhstan

³EMEA Real World Methods & Evidence Generation, IQVIA Ltd

⁴University of Central Lancashire, Preston, England, UK

Current affiliations:

Silviya Nikolova is now affiliated with Health Economics and Outcomes Research Ltd., Cardiff, UK

Corresponding author:

Daliya Kaskirbayeva, University of Leeds, Maqsut Narikbayev University.

Emails: d.kaskirbayeva@gmail.com; d_kaskirbayeva@kazguu.kz



These challenges include unreliable medical records needed for the case-mix system, a shortage of well-trained personnel, limited financial literacy among healthcare providers, and the absence of robust evaluation mechanisms and performance measurement indicators. The most recent studies provide overviews of ongoing reforms in European countries such as Austria, Denmark, France, Germany, Bulgaria, Croatia, Czechia, Estonia, Latvia, Lithuania, Hungary, Poland, and Romania, as well as in Australia, in Canada (Ontario), and the United States.^{1,2} These studies report that many of these countries are moving away from relying predominantly on DRG-based payment systems in favour of more diversified financing methods aimed at adapting existing models to the local context.^{1,2} Some countries that aimed to enhance hospital efficiency without compromising quality of care through DRG-based payment systems experienced unintended side effects, such as insufficient cost control and limited integration with other health services.² The evidence regarding the impact of DRG implementation on quality of care remains inconclusive. Existing studies suggest that DRG implementation has had only a modest effect on health system efficiency, particularly in terms of cost containment, and may be associated with a potential deterioration of quality of care.⁷⁻⁹ A systematic literature review focussing on the Chinese context reveals mixed effects of DRG reforms on quality of care, cost containment, and efficiency.⁵ Therefore, some countries have adjusted their existing payment models to better address system needs and mitigate potential negative consequences. These adjustments typically involve modifications to tariffs and the integration of multiple payment methods.^{1,2}

Despite the broad evidence synthesised on new provider payment reforms across diverse geographical regions – including North America, Europe, the Western Pacific, and Southeast Asia – a paucity of evidence remains regarding their implementation and impact in other regions, such as South America, Africa, the Middle East and Central Asia, particularly in low-income and developing settings. Therefore, generalisations on existing studies may lack robustness.

A significant limitation in the current body of evidence assessing the impact of provider payment reforms on system-level outcomes is the generally low methodological quality of many studies. For example, a systematic literature review focussing on low- and middle-income countries found that much of the existing research is constrained by weak study designs.⁴ One key factor contributing to these limitations is the lack of piloting phases during the reform rollouts – an element identified as a critical facilitator of successful implementation.⁶ Moreover, well-structured pilot studies enable the application of quasi-experimental designs, which are widely regarded as the most rigorous methodological alternative to randomised control trials for establishing causal inference in complex health policy settings.

Despite the growing body of synthesised evidence regarding the impacts of provider payment reforms, there remains a significant gap in understanding how quickly hospitals adapt to newly introduced payment systems, and whether such reforms serve as a panacea for improving system-level outcomes. This gap limits insights into the broader and long-term effects of the reforms, particularly regarding their systemic implications and their influence on healthcare efficiency and quality of care over time. Consequently, there is an increasing demand for more robust, context-specific evidence to inform policy decisions and strengthening the resilience of healthcare systems, particularly in low- and middle-income countries.

To address a literature gap, this study focuses on examining the impact of the transition to DRG on hospital activity in Central Asia, specifically in Kazakhstan. The study employs a holistic methodology by utilising hospital-level data across short time intervals, leveraging weekly data that is particularly sensitive to policy changes especially when reforms are implemented unevenly throughout the country. A controlled interrupted time series design is employed to account for secular trends (i.e. changes observed only in the control group) by assessing shifts in either the intercept, the slope, or both. In the context of multiple concurrent hospital payment reforms, including the introduction of a mandatory social health insurance scheme following the adoption of DRG, the primary objective of this study is to evaluate the immediate response of hospitals to the implementation of DRG. The findings aim to provide valuable insights into the responsiveness of hospitals in upper-middle and low- and middle-income countries to newly introduced provider payment systems.

Implementation of the DRG payment system and hospital system

Since gaining independence after the collapse of the Soviet Union in the 1990s, the Republic of Kazakhstan has made several attempts to reform the payment system used to compensate public hospitals. During this transitional period, health facilities were reimbursed through line-item budgets, which covered specific categories primarily related to staff and bed costs. The first health financing reform was introduced in 1996, implementing a new payment system intended to reimburse hospitals based on case-based payments. However, this payment mechanism did not enhance hospital performance, mainly because state budget planning remained highly centralised. Consequently, healthcare providers were not granted the autonomy essential for managing and reallocating resources more effectively, and local healthcare authorities were unable to transfer resources between providers based on their performance. In 2007, the FFS model replaced case-based payments. Hospitals were reimbursed for each service based on actual costs, which included medical

expenses like medicines and supplies, alongside non-medical costs such as utilities and staff training. This incentivised hospitals to increase service volume, as they were paid for each service delivered. In 2012, the payment system shifted back to a case-based model to control hospital costs and enhance productivity accountability. Before nationwide implementation, the system was piloted in 21 hospitals over 17 weeks, starting in September 2011.¹⁰ The pilot aimed to adapt the DRG system and calculate costs, initially establishing 180 groups based on clinical and cost similarities. The clinical algorithm used the International Statistical Classification of Diseases, while the costing relied on historical prices (FFS tariffs) and volume, due to the absence of a unified accounting system. After the pilot, clinical and cost homogeneity was reassessed, increasing the number of groups from 180 to over 400.¹¹ In 2013, global budgets (GB) were introduced for district hospitals to contain unnecessary hospital services.

The DRG tariffs are set nationally, but certain hospitals receive adjustments through scale-up factors to help maintain financial stability. These adjustments support hospitals in specific regions, including ecological disaster zones, areas with extreme summer and winter temperatures (a hardship factor resulting in high energy costs), and economically disadvantaged rural areas. Hospitals operating autonomously are allowed to retain 100% of any surplus generated. This concept is designed to incentivise hospitals to reduce the length of stays where feasible by admitting day-case patients and implementing innovative, cost-effective surgical treatments. These incentives aim to enhance the efficiency of resource utilisation in hospitals.

The DRG system is revised annually to adjust diagnostic categories and recalibrate tariffs based on complications and resources under new medical protocols. Tariffs for day cases within the same group are set at one-quarter of the corresponding inpatient tariffs, while surgical case tariffs are generally higher than those for non-surgical cases. By setting higher tariffs for surgical procedures, the DRG system incentivises the admission of patients requiring more complex treatments and supports more efficient management of hospital admissions compared to FFS schemes. Thus, it encourages increased surgical activity compared to the FFS model.

One key component of the DRG reform was the healthcare technology programme,¹² which enables providers to adopt the best medical practices and technologies from abroad once approved by the health technology assessment committee. The list of these technologies and their associated tariffs is updated annually. Hospital cases that utilise these technologies receive an additional scale-up factor on top of the standard DRG tariff. As a result, hospitals may be motivated to increase admissions for surgical care or for more expensive and profitable diagnostic groups.

The Ministry of Health of the Republic of Kazakhstan (MoH) and public hospitals agree on the anticipated volume of care for the forthcoming year based on historical

Table 1. Summary of expected results of a shift from FFS to DRG.

| Outcome variable | Expectation from the theory | Policymakers expectations |
|--------------------|-----------------------------|---------------------------|
| Overall admissions | Ambiguous | A decrease |
| Inpatient cases | Not recorded | A decrease |
| Day cases | Not recorded | An increase |
| Surgery cases | Not recorded | An increase |
| ALOS | A decrease | A decrease |
| SMR | An increase | Not recorded |

data and adjusted for local socio-demographic factors. Between June and July, during mid-term performance evaluations, hospitals and the MoH review and revise contractual terms to address any discrepancies in care provision. Following these discussions, the MoH may adjust funding levels accordingly.

Under Kazakhstan's FFS system, providers were compensated per procedure, which encouraged longer patient retention. Conversely, the DRG system incentivises hospitals to reduce the average length of stay (ALOS). While policymakers anticipated an increase in day cases and shorter ALOS, the DRG model may lead to premature discharges or inadequate care. This may elevate hospital-related morbidity or mortality if the quality of care per patient diminishes due to the shorter stays. A variety of quality indicators related to care, such as readmission rates, should ideally be employed. However, this study lacked sufficient data on those indicators; thus, the assessment of care quality was based solely on mortality rates.

In summary, the measures of healthcare outputs utilised in this study are: (i) the number of inpatient cases; (ii) the number of day cases; (iii) the number of surgical cases, categorised by two age groups; (iv) the ALOS; and (v) standardised in-hospital mortality rates (SMR). Table 1 outlines the expected outcomes of DRG implementation based on theoretical assumptions and policymaker expectations.

Kazakhstan's hospital system is hierarchical, with Republican Centres and multi-profile hospitals at the regional, city, and district levels, alongside mono-profile hospitals. Patients are referred from lower to higher levels of care. This study adopts the hospital classification proposed by Katsaga et al.¹³ (2012) and the OECD,¹⁴ which applies to public hospitals managed by the MoH in 2011 and 2012. A summary of this classification is provided in Table S2 of the Supplemental File.

Methods

Data

The following analyses are based on administrative hospital data collected for the 2011/12 and 2012/13 periods, as

provided by the MoH. Participation in the pilot phase of the policy implementation was determined according to the relevant legislation.¹⁰ Data on aggregated primary healthcare visits were extracted from the National Statistics Yearbook.

The original dataset consisted of an unbalanced panel of 361 hospitals for the 2011 and 2012 fiscal years. Hospital episodes with an average length of stay exceeding 90 days were excluded, as they likely corresponded to long-term care,¹⁵ representing 0.01% of the data. Additionally, low-cost episodes, linked to preliminary diagnostics rather than full inpatient stays, were also excluded. These episodes were identified using a threshold based on the country's monthly calculation index for budget planning, set at 1618 national currencies (approximately 10.5 USD), also accounting for 0.01% of all episodes. Furthermore, episodes recorded in December, which represent 5% of all episodes, were omitted. December shows a notable decline in admissions (Supplemental File, Figure S1), which could distort the model estimates by capturing changes in admission levels immediately following the introduction of the DRG in January. This December decline is likely due to the end of the purchaser-provider annual contracting period and the occurrence of four consecutive public holidays.

After balancing the dataset, the final (unmatched) dataset consisted of 1,948,956 and 2,095,175 admitted patients from 331 hospitals in 2011 and 2012, respectively. Hospitals that participated in the pilot phase of DRG implementation between September 2011 and January 2012 were classified as piloted hospitals ($n=19$), which included eight regional hospitals, four city hospitals, and seven district hospitals. Hospitals that did not participate in the pilot phase were categorised as non-piloted hospitals ($n=312$). Republican-level hospitals were excluded from the study, as they did not participate in the piloting phase. All hospitals in this study were public, as the DRG system was implemented exclusively in public hospitals. The time unit for the study was 1 week, with the observation period spanning 48 weeks in both 2011 and 2012.

Controlled interrupted time series with matching technique

Interrupted time series (ITS) has become an increasingly popular design for assessing healthcare interventions.¹⁶⁻¹⁹ Data are routinely collected at equally spaced time intervals before and after the introduction of an intervention at a clearly defined point in time. Differences in levels and slopes before and after the intervention are attributed to the policy effect.²⁰ ITS is particularly useful because it allows for the investigation of and adjustment for potential bias arising from secular trends, seasonal effects, and autocorrelation,²¹ as well as parallel trends. However, it does not facilitate the control of time-varying confounders.¹⁸ One

method to address concurrent confounding events is to incorporate a control group or control series. This design is termed controlled interrupted time series (CITS) and is regarded as a more powerful technique than Difference-in-Differences (DID),^{22,23} which is currently the most commonly used method in public health evaluations.²⁴ While CITS enables the control of other time-varying factors by comparing the outcome of interest between intervention and control groups simultaneously, selection bias may persist between the two groups. To mitigate selection bias, a matching approach can be employed to account for pre-treatment characteristics that differ between treatment and control groups.¹⁸

Hospitals in Kazakhstan exhibit considerable heterogeneity, particularly in terms of size and the range of services provided. To address this variability, we identified 'matched' pairs of treatment and control hospitals with balanced baseline characteristics using Mahalanobis distance matching (MDM). Hospitals involved in the pilot were matched with those not-participating hospitals based on proxies for hospital size and physician workload. In the absence of actual data, we used the average number of beds required for inpatient cases as a proxy for size and the ratio of full-time equivalent physicians to beds as a proxy for workload. According to the OECD¹⁴ review, there has been a trend since 2006 towards a reduction in the number of health facilities and beds in the country, suggesting that the number of beds was not affected by the reform in 2012. Similarly, changes in the healthcare workforce were likely associated with broader structural reforms, such as hospital closures, rather than the DRG intervention itself, supporting the strong ignorability assumption. We conducted a multiple-group CITS analysis to evaluate the impact of the pilot phase on outcome levels and trends, relative to the pre-pilot period, including the effects of nationwide implementation. The analysis examined the immediate and longer-term impacts of each intervention phase to determine how quickly hospitals adapted to changes in the payment system. The CITS model is defined as:

$$y_{it} = \beta_0 + \beta_1 T_t + \beta_2 Z_i + \beta_3 Z_{it} T_t + \beta_4 X_t + \beta_5 X_{it} T_t + \beta_6 Z_{it} X_t + \beta_7 Z_{it} X_{it} T_t + \beta_8 W_t + \beta_9 W_{it} T_t + \beta_{10} Z_{it} W_t + \beta_{11} Z_{it} W_{it} T_t + \beta_{12} H_t + \beta_{13} \text{season} + \beta_{14} \text{holiday} + \varepsilon_{it} \quad (1)$$

where y_{it} is the aggregated outcome variable measured at time t (week); T_t is the time since the start of the study (January 2011 by weeks $t \in [0;96]$); a dummy variable Z_{it} takes a value of 1 if hospital i participated in the piloting phase, and equals 0 otherwise; a dummy variable X_{it} indicates the piloting phase and takes a value of 1 for week $t \in [36;48]$ and 0 otherwise; a dummy variable W_t represents the nationwide implementation of DRG and takes a value of 1 for week $t \in [49;96]$, and 0 otherwise; H_t is the

vector of hospital-specific confounders, including covariates such as the regional location (regions in Kazakhstan vary substantially in terms of demographic structure, epidemiological profiles, geographical features, and weather conditions) and speciality. To account for differences in hospital functions, the national classification of hospital specialities was adopted. This classification includes general hospitals, teaching hospitals, children's hospitals, childbirth and neonatal centres, research centres, and cardio surgery centres. *season* is a categorical variable representing seasonality; *holiday* is a dummy variable for holidays. In Kazakhstan, when a public holiday falls on a weekend, the following Monday is officially designated as a public holiday. It was assumed that public holiday may lead to a decline in hospital admissions – excluding emergency cases – due to the temporary closure of businesses and institutions. Additionally, hospital personnel may adjust their workload schedules around public holidays, potentially influencing patterns of service delivery. ε_{it} captures random variation.

The coefficients of interest, β_6 and β_{10} , indicate the difference between piloted and non-piloted hospitals in the level of the outcome variable immediately following the start of the piloting and nationwide phases, respectively. Similarly, β_7 and β_{11} indicate the differences in post-intervention trends between two groups following the piloting and nationwide phases, relative to the corresponding pre-intervention periods.

Using time series panel data requires an examination of serial correlation, heteroscedasticity, stationarity, and seasonality.^{17,25} By examining both the autocorrelation and partial autocorrelation functions for each outcome variable, orders of autocorrelation were obtained. The augmented Dickey-Fuller test suggests that the time series were stationary. Adjustment for seasonality was also required. Segmented regression for the CITS analysis was performed using the GLS function from the **nlme** package in R,²⁶ that performs a generalised least squares model with first-order autoregressive (AR(1)) errors, allowing the residuals to be autocorrelated and defined as: $\varepsilon_t = \rho\varepsilon_{t-1} + u_t$, where the autocorrelation parameter is such that $|\rho| < 1$, and the disturbances u_t are independently and identically distributed as $N(0, \sigma^2)$.

The balance of the two groups (piloted and non-piloted hospitals) on the level and trajectory of outcome variables in the pre-piloting phase is verified by both visual inspection and checking p -values on β_2 and β_3 , as suggested by Linden.²⁷

Considering that the model accounts for the time required for the policy's effects to develop, the outcome five months post-policy introduction was compared against an estimated counterfactual of no policy implementation. All sectors of the economy undergo mid-term performance evaluations annually from June to July, during which all

ministries report on the mid-term national budget execution, necessitating financial reporting by health facilities. The policy impact needed to be measured prior this reporting period, when the volume of care provided within the purchaser-provider contract is reconsidered. The model estimates were used to compute both absolute and relative changes in outcomes five months after DRG implementation, just before the mid-term performance evaluation. Therefore, the expected value of the outcome variable five months post-intervention is denoted as \hat{Y}_5^{after} . The counterfactual outcome in the absence of the policy is represented as $\hat{Y}_5^{no\ policy}$, which reflects the projected level of the outcome had the policy not been introduced. Thus, the absolute change is estimated as:

$$Absolute\ change = \hat{Y}_5^{after} - \hat{Y}_5^{no\ policy} = \hat{\beta}_8 + \hat{\beta}_9 * 22,$$

The relative change is estimated as:

$$Relative\ change = \frac{\hat{Y}_5^{after} - \hat{Y}_5^{no\ policy}}{\hat{Y}_5^{no\ policy}} = \frac{\hat{\beta}_8 + \hat{\beta}_9 * 22}{\hat{\beta}_0 + \hat{\beta}_1 * (22 + 48)},$$

where 22 weeks corresponds to a five-month post-intervention period, and 48 is the number of baseline weeks. The 95% confidence intervals around the estimated absolute and relative changes are calculated using the delta method, as proposed by Zhang et al.,¹⁶ to account for the uncertainty in the model's parameter estimates.

Results

Descriptive statistics

Table 2 presents the descriptive characteristics of both piloted and non-piloted hospitals for the years 2011 and 2012. In 2012, the patient case mix increased in regional hospitals, whereas ALOS slightly decreased in all hospitals.

Matching results

The assessment of covariate balance (Supplemental file, Table S1) indicates that MDM achieved a good overall balance, significantly outperforming propensity score matching (PSM) and supporting the critiques by King et al.²⁸ PSM led to a reduction in the number of observations due to the exclusion of unmatched treated hospitals, whereas MDM preserved sample size and produced comparable covariate distributions between the treated and control groups (as shown in the density plots in Supplemental File, Figures S2a, S2b, and S2c). Consequently, the MDM-matched sample of hospitals was employed in the CITS models described in Section 3.2.

Table 2. Descriptive statistics for the balanced and unmatched panel for years 2011 and 2012.

| Hospital level | Indicator | 2011 | | | | 2012 | | | | |
|----------------------------------|---------------------------|-------------------|---------|---------|----------|----------|---------|---------|----------|----------|
| | | Mean | S.D. | Min | Max | Mean | S.D. | Min | Max | |
| (a) Piloted hospitals | | | | | | | | | | |
| Regional hospitals (N=8) | Inpatient cases | 9364.25 | 4136.83 | 4484.00 | 17061.00 | 9391.25 | 3861.84 | 4723.00 | 16621.00 | |
| | Day cases | 640.88 | 542.86 | 192.00 | 1812.00 | 892.63 | 639.01 | 159.00 | 1992.00 | |
| | ALOS | 10.01 | 2.27 | 6.18 | 14.01 | 9.43 | 2.02 | 5.64 | 12.01 | |
| | Occupied bed days | 289.47 | 25.42 | 241.06 | 319.16 | 284.64 | 28.35 | 243.46 | 321.58 | |
| | Patient throughput | 31.30 | 10.58 | 21.25 | 54.61 | 32.23 | 9.45 | 23.75 | 51.56 | |
| | Beds | 320.50 | 166.29 | 166.00 | 568.00 | 311.38 | 157.35 | 160.00 | 570.00 | |
| | Day beds | 17.00 | 11.60 | 2.00 | 30.00 | 22.00 | 19.32 | 2.00 | 63.00 | |
| City hospitals (N=4) | Case mix index | 1.73 | 0.92 | 0.91 | 3.86 | 1.88 | 1.77 | 0.79 | 6.10 | |
| | Inpatient cases | 6950.80 | 1454.87 | 4459.00 | 8134.00 | 7123.60 | 1094.88 | 5186.00 | 7881.00 | |
| | Day cases | 700.80 | 378.97 | 232.00 | 1158.00 | 839.60 | 467.35 | 227.00 | 1367.00 | |
| | ALOS | 11.66 | 2.07 | 8.93 | 14.69 | 11.30 | 2.23 | 8.37 | 14.58 | |
| | Occupied bed days | 287.19 | 32.81 | 238.61 | 325.39 | 286.77 | 46.17 | 228.78 | 355.33 | |
| | Patient throughput | 25.70 | 6.62 | 21.00 | 37.16 | 27.00 | 9.51 | 20.30 | 43.22 | |
| | Beds | 291.40 | 102.88 | 120.00 | 380.00 | 293.40 | 105.87 | 120.00 | 382.00 | |
| District hospitals (N=7) | Day beds | 17.80 | 13.79 | 5.00 | 38.00 | 17.80 | 13.79 | 5.00 | 38.00 | |
| | Case mix index | 1.37 | 0.16 | 1.19 | 1.61 | 0.91 | 0.15 | 0.78 | 1.13 | |
| | Inpatient cases | 2623.86 | 924.51 | 1688.00 | 4425.00 | 2611.29 | 830.06 | 1841.00 | 4161.00 | |
| | Day cases | 1687.43 | 895.10 | 327.00 | 2843.00 | 1741.29 | 928.68 | 408.00 | 3210.00 | |
| | ALOS | 9.09 | 0.55 | 8.23 | 9.95 | 8.45 | 0.62 | 7.52 | 9.46 | |
| | Occupied bed days | 215.25 | 18.58 | 189.70 | 248.20 | 223.25 | 32.37 | 192.71 | 287.60 | |
| | Patient throughput | 24.02 | 2.65 | 20.80 | 28.59 | 26.75 | 2.70 | 22.49 | 30.68 | |
| (b) Non-piloted hospitals | Beds | 109.57 | 38.79 | 70.00 | 190.00 | 100.09 | 40.97 | 60.00 | 185.00 | |
| | Day beds | 52.00 | 27.26 | 17.00 | 91.00 | 52.43 | 26.56 | 17.00 | 91.00 | |
| | Case mix index | 0.76 | 0.09 | 0.65 | 0.90 | 0.70 | 0.09 | 0.54 | 0.79 | |
| | Regional hospitals (N=55) | Inpatient cases | 7548.00 | 5911.41 | 581.00 | 24334.00 | 7860.15 | 6081.34 | 679.00 | 25297.00 |
| | | Day cases | 480.18 | 598.41 | 0.00 | 3101.00 | 582.67 | 692.59 | 0.00 | 3647.00 |
| | | ALOS | 11.19 | 4.16 | 4.95 | 21.46 | 10.56 | 3.63 | 4.43 | 18.73 |
| | | Occupied bed days | 296.18 | 42.54 | 182.70 | 460.07 | 303.05 | 39.40 | 222.64 | 446.35 |
| Patient throughput | | 30.77 | 12.77 | 13.03 | 65.01 | 32.72 | 12.79 | 12.73 | 72.78 | |
| Beds | | 233.35 | 173.16 | 40.00 | 730.00 | 231.96 | 173.26 | 40.00 | 705.00 | |
| Day beds | | 12.29 | 13.62 | 0.00 | 60.00 | 16.22 | 17.82 | 0.00 | 77.00 | |
| City hospitals (N=83) | Case mix index | 1.21 | 0.62 | 0.50 | 4.00 | 1.09 | 0.89 | 0.43 | 6.12 | |
| | Inpatient cases | 6263.00 | 4770.48 | 376.00 | 21172.00 | 6839.05 | 5061.63 | 408.00 | 20821.00 | |
| | Day cases | 669.11 | 603.27 | 0.00 | 2469.00 | 779.40 | 674.77 | 0.00 | 2860.00 | |
| | ALOS | 9.45 | 2.59 | 4.61 | 18.82 | 8.96 | 2.36 | 4.33 | 17.03 | |
| | Occupied bed days | 273.42 | 63.15 | 140.44 | 475.07 | 286.06 | 59.55 | 153.75 | 461.26 | |
| | Patient throughput | 31.99 | 13.05 | 10.35 | 71.64 | 34.70 | 12.60 | 16.80 | 74.73 | |
| | Beds | 199.30 | 148.59 | 20.00 | 670.00 | 198.38 | 147.81 | 20.00 | 670.00 | |
| District hospitals (N=174) | Day beds | 16.09 | 14.84 | 0.00 | 65.00 | 17.51 | 15.17 | 0.00 | 75.00 | |
| | Case mix index | 1.03 | 0.40 | 0.42 | 3.80 | 0.82 | 0.45 | 0.30 | 3.39 | |
| | Inpatient cases | 3111.02 | 2577.83 | 318.00 | 15174.00 | 3450.89 | 2926.33 | 309.00 | 17443.00 | |
| | Day cases | 1432.06 | 1130.48 | 0.00 | 7447.00 | 1454.59 | 1171.15 | 0.00 | 9644.00 | |
| | ALOS | 8.70 | 0.89 | 6.27 | 11.39 | 8.46 | 0.80 | 6.05 | 11.82 | |
| | Occupied bed days | 226.61 | 57.92 | 19.55 | 456.67 | 245.37 | 55.09 | 67.74 | 404.44 | |
| | Patient throughput | 26.55 | 7.36 | 2.66 | 58.44 | 29.66 | 7.50 | 9.19 | 58.31 | |
| District hospitals (N=174) | Beds | 118.02 | 87.29 | 15.00 | 485.00 | 112.78 | 82.86 | 15.00 | 487.00 | |
| | Day beds | 46.45 | 39.67 | 0.00 | 257.00 | 49.22 | 40.66 | 0.00 | 257.00 | |
| | Case mix index | 0.76 | 0.17 | 0.41 | 1.52 | 0.71 | 0.10 | 0.39 | 0.99 | |

CITS results

Results of CITS model (equation (1)) are presented in Table 3 and illustrated in Figure 1. The findings indicate that the introduction of DRG did not appear to impact hospital outcomes uniformly across different hospital levels.

The piloting phase of DRG. Contrary to policymakers' expectations, the piloting phase of DRG implementation did not result in a statistically significant decrease in total hospital admissions. The weekly trends in admissions before and after the intervention showed no significant difference.

Regional hospitals responded as anticipated in terms of the number of surgery cases. During the piloting phase, the weekly average number of surgery cases increased significantly by 2.69 (CI: 0.44; 4.94) in the adult group and by 1.39 (CI: 0.33; 2.44) in the children's groups, compared to the pre-piloting period. In contrast, district and city hospitals experienced a decline in the weekly number of surgery cases, albeit these estimates were not statistically significant.

The piloting phase provides no statistically significant evidence of an impact on the ALOS. Contrary to established hypotheses, the DRG piloting phase appears to have no meaningful effect on the quality of care, as proxied by the SMR. The estimated effects for both ALOS and SMR are negligible and do not achieve statistical significance.

The nationwide phase of DRG. The results suggest that the nationwide implementation of DRG was accompanied by modest change in the trends of several outcome variables compared to the piloting period. Although the estimates are not statistically significant across all hospital levels, there was an increase in the weekly trends of average inpatient cases by 1.50 (CI: 0.06; 2.94), 0.47 (CI: -1.21; 2.16), and 0.03 (CI: -0.46; 0.53) cases per hospital per week in regional, city, and district hospitals, respectively. An increase in the weekly trend of average day cases is observed only in district hospitals, with an increase of 0.68 (CI: 0.22; 1.15) per hospital per week. Overall, the number of hospital admissions does not decrease following the implementation of DRG in Kazakhstan, which does not align with policy expectations; however, the expectations regarding day cases appear to be partially fulfilled.

As expected, there was an increase in weekly surgical cases across all levels of hospitals; however, not all estimates are statistically significant. As anticipated, DRG incentivised hospitals to increase surgical cases (i.e. treat more patients with profitable DRG). The quality of care and ALOS appear to have remained unchanged during the nationwide phase compared to the piloting phase. Despite policymakers' expectations of 20% reduction in ALOS (from an average of 10–8 days), the relative change

observed 5 months after the introduction of DRG in non-piloted hospitals was minimal.

Discussion

The DRG payment system theoretically holds significant promise for enhancing hospital efficiency by incentivising facilities to increase the intensity and productivity of service delivery. Between 2011 and 2012, the overall volume of hospital care increased slightly - from 15.8 to 15.9 hospital admissions per 100 population - even though public health expenditures remained constant at approximately 63% of total health spending.^{29,30} However, aligning with evidence from the implementation of DRG across countries,^{2,3,5,31,32} the findings of the current study suggest that its impact on the public hospital activity in Kazakhstan has been relatively limited. Several factors may contribute to this scenario. First, the rapid implementation of DRG¹¹ may have outpaced the institutional capacity of hospitals to adapt to the new system effectively. Second, under the new payment framework, Kazakhstan's hospitals have been functioning beyond their bed capacity limits. Notably, as of 2012, the average occupancy rate (see Table 2) exceeded the OECD average, suggesting high pressure on infrastructure.³³ Third, the financial frameworks established did not incorporate allocations for capital investments, particularly necessary for the development and modernisation of district hospitals.¹¹

In general, district hospitals tend to accommodate a greater number of patients due to their higher bed capacity; however, they are often equipped with less sophisticated medical technology and have fewer trained healthcare professionals.¹⁴ This situation prompts them to prioritise the admission of patients whose cases are financially 'desirable', as hospitals receive higher reimbursements under the DRG payment system for more severe cases. Consequently, this phenomenon has contributed to a challenge in retaining healthcare personnel in underserved rural areas, compelling these institutions to increase overall patient admissions to achieve financial sustainability, particularly given that nearly half of Kazakhstan's population resides in rural locales.

Moreover, the implementation of the DRG system necessitated substantial advancements in hospital management processes, which encompass improvements in financial accounting, the standardisation of medical services, and enhancements to data management systems. Such improvements require comprehensive training for hospital staff in the accurate registration and coding of patients according to DRG codes, alongside the effective use of newly adopted software and billing systems. This transition could lead to an extended adaptation phase for the system, resulting in limited short-term benefits.

Furthermore, the level of autonomy in most hospitals remained notably low, meaning that these institutions

Table 3. Summary results.

| Outcome variable | Pilotting phase | | Nationwide phase | | Absolute change | Relative change |
|---------------------------|-----------------------|---------------------|------------------------|---------------------|------------------------|-----------------------|
| | Change in level | Change in trend | Change in level | Change in trend | | |
| Regional hospitals | | | | | | |
| Inpatient cases | -9.28 (-40.89; 22.33) | 4.32 (-0.06; 8.70) | 24.84 (-28.11; 77.79) | 1.50 (0.06; 2.94) | 5.93 (-43.11; 54.96) | 0.06 (-0.50; 0.63) |
| Day cases | 0.99 (-8.41; 10.39) | 0.41 (-0.66; 1.49) | 2.37 (-6.61; 11.35) | 0.28 (-0.01; 0.57) | -0.80 (-13.57; 11.96) | -0.36 (-4.65; 3.92) |
| Surgery cases (age ≥ 15) | -1.65 (-17.66; 14.37) | 2.69 (0.44; 4.94) | 16.79 (-11.00; 44.59) | 1.23 (0.48; 1.97) | -13.19 (-37.37; 10.99) | -0.42 (-0.98; 0.13) |
| Surgery cases (age < 15) | -5.18 (-12.54; 2.18) | 1.39 (0.33; 2.44) | 11.88 (-1.48; 25.24) | -0.06 (-0.41; 0.29) | -9.31 (-20.55; 1.94) | -0.91 (-1.91; 0.09) |
| ALOS | -0.18 (-0.80; 0.45) | 0.07 (-0.02; 0.15) | 0.41 (-0.53; 1.35) | 0.00 (-0.03; 0.03) | -0.62 (-1.52; 0.28) | -0.06 (-0.14; 0.02) |
| City hospitals | | | | | | |
| Inpatient cases | 0.00 (-0.78; 0.79) | -0.07 (-0.16; 0.02) | -0.24 (-0.97; 0.49) | -0.01 (-0.04; 0.01) | 0.05 (-0.81; 0.90) | 0.06 (-1.09; 1.21) |
| Day cases | -5.07 (-42.22; 32.08) | -1.19 (-6.43; 4.05) | -32.58 (-96.33; 31.17) | 0.47 (-1.21; 2.16) | 78.20 (22.61; 133.79) | 0.43 (0.00; 0.86) |
| Surgery cases (age ≥ 15) | -2.16 (-13.63; 9.32) | -0.39 (-1.81; 1.03) | -1.06 (-15.31; 13.20) | -0.42 (-0.85; 0.01) | -3.33 (-18.89; 12.22) | -0.11 (-0.58; 0.36) |
| Surgery cases (age < 15) | -0.62 (-17.97; 16.73) | 0.60 (-1.81; 3.02) | 0.70 (-28.18; 29.57) | 0.34 (-0.43; 1.12) | 21.19 (-4.52; 46.90) | 0.73 (-0.77; 2.23) |
| ALOS | 2.83 (-1.50; 7.15) | -0.07 (-0.61; 0.48) | 2.29 (-3.45; 8.03) | 0.13 (-0.04; 0.30) | -6.45 (-12.42; -0.49) | -0.63 (-0.94; -0.31) |
| SMR | 0.45 (-0.40; 1.30) | -0.09 (-0.21; 0.03) | -0.62 (-1.99; 0.75) | 0.01 (-0.03; 0.05) | -0.84 (-2.09; 0.40) | -0.09 (-0.21; 0.03) |
| District hospitals | | | | | | |
| Inpatient cases | -0.50 (-1.15; 0.14) | -0.01 (-0.09; 0.06) | -0.87 (-1.48; -0.27) | -0.02 (-0.04; 0.00) | 0.28 (-0.42; 0.98) | 0.37 (-0.95; 1.70) |
| Day cases | 7.06 (-4.58; 18.70) | -0.38 (-1.86; 1.10) | -0.64 (-14.18; 12.91) | 0.03 (-0.46; 0.53) | 30.77 (12.85; 48.69) | 0.98 (-0.10; 2.07) |
| Surgery cases (age ≥ 15) | 7.90 (-4.39; 20.19) | -0.38 (-1.88; 1.12) | -3.64 (-16.82; 9.55) | 0.68 (0.22; 1.15) | 23.80 (6.02; 41.58) | 1.12 (-0.56; 2.79) |
| Surgery cases (age < 15) | 1.95 (-1.92; 5.83) | -0.06 (-0.54; 0.43) | 0.99 (-3.41; 5.39) | 0.00 (-0.15; 0.16) | 4.95 (-0.89; 10.79) | 0.86 (-0.94; 2.65) |
| ALOS | 0.48 (-0.65; 1.60) | 0.07 (-0.06; 0.20) | 1.29 (0.23; 2.35) | 0.03 (-0.01; 0.06) | -3.09 (-4.66; -1.52) | -0.70 (-0.90; -0.50) |
| SMR | 0.15 (-0.48; 0.77) | -0.03 (-0.10; 0.05) | -0.28 (-0.92; 0.36) | 0.00 (-0.02; 0.02) | 0.07 (-0.81; 0.95) | 0.01 (-0.09; 0.11) |
| SMR | 0.43 (-0.69; 1.55) | -0.07 (-0.20; 0.05) | 0.14 (-0.87; 1.15) | -0.01 (-0.04; 0.03) | 0.74 (-0.50; 1.98) | -3.68 (-24.84; 17.47) |

95% Confidence intervals are in parenthesis.

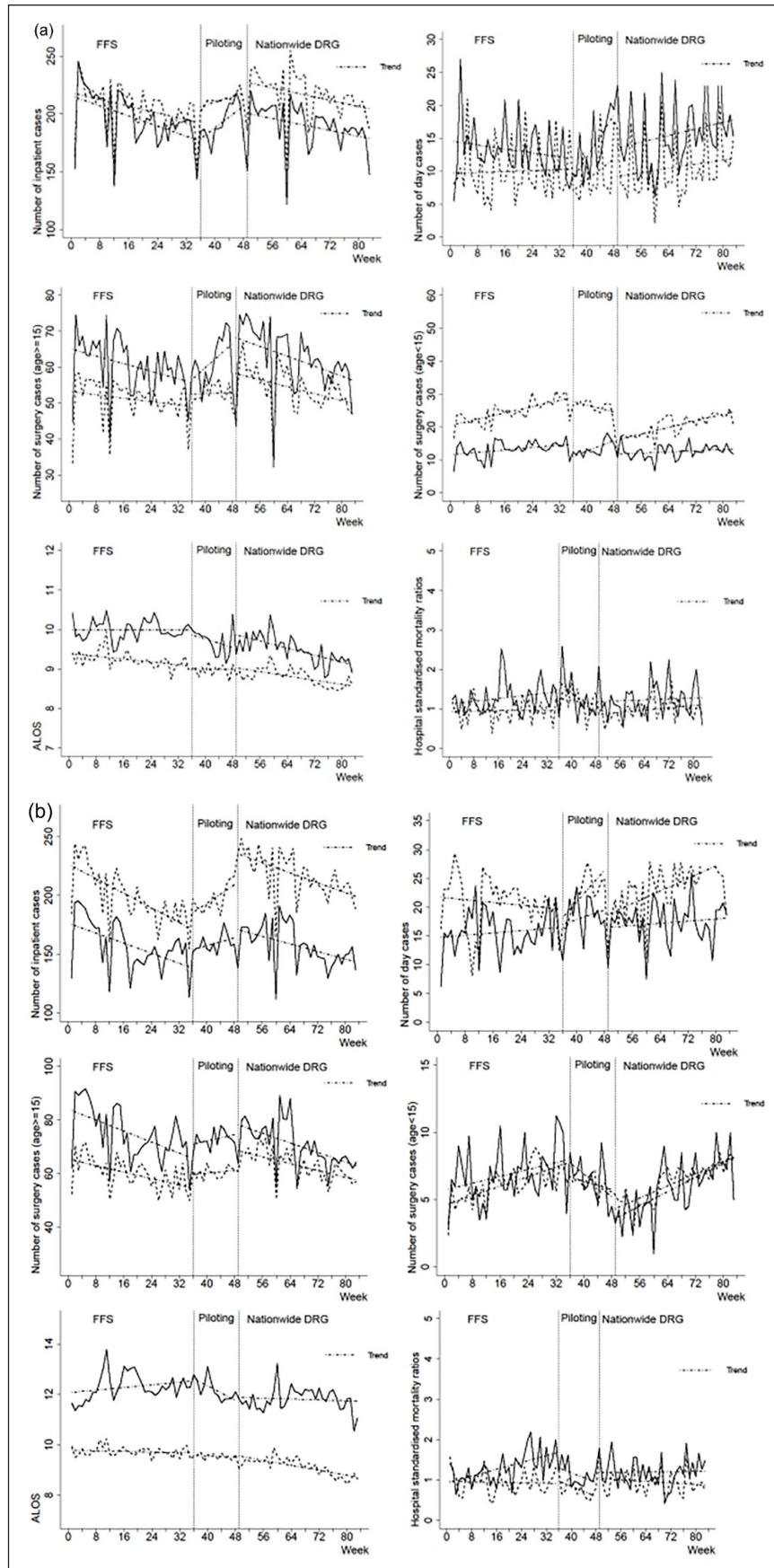


Figure I. (Continued)

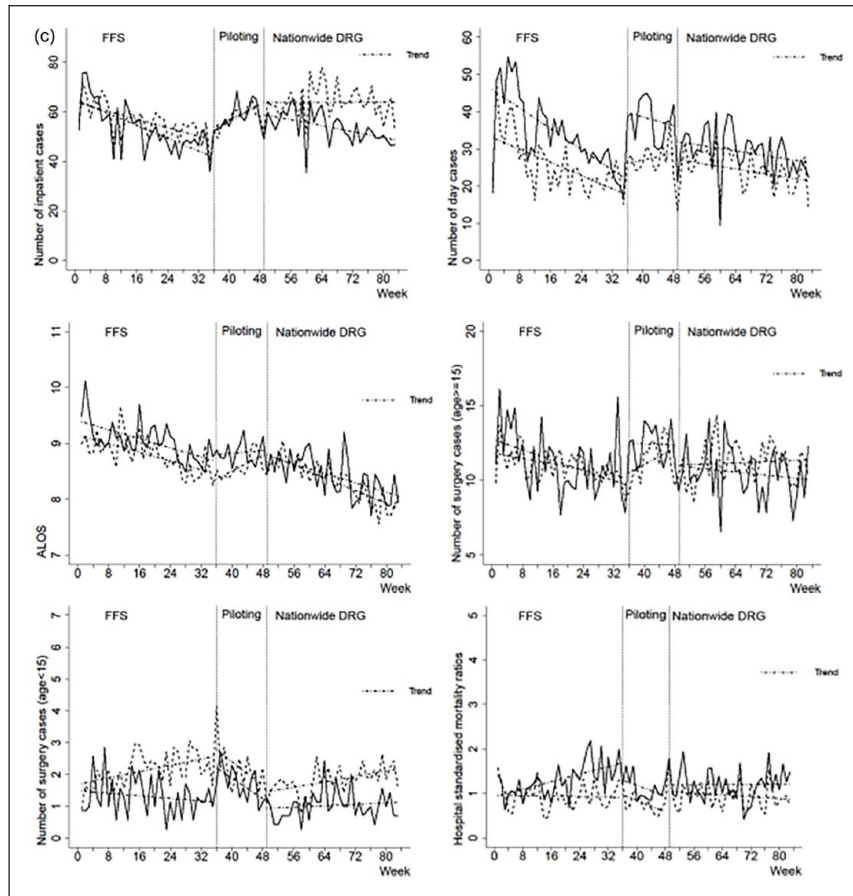


Figure 1. Weekly trends occurred in (a) regional hospitals, (b) city hospitals, (c) district hospitals, between 2011 and 2012 in Kazakhstan. Piloted hospitals (solid line) versus non-piloted hospitals (dashed line).

struggled to reallocate surplus funds across different budgetary categories. Although a handful of hospitals achieved financial autonomy in 2011, with additional facilities gaining such status the following year, these autonomous entities continued to operate as they had under the command economy model, lacking the cultural preparedness necessary for effective financial autonomy, which ultimately resulted in inefficient resource utilisation. Additionally, not all public hospitals attained national accreditation, restricting their capacity to function at optimal levels, as many were not authorised to deliver certain services that they were otherwise capable of providing.

Resistance to the new payment model was also evident. Many hospitals demonstrated reluctance to share cost data essential for the development of accurate DRG tariffs, fearing that greater financial transparency would lead to reduced funding allocations.¹¹ This reluctance was likely rooted in institutional distrust, shaped by past experience of failed reforms marked by allegations of corruption and misappropriation of funds.^{34,35} The absence of a robust accounting system further hampered the reform process. Due to limited availability of cost data, the initial DRG grouper relied heavily on estimations and required

further adjustments and refinements.¹¹ Additionally, some hospitals lacked the institutional and managerial capacity to adapt to the new system, further compounding the challenges of DRG implementation.

Lastly, primary care facilities were not sufficiently equipped to manage the increased patient influx that might have been redirected from secondary care settings, as no infrastructural improvements were made to meet this potential demand. Detailed data tracking the spillover effects between primary and secondary care were unavailable; therefore, aggregate data were sourced from the statistical yearbooks of the MoH. Despite regional variations in the number of visits to primary healthcare facilities, there was a significant decline in figures after 2012 (Supplemental File, Figures S3a, S3b). Moreover, the proportion of governmental expenditures on secondary care rose, while the proportion allocated to primary care fell after the nationwide DRG implementation. At the same time, household spending increased on both levels of care (Supplemental File, Figures S4a, S4b).

The model's primary strength lies in its capacity to facilitate the observation of initial DRG adopters, the pilot hospitals, thus enabling a thorough analysis of their swift

responses to the developed policy of nationwide DRG implementation. The incentives embedded within the policy may be subject to improvements, potentially resulting in changes in providers' responses over the long term. Another significant advantage is that the CITS model permits the control of time-varying confounders, a capability that the DID model lacks, alongside a reduced risk of selection bias achieved through the embedded matching techniques, which similar studies have failed to implement, according to the review.⁴

A general limitation relates to assessing the quality of care in our analysis. Given that hospitals are incentivised to reduce ALOS under DRG,³² we did not measure whether patients are more likely to die outside the hospital or be readmitted after rapid discharge. Although the literature has criticised the use of in-hospital mortality as a quality measurement,^{4,36} this indicator is widely used in the absence of other quality indicators. Moreover, this study is limited by the absence of clinical outcome and patient-reported measures due to the unavailability of such data. These data are typically collected in paper-based formats, and the data provider lacked sufficient resources to conduct the necessary de-identification procedures. Matching can be improved by using the actual size of a hospital instead of bed numbers as a proxy. The fourth limitation is that the outcome variables were not adjusted for the case-mix index. As hospitals were reimbursed under FFS in 2011, the datasets lacked the necessary variables to robustly calculate the case mix index.

The study lacked information on some of the potential unintended consequences of DRG implementation, such as upcoding and cream-skimming. However, one indication of such provider behaviour was a rapid shift from DRG to GB in district hospitals just one year after the nationwide implementation of DRG, while regional and city continued using DRG system. According to a telephone correspondence with a former policymaker involved in the DRG rollout, district hospitals had significantly increased the number of unnecessary admissions, whereas regional and city hospitals increased the number of hospital episodes coded as complicated pregnancies and caesarian sections in an effort to maximise funding. This led to unnecessary travel for rural patients to the nearest regional or city level hospitals, resulting in financial benefits for those hospitals, but placing an extra burden on rural patients. Although the DRG payment formula included adjustment coefficients for remotely located hospitals, these measures proved insufficient. Ultimately, this led to a later reform aimed at preventing the harmful effects. Under the GB model, district hospitals received a fixed budget covering both primary and secondary care, allowing for more efficient and integrated management of healthcare services.

Further empirical research is needed to gain a deeper understanding of how changes in financial incentives affect

hospital activity and quality of care, necessitating the use of extensive data to effectively measure the spillover effects of this policy and the long-term impact of DRG.

Conclusion

This study provides empirical evidence of both the desired and detrimental effects associated with the implementation of DRG in Kazakhstan. The analysis found no significant effects from the transition from FFS to DRG on hospital activity and the quality of care. While policymakers did not specify a desired ceiling or floor range for healthcare impacts, the analysis suggests that hospital activity is likely to increase. This evidence should be approached with caution due to the limitations in the availability of proxies for assessing the quality of care. Our results indicate that the quality of care, as measured by standardised in-hospital mortality rates, has not been affected by the increase in hospital activity.



The experience of introducing the DRG system in Kazakhstan reveals that healthcare providers were not fully prepared for the new national payment system, which is why the DRG system is not yet fully operational. This study indicates that the DRG-based reimbursement system does not enhance horizontal partnerships among providers as policymakers had hoped. Ideally, the introduction of a new provider payment system necessitates a preparatory phase during which an integrated framework is established in advance, enabling the payment system to function at its full potential.

In the context of a significant revision of mandatory health insurance schemes, policymakers are currently considering a potential overhaul of hospital payment systems nationally, moving away from DRG-based payment models towards more diversified financing strategies. Insights drawn from this study may offer valuable understanding of the limited efficacy of the recently implemented payment systems and the lessons learned from them. Importantly, new payment reforms should not be pursued in isolation but rather embedded within broader health system strengthening efforts. This includes foundational investments in effective governance mechanisms, institutional capacity building, and health infrastructure, without which payment reforms are unlikely to achieve their intended objectives.

Acknowledgements

We are grateful to the Islamic Development Bank, IDB Merit Scholarship Programme for High Technology. We also acknowledge the support of Maqсут Narikbayev University for covering the open access publication fees.

ORCID iDs

Daliya Kaskirbayeva  <https://orcid.org/0000-0002-4221-589X>
Tim Ensor  <https://orcid.org/0000-0003-0279-9576>

Ethical approval

No ethical approval was required for this research, as it is based on secondary data (fully anonymised).

Ethical considerations

This study utilised secondary, de-identified data provided by the Ministry of Health of the Republic of Kazakhstan in response to an official data request. As the dataset did not involve direct interaction with human participants or contain identifiable personal information, the Faculty of Medicine and Health Research Ethics Committee at the University of Leeds determined that the study does not constitute research involving human subjects.

Consent to participate

This study does not involve human participants, and therefore, informed consent was not applicable.

Author contributions

| Roles or tasks | Name |
|---|--|
| Conception or design of the work | Daliya Kaskirbayeva, Tim Ensor |
| Data collection | Daliya Kaskirbayeva |
| Data analysis and interpretation | Daliya Kaskirbayeva, Tim Ensor, George Ellison, Silviya Nikolova |
| Drafting the article | Daliya Kaskirbayeva, Tim Ensor, George Ellison, Silviya Nikolova |
| Critical revision of the article | Daliya Kaskirbayeva, Tim Ensor, George Ellison |
| Final approval of the version to be submitted | Daliya Kaskirbayeva, Tim Ensor, George Ellison, Silviya Nikolova |

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was funded by a doctoral award from the Islamic Development Bank, IDB Merit scholarship Programme for High Technology.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data availability statement

Raw data were generated at the Ministry of Health of the Republic of Kazakhstan. Derived data supporting the findings of this study are available from the corresponding author [DK] on request.

Reflexivity statement

The authors include two females and two males and span multiple levels of seniority. While two of the authors specialise in

provider payment reforms in post-soviet countries, the third is a data scientist and the fourth is a health economist. All four authors have extensive experience in causal inferences in qualitative research.

Supplemental material

Supplemental material for this article is available online.

References

1. Ndayishimiye C, Tambor M, Behmane D, et al. Health care provider payment schemes and their changes since 2010 across nine Central and Eastern European countries - a comparative analysis. *Health Policy* 2025; 153: 105261.
2. Milstein R and Schreyögg J. The end of an era? Activity-based funding based on diagnosis-related groups: a review of payment reforms in the inpatient sector in 10 high-income countries. *Health Policy* 2024; 141: 104990.
3. Mehmood A, Ahmed Z, Ghailan K, et al. Implementation of healthcare financing based on diagnosis-related group in three WHO regions; Western Pacific, South East Asia and Eastern Mediterranean: a systematic review. *J Health Manag* 2023; 25: 404–413.
4. Ghazaryan E, Delarmente BA, Garber K, et al. Effectiveness of hospital payment reforms in low- and middle-income countries: a systematic review. *Health Policy Plan* 2021; 36: 1344–1356.
5. Zou K, Li HY, Zhou D, et al. The effects of diagnosis-related groups payment on hospital healthcare in China: a systematic review. *BMC Health Serv Res* 2020; 20: 112.
6. Ndayishimiye C, Tambor M, Behmane D, et al. Factors influencing health care providers payment reforms in central and Eastern European countries. *INQUIRY* 2024; 61: 00469580241287626.
7. Moreno-Serra R and Wagstaff A. System-wide impacts of hospital payment reforms: evidence from Central and Eastern Europe and Central Asia. *J Health Econ* 2010; 29: 585–602.
8. Mathauer I and Wittenbecher F. Hospital payment systems based on diagnosis-related groups: experiences in low- and middle-income countries. *Bull World Health Organ* 2013; 91: 746–756A.
9. Brügger U and Eichler K. Impact of introducing a DRG reimbursement system in an acute inpatient hospital setting: a literature review. In: *HTAi 7th annual meeting*, Dublin, Ireland, 6–9 June 2010.
10. MoH. Decree #489. *Ob utverzhdanii perechnya medicinskih organizacij pilotnogo vnedreniya metodiki oplaty stacionarnykh i stacionarozameshchayushchih medicinskih uslug po kliniko-zatratnym gruppam*. [Russian]. [Online]. The Ministry of Health of The Republic of Kazakhstan, 2011, <https://online.zakon.kz/> (accessed September 2017).
11. Chanturidze T, Esau M, Hölzer S, et al. Introducing diagnosis-related groups in kazakhstan: evolution, achievements, and challenges. *Health Policy* 2016; 120: 987–991.
12. MoH. Decree #1113. *Ob utverzhdanii Gosudarstvennoj programmy razvitiya zdavoohranenija Respubliki Kazakhstan 'Salamatty Kazakhstan' na 2011 - 2015 gody*. [Russian]. [Online]. The Ministry of Health of The Republic of Kazakhstan, 2010, <https://online.zakon.kz/> (accessed March 2021).

13. Katsaga A, Kulzhanov M, Karanikolos M, et al. Kazakhstan health system review. *Health Syst Transit* 2012; 14: 1–154.
14. OECD. *OECD reviews of health systems: Kazakhstan 2018*. Paris: OECD Publishing; 2018. Available at: <https://doi.org/10.1787/9789264289062-en>
15. Weissert WG. Estimating the long-term care population: prevalence rates and selected characteristics. *Health Care Financ Rev* 1985; 6: 83–91.
16. Zhang F, Wagner AK, Soumerai SB, et al. Methods for estimating confidence intervals in interrupted time series analyses of health interventions. *J Clin Epidemiol* 2009; 62: 143–148.
17. Lagarde M. How to do (or not to do) . . . assessing the impact of a policy change with routine longitudinal data. *Health Policy Plan* 2012; 27: 76–83.
18. Bernal JL, Cummins S and Gasparrini A. The use of controls in interrupted time series studies of public health interventions. *Int J Epidemiol* 2018; 47: 2082–2093.
19. Bernal JL, Soumerai S and Gasparrini A. A methodological framework for model selection in interrupted time series studies. *J Clin Epidemiol* 2018; 103: 82–91.
20. St. Clair T, Cook TD and Hallberg K. Examining the internal validity and statistical precision of the comparative interrupted time series design by comparison with a randomized experiment. *Am J Eval* 2014; 35: 311–327.
21. Ramsay CR, Matowe L, Grilli R, et al. Interrupted time series designs in health technology assessment: lessons from two systematic reviews of behavior change strategies. *Int J Technol Assess Health Care* 2003; 19: 613–623.
22. Cook TD, Campbell DT and Shadish W. *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin, 2002.
23. Sommersguter-Reichmann M. 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* 2000; 3: 309–321.
24. Wing C, Simon K and Bello-Gomez RA. Designing difference in difference studies: best practices for public health policy research. *Annu Rev Public Health* 2018; 39: 453–469.
25. Jandoc R, Burden AM, Mamdani M, et al. Interrupted time series analysis in drug utilization research is increasing: systematic review and recommendations. *J Clin Epidemiol* 2015; 68: 950–956.
26. Pinheiro J, Bates D, DebRoy S, et al. *nlme: linear and non-linear mixed effects models*. R package version 3.1-139. 2019.
27. Linden A. Conducting interrupted time-series analysis for single- and multiple-group comparisons. *The Stata Journal* 2015; 15 2015; 15: 480–500.
28. King G and Nielsen R. Why propensity scores should not be used for matching. *Polit Anal* 2019; 27(4): 435–454.
29. MoH. Decree #57. *Ob utverzhenii Pravil oplaty za okazannye medicinskie usluzhi v ramkah garantirovannogo ob'ema besplatnoj medicinskoj pomoshchi, osushchestvlyaemyh za schet sredstv respublikanskogo byudzheta*. [Russian]. [Online]. The Ministry of Health of The Republic of Kazakhstan, 2012, <https://online.zakon.kz/> (accessed September 2017).
30. MoH. *Zdorov'e naseleniya Respubliki Kazakhstan i deyatel'nost' organizatsiy zdravookhraneniya v 2012 g.: statisticheskiy sbornik [Health of the population of the Republic of Kazakhstan and activities of healthcare organizations in 2012: statistical yearbook]*. 2013.
31. Bredenkamp C, Bales S and Kahur K. *Transition to diagnosis-related group (DRG) payments for health: lessons from case studies*. World Bank Publications, 2019.
32. Busse R, Geissler A, Aaviksoo A, et al. Diagnosis related groups in Europe: moving towards transparency, efficiency, and quality in hospitals? *BMJ* 2013; 346: f3197.
33. OECD. *Health at a Glance 2023: OECD indicators*. Paris: OECD Publishing; 2023. Available at: <https://doi.org/10.1787/7a7afb35-en>
34. World Health Organization Regional Office for Europe, European Observatory on Health Systems and Policies. *Health care systems in transition: Kazakhstan*. Copenhagen: WHO Regional Office for Europe; 1999. Available at: <https://iris.who.int/handle/10665/108257>
35. Kulzhanov M, Rechel B, World Health Organization Regional Office for Europe, European Observatory on Health Systems and Policies. *Kazakhstan: Health System Review*. Health Syst Transit [Internet]. 2007 Nov 14;9(7). Available at: <https://eurohealthobservatory.who.int/publications/i/kazakhstan-health-system-review-2007>
36. Pitches DW, Mohammed MA and Lilford RJ. What is the empirical evidence that hospitals with higher-risk adjusted mortality rates provide poorer quality care? A systematic review of the literature. *BMC Health Serv Res* 2007; 7: 91.