A STATISTICAL STUDY ON KEY PERFORMANCE INDICATORS IN HEALTH SYSTEM

Lucian MIRESCU
University of Craiova, Craiova, Romania
ORCID: https://orcid.org/0000-0003-3780-2050
Email: mirescu.lucian.m5t@student.ucv.ro
Ana-Maria Camelia POPESCU
University of Craiova, Craiova, Romania
ORCID: https://orcid.org/0000-0003-1734-0944
Email: popescu.ana.g6w@student.ucv.ro

DOI: 10.52846/MNMK.21.1.02

Abstract:
In this paper, a statistical study on key performance indicators in health system is given and two regression models with the average cost of hospitalization as the dependent variable are found. Using linear regression, it was found that a decrease in the average cost per day of hospitalization can be achieved by increasing the bed utilization rate and increasing the number of cases treated in a calendar year. Also, the decrease in the case-mix index, which can be achieved by increasing the number of resolved cases, leads to a decrease in the average cost of hospitalization per day. The obtained results were interpreted and certain measures are proposed to improve health management.

Keywords: key performance indicators, health system, econometric models, linear regression.

1. Introduction

In the last decades, most medical institutions have implemented certain key performance indicators that are specific to their activity, in order to monitor, measure and manage the performance of healthcare systems and at the same time, to ensure efficiency, effectiveness, but also equity and quality superior medical services offered to patients. Managers of medical institutions want to obtain and manage results in accordance to certain established objectives and quality standards and seem to be aware of the beneficial effect of using performance indicators on their activity, with the aim of monitoring and improving performance. However, they do not always use the statistical analysis of indicators as the basis of decisions and managerial strategies. The development of current medical services is characterized by great complexity and diversity, which implies an approach to them on a scientific basis. This approach is much more effective if one resorts to the use of appropriate statistical-econometric methods and models.

However, we do not want it to be understood that statistical analysis excludes intuition, but on the contrary, it channels it into making the best decisions, in accordance to the established goal. Ordinary reasoning, which is more or less

empirical and intuitive, is complemented by scientific reasoning, which requires that the relationships between the different indicators included in the study to be precisely defined and quantified. Thus, it becomes very important to find and develop some strategic performance indicators, as well as the link between them, with the aim of reflecting the real performance of medical service institutions.

Performance indicators are thus used by medical service institutions to monitor and evaluate their performance in relation to specific quality standards, showing certain trends and evolution over time. We can also say that the key performance indicators also help to compare the results with other similar medical service institutions, leading to the improvement of the medical services offered by identifying the problems and areas in which certain improvements are needed.

Indicator systems, identified during the search for information on the management of financial resources in medical services (see Fuchs, (1986), Leggat et all. (1998), Parmenter (2010), Bergeron (2017)) revealed the existence of several indicators frequently used to measure hospital performance. For the analysis, we selected a number of 5 representative performance indicators, namely: average duration of hospitalization, bed utilization rate, the number of cases in a calendar year, the case-mix index and the average cost per day of hospitalization.

The main goal of the work is to study the dependence between indicators and find valid econometric models, which have as dependent variable the average cost per day of hospitalization and independent variables the other indicators. We are thus interested in ascertaining the positive or negative influence of exogenous variables on the endogenous variable and measures that can be taken to decrease the average cost of hospitalization.

For this analysis we have selected a set of data collected to verify the assumptions regarding the functionality of the model. The data collected in Annex 1 are from the County Emergency Hospital in Craiova with a number of 1518 beds and refer to the period 2012-2021. The work is structured as follows. The second section presents the current state of knowledge in the field, making a brief analysis of some scientific works in the field of performance indicators in the health system. The third section contains the research methodology, which includes the main statistical tests used in order to validate multiple linear regression econometric models. In the fourth section, which also contains the novelty of the work, two valid multiple linear regression models are presented and the results obtained are interpreted.

The novelty of our paper consists in finding the correlation between the average hospitalization cost per day and the other independent indicators, as well as their influence on the dependent variable. Thus, a decrease in the average cost per day of hospitalization can be achieved by increasing the bed utilization rate, increasing the number of cases treated in a calendar year and the decrease in the case-mix index, which can be achieved by increasing the number of resolved cases. The work ends with the conclusions that include recommendations for improving the performance of health management, as well as future research directions.

2. Literature Review

Key performance indicators in the healthcare system help to define and measure organizational performance and objectives that are fundamental to the current and future success of an organization (see for instance Fuchs (1986), Leggat et all. (1998) or Parmenter (2010), Bergeron (2017)). The performance of health

indicators over the years in India is studied by Bhuyan, Kalita and Goswami (2018) and the influence of state public expenditures on the performance index of health indicators is examined with the help of the linear regression model. The study shows that the performance of the health indicators showed a positive trend over time. It was found that there is a positive and significant impact of public health expenditure on the performance index of health indicators.

Martin and Smith (2005) showed that the traditional approach to statistical analysis regarding organizational performance has been based on developing a separate regression model for each performance indicator, but this can remove valuable information by ignoring potentially important relationships between individual performance measures. Thus, it is proposed to simultaneously model an organization's performance measures using apparently unrelated regression methods. The method is illustrated using an example from English public hospitals.

In their work, the authors Barliba, Nestian and Tiţă (2012) proposed to create and validate a hospital performance management model through the lens of seven key performance indicators, using a sample of five county emergency hospitals. Among the performance indicators are the degree of occupancy of beds per hospital, average length of stay per hospital and average values of the case mix, which are also indicators used in the present work. Finally, the article highlights some conclusions regarding the strengths and weaknesses of the model, resulting from testing the relevance of key performance indicators. Burlea-Schiopoiu and Ferhati (2021) identified a set of key performance indicators in the health system with the aim of providing health managers with recommendations for the evaluation, monitoring and control of critical factors that influence the performance of the health sector during a pandemic crisis.

Chang et all (2019) set out to characterize the past, present and projected future of global health spending, focusing on the equity of spending between countries. Thus, they estimated domestic health expenditure for 195 countries and territories from 1995 to 2016. They estimated future health expenditure scenarios using an ensemble of linear mixed-effects models with time series specifications to project domestic expenditure in healthcare from 2017 to 2050. Data were standardized and converted to inflation-adjusted 2018 US dollars.

In their paper, Bokhari, Gai and Gottret (2007) found econometric models that link a country's per capita government health expenditure and per capita income to two health outcomes, such as under-five mortality and maternal mortality. For developing countries, the results show that while economic growth is certainly an important contributor to health outcomes, government spending on health is an equally important factor.

Our work studies the dependence between certain performance indicators in the health system using linear regression analysis and proposes measures that can lead to a decrease in the average cost of hospitalization per day at the Craiova County Emergency Hospital.

3. Research methodology

In this paper we proposed to test several models of multiple linear regression,

$$Y = a_0 + a_1X_1 + \cdots + a_kX_k + e,$$

to find the statistical-mathematical relationship and the influence between the variables. We will use the average cost per day of hospitalization (CM) as endogenous variable and:

- average duration of hospitalization (DMS)
- bed utilization rate (RUP)
- the number of cases in a calendar year (NC)
- the case-mix index (ICM)

as exogenous variables.

The case-mix index or the case complexity index is a number (no unit) that expresses the resources needed by the hospital in accordance with the patients treated. ICM for hospital A = Total number of weighted cases /Total number of resolved cases hospital A. Average costs per day of hospitalization were standardized and adjusted with inflation. For the econometric models we have rigorously gone through the stages of specification, parameterization, testing and decision, focusing on their validation (see Brooks (2008), Maddala, G. S. (2001), Wooldridge, J. (1999)). However, for the stationarity of the data series, the Dickey-Fuller test could not be used, due to the relatively small number of observations. Multicollinearity testing was performed with Klein's test and variance inflation factor (*VIF*) calculation. Thus, if $R_y^2 < R_k^2$ there is multicollinearity, where R_y^2 represents the R-squared resulting from the regression model between variable *Y* and the independent variables, and R_k^2 represents the R-squared resulting from the regression model between variable X_k and the others independent variables. Variance inflation factor for the X_k variable is $VIF = \frac{1}{1-R_k^2}$ and if *VIF*>10, then the X_k variable must be eliminated.

Thus, to achieve the assessment of the parameters we used the least squares method, where the determination coefficient R^2 shows us the percentage by which the influence of the significant factors is explained. *The t-Student test* was used to test the null hypothesis H_0 (the coefficients are not significantly different from 0) and the alternative hypothesis H_1 (the coefficients are significantly different from 0). *The F test* verifies the null hypothesis H_0 **(**all the coefficients are not significantly different from 0) and the alternative hypothesis H_1 **(**there is at least a coefficient different from 0).

Moreover, we used: Durbin-Watson test to check the autocorrelation of first order or independence of the model errors, Breusch-Godfrey test for checking the higher order autocorrelation of the errors, Jarque-Bera test in order to see if the model errors follow a normal distribution or not and White test to check homoscedasticity or heteroscedasticity of the regression model.

4. Results and discussion

Making an empirical analysis of the evolution of the five indicators, it can be found in Figure 1, that the average duration of hospitalization is between 5.76 and 6.01 days, the highest value being in 2014 and the lowest in 2018. Between 2014 and 2018, the trend is decreasing, and starting with 2018, an increase in the average duration of hospitalization until 2020 is observed, also due to the Covid crisis, the fluctuations being, however, quite small and insignificant.

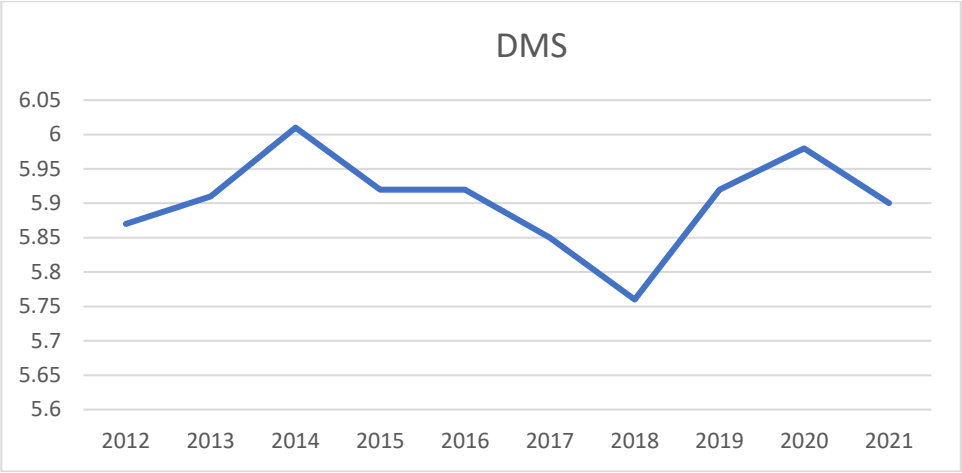

Figure 1. Average duration of hospitalization

Regarding the bed utilization rate, there is a constant weighted decrease until 2019, followed by a sharp decrease, by 35% in 2020, and which was maintained in 2021, due to the restrictions imposed by the Covid pandemic regarding admissions to the Craiova Emergency Hospital.

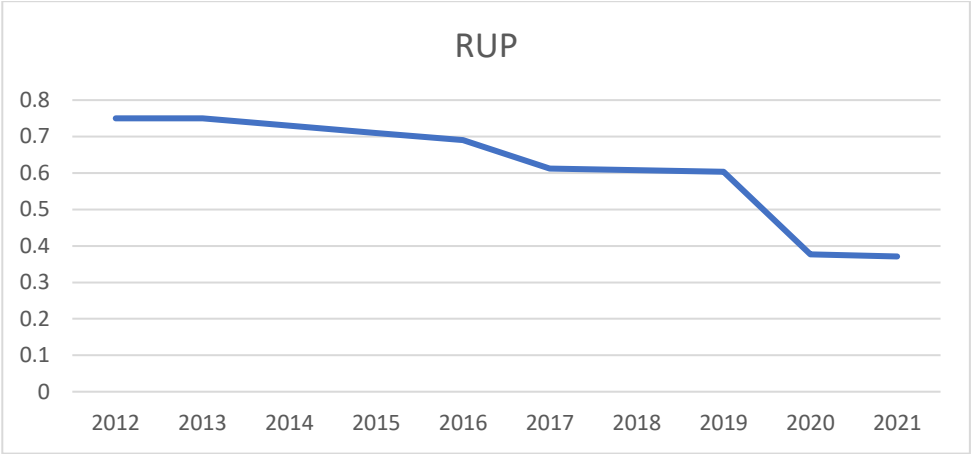

Figure 2. Bed utilization rate

Analyzing the Case Complexity Index, an increase in it can be observed in 2019 and continued in 2020 and 2021. We can say that this increase in the last two years analyzed is also due to the Covid crisis, which generated a decrease in the number of solved cases among patients.

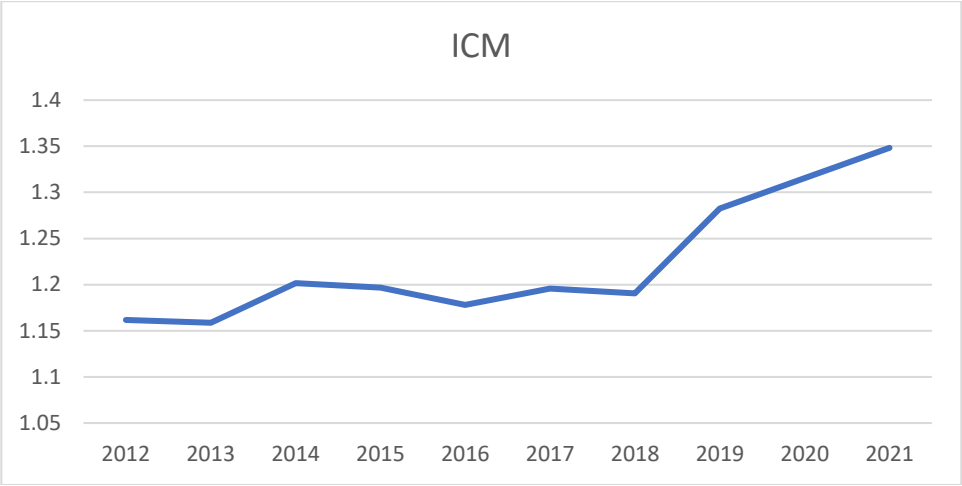

Figure 3. Case complexity index

The number of cases in a calendar year is closely related to the bed utilization rate and had a downward trend, accentuated during the pandemic, falling by 40% in 2020 compared to 2019.

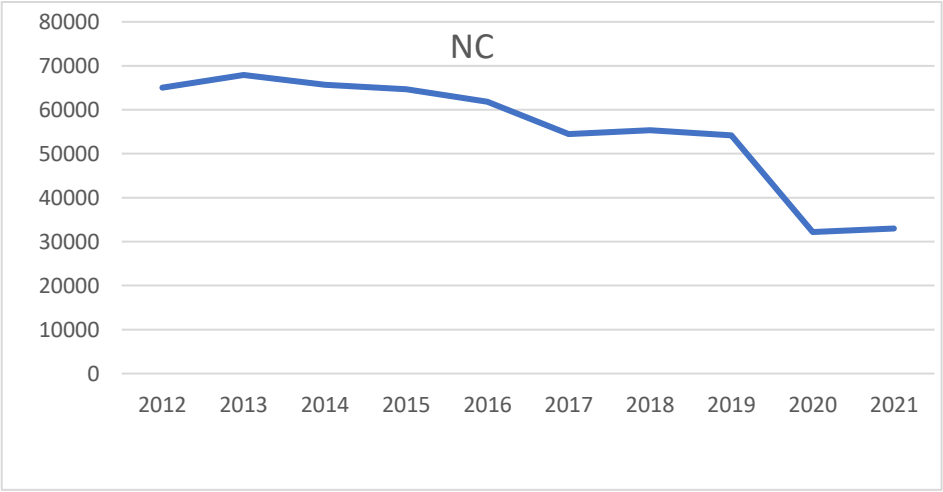

Figure 4. Number of cases in a calendar year

Finally, the evolution of the average cost per day of hospitalization (RON) is an increasing one, due to inflation and accentuated in 2020 and 2021 by the Covid crisis. We specify that these are comparable prices, being adjusted with the inflation rate. Between 2019 and 2020, the increase was 72% of the average hospitalization cost. We forecast an increase of at least 30% in 2022, mainly due to inflation and the increase in the price of electricity and natural gas.

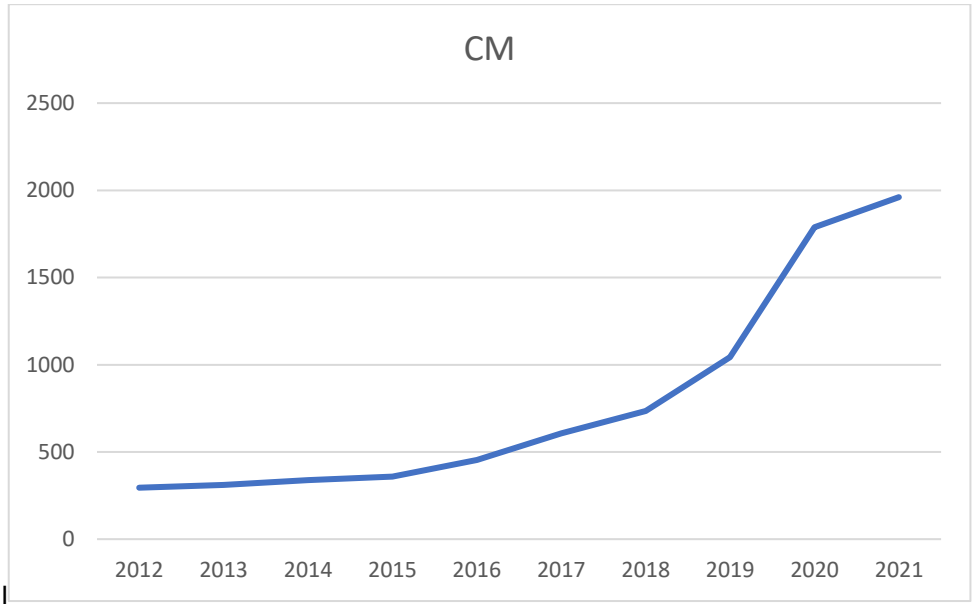

Figure 5. Average cost per day of hospitalization (RON)

In what follows, we will use multiple linear regression, having the average cost per day of hospitalization as an endogenous variable and the other indicators as exogenous variables, with the aim of finding the connection between these indicators and possible measures that can be taken to lower the cost of hospitalization per day of patients.

4.1 Econometric models

In this section we will present two linear regression models. It should be noted that several combinations of independent variables were tested, but only these models passed all regression validation tests. It is found that the coefficient of determination R-squared, which shows the percentage in which the performance indicators in the health system influence the average cost of hospitalization, has values higher than 95%.

The first econometric model uses as exogenous variable the number of cases treated in a calendar year, estimating the parameters using the least squares method.

Analyzing the data from Table 1 from a statistical point of view, we find that the probability related to the t-Student test are less than 0.05 and confirm that the parameter corresponding to independent variable is significantly different from 0, and the value of the test F=201.7 and the attached probability attests to the validity of the model. The values of the Jarque-Bera test and its associated probability greater than 0.05 (JB=0.08; the associated probability is 95.7%) confirm the normal distribution of errors (Skewness = -0.12, Kurtosis = 2.61). Regarding the Durbin Watson test (DW=2.14), it is found that there is no autocorrelation of first order at the level of the series of residuals, the errors being independent. The Breusch-Godfrey test confirms the absence of higher order autocorrelation of the errors (F-statistic=0.08; the probability is 87.5%). Finally, the validity of the model is also confirmed by the White

test (F-statistic=0.26, the attached probability is 89%), confirming the homoscedasticity of the model.

Dependent Variable: CM
Sample: 2012-2021

Table 1. Regression model 1

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3373.901	186.4423	18.09622	0.0000
NC	-0.046632	0.003283	-14.20305	0.0000
R-squared	0.961855	Mean dependent var		789.1800
Adjusted R-squared	0.957087	S.D. dependent var		618.7171
S.E. of regression	128.1700	Akaike info criterion		12.72145
Sum squared resid	131420.4	Schwarz criterion		12.78197
Log likelihood	-61.60724	F-statistic		201.7267
Durbin-Watson stat	2.143395	Prob(F-statistic)		0.000001
Breusch-Godfrey Serial Correlation LM Test:				
F-statistic	0.136325	Probability		0.875187
Obs*R-squared	0.434664	Probability		0.804663
Jarque-Bera test	0.086096	Prob. (J-B)		0.957865
Skewness	-0.123860	Kurtosis		2.618862
White Heteroskedasticity Test:				
F-statistic	2.385202	Probability		0.162208
Obs*R-squared	4.052881	Probability		0.131804

Source: Authors' contribution with software EViews.

The regression equation has the form:

$$\mathrm{CM} = -0.046632 \cdot \mathrm{NC} - 1804.131 + \mathrm{e}$$

The proposed econometric model explains 96.18% the variation of the dependent indicator studied according to the coefficient of determination.

The increase of NC by 1% leads to the decrease of the CM by 0.04%. The first econometric model shows a close correlation between the average cost per day of hospitalization on the one hand and the number of cases treated in a calendar year on the other hand. Thus, an increase of the number of treated cases, leads to a decrease in the average cost per day of hospitalization of patients.

The second econometric model uses as exogenous variables the case-mix index and the bed utilization rate in a calendar year, the estimation of the parameters being done by the method of least squares.

Dependent Variable: CM
Sample: 2012- 2021

Table 2. Regression model 2

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1443.795	1468.871	-0.982928	0.3584
ICM	3281.268	981.3686	3.343563	0.0124
RUP	-2870.305	465.5488	-6.165422	0.0005
R-squared	0.986890	Mean dependent var		789.1800
Adjusted R-squared	0.983144	S.D. dependent var		618.7171
S.E. of regression	80.32805	Akaike info criterion		11.85344
Sum squared resid.	45168.16	Schwarz criterion		11.94422
Log likelihood	-56.26720	F-statistic		263.4699
Durbin-Watson stat	1.698267	Prob(F-statistic)		0.000000
Breusch-Godfrey Serial Correlation LM Test:				
F-statistic	0.056134	Probability		0.946000
Obs*R-squared	0.219604	Probability		0.896012
Jarque-Bera test	0.437579	Prob. (J-B)		0.803491
Skewness	-0.215290	Kurtosis		2.070059
White Heteroskedasticity Test:				
F-statistic	0.511725	Probability		0.732004
Obs*R-squared	2.904680	Probability		0.573902

Source: Authors' contribution with software EViews.

The regression equation has the form:

$$\mathrm{CM} = 3281.268 \cdot \mathrm{ICM} - 2870.305 \cdot \mathrm{RUP} - 1443.795 + \mathrm{e}$$

From a statistical point of view, the data in the previous Table 2, show us that the probabilities attached to the t-Student test are less than 0.05 and the calculated values confirm that the parameter corresponding to each variable is significantly different from 0. The F-statistic for the proposed model has the value of F = 263.4 and therefore, it is accepted that overall, the studied multiple linear regression model is valid and verifies the existence of at least one parameter that corresponds to an exogenous variable different from zero. The econometric model verifies the test of independence of errors of first order, using the Durbin-Watson test (DW=1.69), and high order from Breusch-Godfrey test (F-statistic = 0.05; Probability = 0.94), also the normality of the distribution of errors, from the Jarque-Bera test (JB=0.43; the associated probability is 80.3%) with Skewness = - 0.21 and Kurtosis = 2.07 and the White test leads to homoscedasticity model (F-statistic = 0.51, attached probability of 73.2%).

Multicollinearity is also found to be absent, because we obtain that $R^2_{CM} = 0{,}98 > R^2_{ICM} = 0{,}81$ and $R^2_{CM} = 0{,}98 > R^2_{RUP} = 0{,}81$, and variance inflation factor is given by *VIF*=5,26 < 10. The proposed multifactorial econometric model explains 98.31% the variation of the dependent indicator studied according to the adjusted coefficient of determination. The increase of 0.01% of the ICM leads to the increase of the dependent variable by 32.81% if the other remains constant. At the same time, the increase of RUP by 0.01% in the conditions where the ICM remains constant, leads to the decrease of CM by 28.7%.

The second econometric model indicates an important link between the endogenous variable and the exogenous variables, given in this case by the case-mix index and the bed utilization rate. Thus, a decrease in the ICM leads to a decrease in the average cost per day of hospitalization, as well as an increase in the bed utilization rate, implies a decrease in the average cost per day of hospitalization at the Craiova Emergency Hospital.

5. Conclusions

From the two validated econometric models, it can be seen that a decrease in the average cost per day of hospitalization can be achieved by increasing the bed utilization rate and increasing the number of cases treated in a calendar year. Also, the decrease in the case-mix index, which can be achieved by increasing the number of resolved cases, leads to a decrease in the average cost of hospitalization per day. The implementation of these results can lead to efficient management at the Emergency County Hospital in Craiova.

The only indicator that does not appear in any econometric model is the average length of hospitalization. Consequently, we cannot draw the conclusion that this performance indicator has or does not have an influence on the average cost per day of hospitalization.

In the future, we propose to do a statistical analysis of other hospitals in Romania, a comparison between them, regarding the performance indicators of the health system, trying to find data for a longer period of time.

REFERENCES

Barliba, I., Nestian, A.S. and Tiţă S.M., 2012. Relevance of Key Performance Indicators in a Hospital Performance Management Model, *Journal of Eastern Europe Research in Business & Economics,* 674169. https://doi.org/10.5171/2012.674169.

Bergeron, B., 2017. *Performance Management in Healthcare. From Key Performance Indicators to Balanced Scorecard*, Taylor & Francis. https://doi.org/10.4324/9781315102214.

Bhuyan, R., Kalita, N. and Goswami, G., 2018. *Health Performance Index and Healthcare Expenditure in Assam: Are There any Structural Change*? In: De, U., Pal, M., Bharati, P. (eds) Issues on Health and Healthcare in India. India Studies in Business and Economics, 229-239, Springer, Singapore. https://doi.org/10.1007/978-981-10-6104-2_12.

Brooks, C., 2008. *Introductory Econometrics for Finance*, 2nd Edition, Cambridge University Press.

Bokhari F., Gai, Y. and Gottret, P., 2007. Government health expenditures and health outcomes, *Health Econ*., 16(3), pp. 257-273. 10.1002/hec.1157.

Burlea-Schiopoiu A. and Ferhati K., 2021. The Managerial Implications of the Key Performance Indicators in Healthcare Sector: A Cluster Analysis, *Healthcare*, *9*(1),19. https://doi.org/10.3390/healthcare9010019.

Chang, A. et all., 2019. Past, present, and future of global health financing: a review of development assistance, government, out-of-pocket, and other private spending on health for 195 countries, 1995–2050 - Global Burden of

Disease, Health Financing Collaborator Network, *Lancet*, 393(10187), pp. 2233-2260, 10.1016/S0140-6736(19)30841-4.

Fuchs, V., 1986. *The Health Economy*, Harvard University Press, London, UK.

Leggat, S. G. et all., 1998. A Review of Organizational Performance Assessment in Health Care, *Health Services Research Management*, 11(1), pp. 3–23. 10.1177/095148489801100102.

Maddala, G. S., 2001. *Introduction to Econometrics*, John Wiley & Sons.

Martin, S. and Smith, P., 2005. Multiple Public Service Performance Indicators: Toward an Integrated Statistical Approach, *Journal of Public Administration Research and Theory*, 15(4), pp. 599 – 613. https://doi.org/10.1093/jopart/mui036.

Parmenter, D., 2010. *Key performance indicators (KPI): developing, implementing, and using winning KPIs*, John Wiley & Sons.

Wooldridge, J., 1999. *Introductory Econometrics: A modern Approach*, South-Western College Pub.

Annex 1.

Years	DMS	RUP	NC	ICM	CM
2012	5.87	0.75	64995	1.1618	294.95
2013	5.91	0.75	67913	1.1587	311.8
2014	6.01	0.73	65717	1.2015	338.94
2015	5.92	0.71	64654	1.1967	357.87
2016	5.92	0.69	61793	1.1779	454.5
2017	5.85	0.6123	54478	1.1958	607.3
2018	5.76	0.6078	55345	1.1907	734
2019	5.92	0.603	54183	1.2826	1042.82
2020	5.98	0.3766	32191	1.3154	1788.95
2021	5.9	.3711	33009	1.3483	1960.67

Source: Craiova Emergency Hospital.