

DESCRIBING SERBIAN HOSPITAL ACTIVITY USING AUSTRALIAN REFINED DIAGNOSIS RELATED GROUPS: A CASE STUDY IN VOJVODINA PROVINCE

OPIS SRBSKE BOLNIŠNIČNE DEJAVNOSTI Z UPORABO AVSTRALSKIH SKUPIN PRIMERLJIVIH PRIMEROV DIAGNOZ: PRIKAZ ZA VOJVODINO

Aleksandar P MEDAREVIC^{1*}

¹University of Belgrade, Faculty of Medicine, Dr Subotica 8, 11000 Belgrade, Serbia

Received: Mar 21, 2019
Accepted: Oct 16, 2019

Original scientific article

ABSTRACT

Keywords:

average length of stay, outliers, reduction of variance, AR-DRG, casemix, healthcare financing, clinical coding

Introduction: AR-DRG system for classification hospital episodes was implemented in Serbia to improve efficiency and transparency in the health system.

Methods: L3H3, IQR, and 10th-95th percentile methods were used to identify outlier episodes in the classification. Classification efficiency and within-group homogeneity were measured by an adjusted reduction in variance (R2) and a coefficient of variation (CV).

Results: There were 246,131 hospital episodes with a total 1,651,913 bed days from 14 hospitals. All episodes were classified into 652 groups of which 441 had CV lower than 100%. "Medical groups" accounted for 51% of groups and for 72% of episodes. Chemotherapy and vaginal delivery were the highest volume groups, with 5% and 4% of total episodes. Major diagnostic category 6 (MDC 6, Diseases of the digestive system) was the highest volume MDC, accounting for 11% of episodes. "Day-cases" and "prolonged hospitalisation" accounted for 21% and 3% of episodes, respectively. The average length of stay varied from 5.6 to 8.2 days. Adjusted R2 was 0.3 for untrimmed data. Trimming by L3H3, IQR, and 10th-95th percentile method improved the value of adjusted R2 to 0.61, 0.49, and 0.51, identifying 24%, 7%, and 7% of total cases as outliers, respectively. Mental diseases (MDC 19) remained the lowest adjusted R2 in untrimmed and trimmed datasets.

Conclusion: A long length of stay and a small percentage of "day-cases" characterized hospital activity in Vojvodina. Trimming methods significantly improved DRG efficiency. Future studies should consider cost data.

IZVLEČEK

Ključne besede:

povprečno trajanje hospitalizacije, odstopanja, zmanjšanje neskladij, AR-DRG, razvrščanje casemix, financiranje zdravstva, klinično kodiranje

Uvod: Klasifikacijski sistem hospitalizacij AR-DRG je bil v Srbiji vpeljan za izboljšanje učinkovitosti in preglednosti zdravstvenega sistema.

Metode: Za prepoznavanje odstopanj v klasifikaciji so bile uporabljene metode L3H3, IQR, in 10.-95. percentila. Učinkovitost klasifikacije in homogenosti znotraj skupine primerljivih primerov diagnoz sta bili izmerjeni s prilagojenim zmanjšanjem neskladij (R2) in koeficientom razlik (CV).

Rezultati: Zabeleženih je bilo 246.131 hospitalizacij s skupnim številom 1.651.913 preležanih dni v 14 bolnišnicah. Vse hospitalizacije so bile razvrščene v 652 skupin, od tega jih je imelo 441 vrednost CV nižjo od 100 %. »Kliničnih skupin« je bilo 51 % od vseh skupin in 72 % od vseh hospitalizacij. Kemoterapija in vaginalni porod sta po obsegu največji skupini s 5 % in 4 % vseh hospitalizacij. Skupina primerljivih primerov diagnoz 6 (MDC 6, boleznih prebavnega sistema) je največja, in sicer zavzema 11 % vseh hospitalizacij. »Dnevne obravnave« obsegajo 22 % in »podaljšane hospitalizacije« 3 % hospitalizacij. Povprečno trajanje hospitalizacije se giblje med 5,6 in 8,2 dni. Prilagojeni R2 se obravnava za 0,3 neprirezanih podatkov. Aplikacija metod L3H3, IQR, in 10.-95. percentila je izboljšala vrednost prilagojenega R2 na 0,61, 0,49 in 0,51, ekstremne vrednosti so se pojavljale v 24 %, 7 % in 7 % vseh primerov. Duševne bolezni (MDC 19) imajo najnižji prilagojen R2 v neprirezanih in prirezanih podatkovnih setih.

Zaključek: Bolnišnično dejavnost v Vojvodini označuje dolgotrajna hospitalizacija in nizek delež dnevnih obravnav. Metode prirezovanja so opazno izboljšale učinkovitost DRG. Prihodnje raziskave naj vključijo še stroškovni vidik.

*Corresponding author: Tel. + 38 10 642 075 798; E-mail: aleksandar.m.gplus@gmail.com

1 INTRODUCTION

Despite decades of transition, Serbia has maintained an archaic organisation of the health system inherited from the former Yugoslavia, with national social health insurance, universal coverage, public hospitals and physicians as state employees (1). The National Health Insurance Fund collects insurance revenues and distributes them to providers according to a contract between the Fund and providers. An integral part of the contract is a business plan with the planned amount of all services based on previous-year inputs rather than service delivery. Therefore, hospitals are incited to gradually increase inputs through increasing the bed occupancy rate, performing unnecessary procedures, and employing new staff, rather than focusing on the results or the quality of care.

To increase the efficiency of hospital healthcare, the World Bank has recommended the introduction of a prospective payment system based on diagnostic-related groups (DRGs) as one of the priorities in the reform of public finances (2). The DRG is a case-mix system created at Yale University for classifying hospital episodes in groups that are relatively homogeneous with respect to resource use and clinical conditions (3). Similar to other European countries, the main arguments for the implementation of the DRG system were to improve transparency and efficiency without deterioration in the quality of care (4, 5). The system improves transparency by condensing the confusingly large number of individual patients and thousands of procedures into a manageable number of groups. This allows for group analyses for comparison between states, regions, hospitals and departments (4).

DRG systems are adopted as a reimbursement base for a prospective payment system (5, 6). The prospective payment system incites hospitals to improve efficiency, to limit the services per patient, to treat more patients and to produce sufficient services meeting patient needs. The Australian redefined version of DRG (AR-DRG 6.0) was chosen as a case-mix model for implementation in Serbia (7). Different versions of AR-DRG, in original or in modified form, have been previously implemented in Germany, Romania, Slovenia, Croatia, Ireland, New Zealand etc. (8).

The DRG-based payment system consists of the classification system and the payment formula that is a product of hospital base rate and relative weights adjusted for outliers (9). The method for determining outliers is an integral element of any payment system and can be as important as the patient classification system itself (10). Outliers are essential because they can lead healthcare providers to expensive losses. The methods for handling outliers and weights are specific

for every country or even providers according to the distribution of costs, the length of stay and the type of hospital (i.e. level of care) in the data sample (11). In general, there are two types of methods for detecting outliers: parametric and non-parametric (11). Parametric methods are based on a normal distribution of episodes around arithmetic mean, whereas non-parametric are based on inter-quartile range. Standing procedure before the implementation of a certain DRG scheme involves testing several classifications during the pilot stage and subsequently adopting the most effective one according to relevant statistical parameters (12).

The study aims to describe hospital activity using the AR-DRG, to examine case-mix performances using relevant statistics and to estimate the data quality in one Serbian province. Findings could be generalised at a national level. Results will provide precious information for the creators of the new hospital payment system.

2 METHODS

The autonomous province Vojvodina is located in the northern part of Serbia, with a population of 1.9 million and a total area of 21,506 square kilometres (13). Demographic and clinical data has been obtained from five university hospitals and nine general hospitals in 2016. Patients with severe conditions were transported from general to university hospitals. All hospitals might be considered as part of a complete system.

The National Institute of Public Health collected hospital records that contain age, gender, admission and discharge date, discharge status, birthweight for new-borns, the principal diagnosis, secondary diagnoses and performed procedures. The International Classification of Diseases (ICD) 10th revision and the Australian Classification of Health Interventions (ACHI) 7th edition were used for coding diagnoses and procedures. The collected data was input into computer-based software named "grouper" created by Laeta (Laeta Pty Ltd, Randwick, New South Wales, Australia), which classifies patients into AR-DRG groups according to the algorithm (7).

The AR-DRG 6.0 classification system contains 698 groups with unique alphanumeric codes classified into 23 major diagnostic categories (MDCs). The AR-DRG version 6.x definitions manual contains full names for DRGs and MDCs (7). The first character in the DRGs code refers to the major diagnostic category. According to the next two characters, all groups can be separated into "surgical DRGs" (from 01 to 39), "medical DRGs" (from 60 to 99) and "other DRGs" (from 40 to 59). The last character represents the resource consumption. Cases with higher and variable costs are grouped into a pre-MDC category.

Error cases are assigned to error DRGs (960Z, 962Z and 963Z). Groups 801A, 801B and 801C, contain operating-room procedures unrelated to the principal diagnosis.

Trimming is a method of identifying cases based on the length of stay (LOS). LOS is calculated as a difference between admission and discharge date. All cases between lower and upper threshold are determined as inliers, whereas cases out of range are determined as outliers. Three trimming methods were used: L3H3, IQR and 10th-95th percentile method. The L3H3 method is based on the average length of stay (ALOS) of each DRG. In such a method, a lower threshold (L3) is ALOS divided by three, whereas an upper threshold (H3) is ALOS multiplied by three. In the interquartile range (IQR) method, the lower threshold is calculated by equation $Q1-1.5*IQR$, whereas the upper threshold as $Q3+1.5*IQR$. The IQR is a difference between the first quartile (Q1) and the third quartile (Q3) of the distribution of LOS. In the final, 10th-95th method, the lower threshold being equal to the 10th percentile of LOS and the upper threshold to the 95th percentile of LOS.

The coefficient of variation (CV) measures variation in LOS for each DRG, as the ratio of the standard deviation to the mean. The coefficient of variation below 100% reflects acceptable within-group homogeneity.

The total sum of squares (SST) of the LOS is defined as the sum of the squared deviations of each observation from the mean of all observations (6). The error sum of squares (SSE) is defined as the sum of squared deviations of each observation from the mean of the group into which the observation has been classified. The difference between SST and SSE is the regression sum of squares (SSR). SSR is a variation between the mean of each group and the mean of the observed dataset. The ratio of SSR to SST provides a measure of the reduction in variation (RIV) measured with the coefficient of multiple determination (R^2). R^2 represents the fraction of variation in LOS explained by the DRG. In other words, R^2 is a summary measure of the extent to which the DRG system can predict the value of an outcome variable based on the characteristics of individual patients. R^2 ranges between zero and one. The coefficient of multiple determination takes maximum value only if the number of hospital episodes and number of DRG groups are equal. Since the number of groups and sample size affect R^2 , adjusted R^2 was used (14). The statistical significance of adjusted R^2 can be measured using F statistic (15).

3 RESULTS

Data was obtained from 14 hospitals in Vojvodina province containing 246,131 hospital discharges with 1,651,913 inpatient days. Women accounted for 56% of all episodes. The average age of hospital patient was 50 years. University hospitals recorded 100,334 (40.8%) discharges, while general hospitals recorded 145,797 (59.2%) discharges.

A classification of all discharges using AR-DRG version 6.0 resulted in 652 discrete DRG groups, of which 8.3% had five episodes or less. There were 333 "medical", 280 "surgical" and 39 "other DRGs", accounting for 71.7%, 25.6% and 2.7% of discharges, respectively. R63Z (chemotherapy) accounted for the majority of inpatient discharges (4.9 per cent) with ALOS of 1.3 days. E71B (respiratory neoplasms without catastrophic complication and comorbidity) accounted for the majority of bed days (2.4 per cent) with ALOS of 7.9 days. The 20 highest volume DRGs accounted for 34.1% of discharges, while 142 highest volume DRGs accounted for 79.8% of discharges (Table 1).

Table 1. The highest volume DRGs with the average length of stay and the coefficient of variation for untrimmed data in Vojvodina province during 2016.

Diagnostic related group	ALOS	% of total episodes (%)	CV
R63Z	1.31	4.9	98%
O60Z	4.43	3.6	69%
G60B	3.32	3.0	195%
E71B	7.85	2.0	126%
Z64A	7.11	1.9	139%
O66Z	3.39	1.8	102%
J62B	2.84	1.7	233%
J11Z	2.26	1.7	172%
R61C	1.00	1.6	0%
N09Z	1.77	1.5	154%
O01B	6.13	1.3	48%
G10B	4.29	1.1	70%
961Z	10.68	1.1	101%
D11Z	2.94	1.1	71%
G67B	4.69	1.0	83%
G70B	4.44	1.0	104%
C16Z	2.52	1.0	87%
Q61B	3.42	0.9	178%
J62A	3.57	0.9	186%
K60B	6.48	0.9	87%

Approximately 1.1% of episodes were identified as errors into three DRGs. 961Z (unacceptable principal diagnosis), 963Z (neonatal diagnosis not consistent with age and/or weight), 960Z (ungroupable) accounted for 95.8%, 2.2% and 1.9% of errors, respectively.

Episodes with operating-room procedures unrelated to principal diagnosis were grouped into 801A, 801B and 801C DRGs. 801C (operating room procedures unrelated to the principal diagnosis without catastrophic complication and comorbidity) accounted for 61.6% of unrelated episodes, 801B (operating room procedures unrelated to the

principal diagnosis with severe or moderate complication and comorbidity) for 23.9% and 801A (operating room procedures unrelated to the principal diagnosis with catastrophic complication and comorbidity) for 14.5% of unrelated episodes.

MDC 6 (diseases and disorders of the digestive system) was the highest volume MDC, accounting for 11.4% of the total number of episodes. The five highest volume MDCs accounted for 46.3% of the total number of episodes, whereas twelve MDCs accounted for 80.5% (Table 2). The average length of stay was 6.71 days (95CI 6.67-6.75)

Table 2. Variance explained (adjusted R²) for the length of stay and percentage of outliers for untrimmed and trimmed data in Vojvodina province during 2016.

Major diagnostic category	Untrimmed data			L3H3		IQR		10 th -95 th	
	Adjusted R ²	% of total episodes (%)	% of day cases within MDC (%)	Adjusted R ²	% of outliers within MDC (%)	Adjusted R ²	% of outliers within MDC (%)	Adjusted R ²	% of outliers within MDC (%)
Pre-MDC	0.13	<1	1	0.31	23	0.31	5	0.33	11
MDC 01	0.13	5	6	0.38	22	0.27	6	0.28	10
MDC 02	0.11	2	14	0.46	18	0.29	9	0.23	5
MDC 03	0.15	4	4	0.33	13	0.23	7	0.25	7
MDC 04	0.13	6	13	0.31	29	0.23	5	0.26	9
MDC 05	0.14	9	5	0.27	20	0.25	5	0.29	9
MDC 06	0.18	11	13	0.47	30	0.47	7	0.47	7
MDC 07	0.17	3	7	0.41	24	0.32	5	0.34	10
MDC 08	0.21	5	13	0.44	33	0.34	5	0.37	8
MDC 09	0.15	8	41	0.62	22	0.53	11	0.44	5
MDC 10	0.16	3	19	0.45	31	0.35	7	0.32	6
MDC 11	0.37	5	27	0.75	34	0.68	7	0.69	6
MDC 12	0.10	2	27	0.46	37	0.57	10	0.46	6
MDC 13	0.19	5	37	0.59	21	0.49	8	0.41	5
MDC 14	0.16	8	9	0.30	10	0.37	7	0.36	6
MDC 15	0.37	2	11	0.65	21	0.54	3	0.52	7
MDC 16	0.10	2	43	0.40	57	0.37	11	0.25	5
MDC 17	0.27	10	77	0.71	13	0.37	7	0.35	2
MDC 18	0.12	1	4	0.34	18	0.21	5	0.20	10
MDC 19	0.09	1	2	0.18	28	0.16	5	0.18	13
MDC 20	0.11	1	9	0.31	44	0.33	10	0.32	10
MDC 21	0.16	1	21	0.49	38	0.39	7	0.29	7
MDC 22	0.26	<1	2	0.68	33	0.37	4	0.43	11
MDC 23	0.19	4	17	0.56	25	0.52	6	0.49	5
Unrelated DRGs	0.13	1	6	0.29	30	0.24	6	0.29	8
Error DRGs	0.01	1	8	0.06	37	0.02	1	0.01	5
Overall DRGs	0.30	100	21	0.61	24	0.49	7	0.51	7

and the median value was 4 days (Table 3). “Day-cases”, which did not require an overnight stay in the hospital, accounted for 21.4% of total episodes. R63Z with 10,696 episodes accounted for the majority of “day-cases”. After excluded “day-cases”, ALOS increased to 8.24 days (95CI 8.19-8.29) whereas the median value remained 4 days. There were 7,161 (2.9%) episodes lasting more than 28 days defined as “prolonged hospitalisation”. Among “prolonged hospitalisation”, 1,368 (19.1%) were classified in MDC 19 (mental diseases and disorders). Without “prolonged hospitalisation” ALOS dropped down to 5.56 days (95CI 5.54-5.58) with the median value of 3. ALOS amounted to 5.98 days (95CI 5.96-6.00) with the median value of 4 days, excluding “day-cases” and “prolonged hospitalisation” altogether.

Table 3. The average length of stay and percentage of outliers in university and general hospitals for untrimmed and trimmed data in Vojvodina province during 2016.

Trimming method	ALOS	ALOS (university hosp.)	ALOS (general hosp.)	% of outliers (university hosp.) (%)	% of outliers (general hosp.) (%)
Untrimmed data	6.71	6.88	6.59	-	-
L3H3	7.02	6.89	7.09	24	24
IQR	5.65	5.56	5.71	8	6
10th-95th	5.78	5.73	5.82	8	6

For untrimmed data, the highest ALOS was 73.3 days for L02A (operative insertion of peritoneal catheter for dialysis with catastrophic or severe complication and comorbidity). Among DRGs with more than five episodes, the highest ALOS was 52.4 days for P62Z (neonate, admission weight 750-999 g) with seven episodes.

ALOS for “surgical”, “medical” and “other DRGs” was 5.94, 6.98 and 6.84 days, respectively.

The average length of stay seen in university hospitals (6.88 days) was higher than in general hospitals (6.59 days) (Table 3). Among the DRGs seen in both types of hospitals, 282 DRGs had higher ALOS in university examples, whereas 244 DRGs had higher ALOS in general examples.

Outliers accounted for 24% (H3L3 method), 7% (IQR) and 7% (10th-95th) of total episodes, covering 20.6% (H3L3), 21.7% (IQR) and 19.7% (10th-95th) of total bed days (Table 2). The H3L3 method increased ALOS to 7.02 (Table 3) with the maximum ALOS of 105.0 days for L02A group. Within DRGs with more than five episodes, the maximum ALOS was 44.25 days for P63Z (neonate, admission weight 1,000-1,249 g without significant operating room procedure) with 16 episodes.

There were 27 DRGs with CV below 20%, 70 DRGs with CV below 50%, and 441 DRGs with CV below 100% that accounted for 2.7%, 5.9% and 53.6% of total episodes, respectively. G65A (gastrointestinal obstruction with catastrophic or severe complication and comorbidity) had the highest CV of 341%. Nine more groups had CV greater than 200%. In the group of highest volume DRGs, nine groups had CV below 100% (Table 1). Trimming increased homogeneity to 91%, 95% and 100% of DRGs with CV below 100%, also reducing maximum CV in the dataset (Table 4). Maximum variation after the H3L3 trimming method was 96% for L02A. For the IQR method, the highest value for CV was 158% for I79A (pathological fracture with catastrophic complication and comorbidity), whereas for the 10th-95th method this was 153% for V60Z (alcohol intoxication and withdrawal).

Table 4. The total number of DRGs and the number of DRGs with CV <100% in trimmed and untrimmed data in Vojvodina province during 2016.

Trimming method	Total number DRGs *	Number of DRGs with CV <100%	% of total episodes CV <100% (%)
Untrimmed data	639	441	53.6
L3H3	634	634	100.0
IQR	639	607	94.1
10 th -95 th	618	562	83.1

*Number of DRGs after excluding DRGs with one episode

Adjusted R^2 for untrimmed data was 0.30 (Table 2). Exclusion of “day-cases” decreased adjusted R^2 to 0.27, whereas exclusion of “prolonged hospitalisation” increased adjusted R^2 to 0.36. Adjusted R^2 for untrimmed data was 0.30 after exclusion of “day-cases” and “prolonged hospitalisation” together. Data trimming increased adjusted R^2 to 0.61 (L3H3 method), 0.49 (IQR) and 0.51 (10th-95th percentile) (Table 2). The L3H3 trimming method resulted in the maximum adjusted R^2 , as well as the greatest number of outliers. MDC 16 (diseases and disorders of the blood and blood-forming organs) had the highest proportion of outliers for H3L3 and IQR, with 57% and 11% of episodes (Table 2). Even after trimming, adjusted R^2 for some MDCs remained relatively low (less than 0.25). The lowest adjusted R^2 , except for errors, was for MDC 19, in both untrimmed and trimmed data (Table 2). MDC 11 (diseases and disorders of the kidney and urinary tract) had the maximum value for adjusted R^2 . The most significant improvement in adjusted R^2 compare to untrimmed value was for MDC 12 (diseases and disorders of the male reproductive system).

Trimming improved adjusted R^2 for medical rather than for “surgical DRGs”. However, the adjusted R^2 for “surgical DRGs” remained above values for “medical” and for “other DRGs” after all trimming methods (Table 5).

4 DISCUSSION

The average length of stay is a standard measure of hospital activity (15). According to Eurostat, Serbian ALOS was 9.5, being among the highest in Europe (16). Some of the reasons for prolonged hospitalization are the inadequate planning of admissions and discharges, duplicate procedures to fulfil the annual plan, as well as the shortage of mental health and palliative care community centres, lack of within-hospital coordination and archaic definition of daily cases. Therefore, the second volume group and routine procedure O60Z-Vaginal delivery had ALOS of almost five days; or the C16Z-lens procedures that are usually performed during daily cases had ALOS of almost three days.

The proportion of outlier cases is a measure of classification effectiveness (3). A less effective case-mix will detect more outliers, whereas a more effective classification will allow outliers to be assigned to inliers.

The most common trimming method in Australia was H3L3 method (17). Understandable and easily computable, this method was accepted at the beginning of AR-DRG implementation in numerous countries. The H3L3 method is based on the assumption of the normal distribution of LOS. However, the distribution of LOS is right-skewed,

Table 5. Variance explained (adjusted R^2) and the percentage of outliers within medical, surgical and others DRGs in Vojvodina province during 2016.

Trimming method	Medical DRGs		Surgical DRGs		Others DRGs	
	Adjusted R^2	% of outliers (%)	Adjusted R^2	% of outliers (%)	Adjusted R^2	% of outliers (%)
Untrimmed data	0.28	-	0.38	-	0.22	-
L3H3	0.60	29	0.62	11	0.47	16
IQR	0.47	7	0.58	8	0.39	6
10 th -95 th	0.50	6	0.62	8	0.39	8

and the arithmetic mean might become misleading. In addition to skew, with the median higher than mean, prolonged hospitalizations pull the ALOS more to the right. In this research, the percentage of prolonged episodes was 3% of episodes, in contrast with 1% in Ireland (18). The percentile method is also very sensitive to skewed data (19). Aforementioned statistical inconveniences support some non-parametric methods based on the IQR and median value as preferable methods in this stage of implementation (20, 21). The results of the IQR method were satisfactory, classifying 7% of cases as outliers with only 5% of groups with CV above 100% and ALOS of 5.65 days.

The evidence from literature suggests great variability in the proportion of outliers, depending on the algorithm, method (parametric or non-parametric, based on LOS or cost), prior experiences etc. (22). Outliers in Ireland, Germany, Austria and France accounted for 6%, 22%, about 14% and less than 1% of total episodes, respectively (22). According to the proposal from the US, the accepting proportion of outliers should be below 10% (3). Such a proportion was reached by two of three applied trimming methods in this case study and closed to the acceptable ratio suggested by Professor Fedler. Fedler highlighted that optimal threshold depends on providers and their willingness to take risk (23). Such risk differs between US and European hospitals. If hospitals are risk averse, the optimal threshold should be higher and with no more than 5% of total cases beyond the upper threshold; whereas different rules should be applied on the lower threshold (in an email from Fedler S, in September 2019). As outliers are somewhat inevitable, a kind of surcharge is necessary. The common surcharge for a long-stay outlier depends on the number of hospital days beyond the upper threshold adjusted for some types of patients (e.g. new-borns) or additionally paid for new technologies, expensive medications etc. (22, 24). Some countries prefer to amplify short-stay weights without the lower threshold in order to create an incentive for short-stay visits or day-cases (11). A system without the lower threshold might be implemented in Serbia later on, in order to raise the currently small percentage of daily cases (21% in Vojvodina). On the other hand, the lower threshold is an attempt at avoiding inappropriately early discharges, colloquially called "bloody discharges" (22). In conclusion, the implementation of the DRG system is a continuous process of improving (25). Eventually, the method should be chosen by the authority as the balance between efficiency and quality, and between competition and sustainability.

Joint activities of health institutions, the National Health Insurance Fund and the Ministry of Health resulted in the modification in the hospital payment system. According to the Rulebook for 2019, 4% of the hospital reimbursement

should be based on DRGs performance and 1% on quality indicators. Neither of the methods for determining outliers has been included yet (26). Presumably, the authority has planned to cross a point of no return and postpone implementation of the method until stakeholders become familiar with scheme and coding. Since there is no real competition between providers, the financial effects of recent reform are not easy to predict. For instance, it is questionable how hospitals would cover a potential decrease in revenues, even for a single percent in comparison to a previous year. Should hospitals cut down expenditure for medication or for salaries that are guaranteed by the law? Considering this, general hospitals could count on more or less a similar number of patients each year, so it could be presumed that more pressure would be on university hospitals. However, Keeler suggests that large hospitals have a lower risk from prospective payment and consequence of outliers, since they can make transfers between different DRGs (27).

The pre-DRG hospital payment system in Serbia instigated providers to prolong hospitalisation. Additional to poor coding practices and insufficient planning of hospital admissions resulted in great heterogeneous data. Therefore, the implementation of the DRG system should be strengthened with the implementation of solutions in different aspects of healthcare. From a clinical perspective, the utilisation of acute beds should be a privilege for acute patients, who should continue further treatment either in primary healthcare facilities or in nursing homes, afterward. Clinical and integrative pathways in support of knowledge and judgment should be directed at the highest volume conditions and diseases (measuring by the proportion of episodes) and for conditions with insufficient within-group homogeneity in order to reduce LOS and costs. Since there is no "best trimming method", the choice of method must be made based not only on the characteristics of the data sample at hand but also on the goals that health policymakers intend to reach, particularly regarding the announced rationalisation of public health facilities (11, 28). The publishing and comparison of data will certainly improve transparency in clinical practice and spending.

5 LIMITATIONS

Data quality may affect the measuring of DRG performances (3). In the Vojvodina dataset, a bit more than 1% of cases were identified as errors, which is more than in countries with longer experience in DRG implementation (29).

Episodes with the operating room (OR) procedures unrelated to the principal diagnosis were grouped into separate DRGs. Such groups accounted for around 1% of total episodes in comparison to 0.05% in Australia (30). There are no mistakes in the real sense; despite some of

them possibly being the result of miscoding. Therefore, their percentage should be under control, because they could be a result of oversight high-cost episodes. The training of clinicians on the correct usage of ICD 10 and ACHI, to avoid such errors, is necessary. Since proper coding is essential and clinicians are more focused on treatment, the authority should consider training for coders who will review, analyse and accurately assign ICD-10-AM/ACHI codes and DRGs to all inpatient episodes.

6 CONCLUSIONS

A long length of stay, a small percentage of daily cases and a substantial number of long-term episodes characterized hospital activity in Vojvodina with a great heterogeneity of coding practice. AR-DRG could explain 30% of variation for LOS in raw dataset, and between 49% and 61% in trimmed dataset. The percentage of outliers varied from 7% to 24%, depending on the trimming method.

Further studies should test different trimming algorithms and identify factors associated with high length of stay and low R2 for some MDCs using cost data rather than LOS.

CONFLICT OF INTEREST

The author declares that no conflict of interest exist.

FUNDING

No funding has been received for the conduct of this study and/or preparation of this manuscript.

ETHICAL APPROVAL

The study is based on the retrospective analysis of registry data and is, as such, exempt from ethical approval.

REFERENCES

1. Stosic S, Karanovic N. Health care economics in Serbia: current problems and changes. *Vojnosanit Pregl.* 2014;71(11):1055-61. doi: 10.2298/vsp120205002s.
2. World Bank. Serbia - doing more with less: addressing the fiscal crisis by increasing public sector productivity. Accessed March 10th, 2019 at: <https://openknowledge.worldbank.org/handle/10986/3061>.
3. Palmer G, Reid B. Evaluation of the performance of diagnosis-related groups and similar casemix systems: methodological issues. *Health Serv Manage Res.* 2001;14(2):71-81. doi: 10.1258/0951484011912564.
4. Geissler A, Quentin W, Scheller-Kreinsen D, Busse R. Introduction to DRGs in Europe: common objectives across different hospital systems. In: Busse R, Geissler A, Quentin W, Wiley M, editors. *Diagnosis related groups in Europe: moving towards transparency, efficiency and quality in hospitals.* Maidenhead: Open University Press, 2011:9-21.
5. Busse R, Geissler A, Aaviksoo A, Cots F, Häkkinen U, Kobel C, et al. Diagnosis related groups in Europe: moving towards transparency, efficiency, and quality in hospitals? *BMJ.* 2013;346:f3197. doi: 10.1136/bmj.f3197.
6. Tan SS, Geissler A, Serdén L, Heurgren M, Van Ineveld BM, Redekop WK, et al. DRG systems in Europe: variations in cost accounting systems among 12 countries. *Eur J Public Health.* 2014;24:1023-8. doi: 10.1093/eurpub/cku025.
7. Independent hospital pricing authority. *Australian Refined Diagnosis Related Groups, Version 6.x, Addendum.* Darlinghurst NSW, 2012.
8. Hindle D. Implementing DRGs in Slovenia: why the Australian variant was selected. *Aust Health Rev.* 2003;26(3):50-60. doi: 10.1071/ah030050.
9. Mathauer I, Wittenbecher F. Hospital payment systems based on diagnosis-related groups: experiences in low- and middle-income countries. *Bull World Health Organ.* 2013;91(10):746-56A. doi: 10.2471/BLT.12.115931.
10. Goldfield N. The Evolution of Diagnosis-Related Groups (DRGs): from its beginnings in case-mix and resource use theory, to its implementation for payment and now for its current utilization for quality within and outside the hospital. *Qual Manag Health Care.* 2010;19(1):3-16. doi: 10.1097/QMH.0b013e3181ccbcc3.
11. Schreyögg J, Stargardt T, Tiemann O, Busse R. Methods to determine reimbursement rates for diagnosis related groups (DRG): a comparison of nine European countries. *Health Care Manag Sci.* 2006;9(3):215-23. doi: 10.1007/s10729-006-9040-1.
12. Ghaffari S, Jackson TJ, Doran CM, Wilson A, Aisbett C. Describing Iranian hospital activity using Australian Refined DRGs: a case study of the Iranian Social Security Organisation. *Health Policy.* 2008;87(1):63-71. doi: 10.1016/j.healthpol.2007.09.014.
13. Republički zavod za statistiku. *Opštine i regioni u Republici Srbiji.* Beograd: Republički zavod za statistiku, 2017.
14. Harel O. The estimation of R2 and adjusted R2 in incomplete data sets using multiple imputation. *J Appl Stat.* 2009;36:1109-18. doi: 10.1080/02664760802553000.
15. Marazzi A, Paccaud F, Ruffieux C, Beguin C. Fitting the distributions of length of stay by parametric models. *Med Care.* 1998;36(6):915-27.
16. Eurostat Hospital discharges and length of stay statistics - statistics explained. Accessed March 14th, 2019 at: http://ec.europa.eu/eurostat/statistics-explained/index.php/Hospital_discharges_and_length_of_stay_statistics.
17. Duckett SJ. Casemix funding for acute hospital inpatient services in Australia. *Med J Aust.* 1998;169(Suppl):17-21.
18. Activity in acute public hospitals in Ireland: 2014 annual report. Accessed March 17th, 2019 at: <http://hdl.handle.net/10147/600629>.
19. Polyzos N, Karanikas H, Thireos E, Kastanioti C, Kontodimopoulos N. Reforming reimbursement of public hospitals in Greece during the economic crisis: implementation of a DRG system. *Health Policy.* 2013;109(1):14-22. doi: 10.1016/j.healthpol.2012.09.011.

20. Kulinskaya E, Kornbrot D, Gao H. Length of stay as a performance indicator: robust statistical methodology. *IMA J Manag Math.* 2005;16(4):369-81. doi: 10.1093/imaman/dpi015.
21. Hubert M, Vandervieren E. An adjusted boxplot for skewed distributions. *Comput Stat Data An.* 2008;52(12):5186-201. doi: 10.1016/j.csda.2007.11.008.
22. Cots F, Chiarello P, Salvador X, Castells X, Quentin W. DRG-based hospital payment: intended and unintended consequences. In: Busse R, Geissler A, Quentin W, Wiley M, editors. *Diagnosis related groups in Europe: moving towards transparency, efficiency and quality in hospitals.* Berkshire: Open University Press, 2011:75-92.
23. Felder S. The variance of length of stay and the optimal DRG outlier payments. *Int J Health Care Finance Econ.* 2009;9:279-89. doi: 10.1007/s10754-008-9051-1.
24. Scheller-Kreinsen D, Quentin W, Busse R. DRG-based hospital payment systems and technological innovation in 12 European countries. *Value Health.* 2011;14(8):1166-72. doi: 10.1016/j.jval.2011.07.001.
25. Vogl M. Assessing DRG cost accounting with respect to resource allocation and tariff calculation: the case of Germany. *Health Econ Rev.* 2012;2(1):15. doi: 10.1186/2191-1991-2-15.
26. Rule book on the contracting of healthcare from compulsory health insurance with healthcare service providers for 2019. *Off Gazette RS.* 2018;106.
27. Keeler EB, Carter GM, Trude S. Insurance aspects of DRG outlier payments. *J Health Econ.* 1988;7(3):193-214.
28. Markovic Z. Zakon o zdravstvenoj zaštiti i zdravstvenom osiguranju: šta je novo u zakonu. *Nova ekonomija.* 2019;60. Accessed September 24th, 2019 at: <https://novaekonomija.rs/tekstovi/šta-je-novo-u-zakonu>.
29. Averill RF, Muldoon JH, Vertrees JC, Goldfield NI, Mullin RL, Fineran EC, et al. The evolution of casemix measurement using diagnosis related groups (DRGs). Wallingford: 3M Health Information Systems, 1998.
30. Australian hospital statistics 2011-12. Accessed March 14th, 2019 at: <https://www.aihw.gov.au/reports/hospitals/australian-hospital-statistics-2011-12/data>