EPH - International Journal of Medical and Health Science

ISSN (Online): 2456-6063
Volume 03 Issue 03 September 2017

DOI: https://doi.org/10.53555/eijmhs.v4i2.33

IMPORTANCE OF CONSIDERING QUALITY INDICATORS IN PRIMARY HEALTHCARE. APPLICATION OF A TWOSTAGE CLUSTER ANALYSIS.

Dr. Ana Luisa Godoy Caballero[1*], Dr. Luis Regino Murillo Zamorano[2]

[*1]*Department of Economics. Faculty of Business, Finance and Tourism. University of Extremadura, 10071, Cáceres, Spain. Ext.: 51423.*

[2]*Department of Economics. Faculty of Business. University of Extremadura, 06006, Badajoz, Spain. Phone: +34 924 289300 Ext.: 89197. Email: lmurillo@unex.es*

[*]***Corresponding Author:-***

Email: anagodoycaballero@gmail.com ; anagodoycaballero@unex.es Phone: +34 927 257000

Abstract:-

The aim of this paper is to evaluate the efficiency and quality of primary healthcare in Extremadura (Spain), assessing at the same time the importance and influence of the quality indicators in the performance of the health units. This analysis considers a series of quality indicators that may affect the efficiency and activity levels of a series of primary care centres. We build different synthetic indices of quantitative output; output adjusted by quality; input, and costs, applying Principal Component Analysis. Using those indices we run several two-stage cluster analyses. In a first analysis, the output of the health system is obtained from a strictly quantitative point of view and compared to the levels of inputs and costs. In a second analysis, we include an output adjusted by quality to perform such a comparison. The health units in which the region is organised can be clustered in four levels of efficiency and activity: efficientactive, efficient-inactive, inefficient-active and inefficient-inactive. The comparison of both analyses highlights the importance of considering qualitative indicators as they substantially influence the efficiency and activity levels of the different primary healthcare centres.

What is known about the topic?

- *Health output has been linked with quantitative aspects, related to the activity of the different centres, evaluating the health service from a strictly quantitative point of view, widely criticised orientation.*

What does this paper add?

- *The incorporation of qualitative indicators into the measurement of the healthcare efficiency and quality affects the results obtained under a strictly quantitative point of view.*

INTRODUCTION

The primary objective of health policy is to maintain and improve the health of its users increasing patients' satisfaction with the whole system. Therefore, it is necessary to efficiently assign the scarce economic resources to be used where they are needed and where better results are obtained. A huge stream of research focuses on measuring the health systems efficiency. However, most of it mainly deals with hospitals, providing less attention to primary healthcare. The consideration of hospitals instead of primary care obeys to the fact that the former are organisations with clear boundaries where patients are admitted and discharged (Amado and Dyson, 2008). Contrarily, primary care bounds are not that explicit, and health is delivered in an open community-based system. These primary healthcare characteristics make it difficult to appropriately establish a measurement of the service delivered, especially when the output of the service needs to be defined. Health output has traditionally been linked to quantitative aspects, related to the centres' activity (e.g., number of consultations) (Chilingerian and Sherman, 1997; Goñi, 1999), producing models evaluating the service from a strictly quantitative point of view. This orientation is mainly motivated by the lack of information about other aspects of the health services and it has been widely criticised (Puig-Junoy and Ortún, 2004; Amado and Dyson, 2009). Firstly, because the number of consultations may be affected by elements that providers cannot control (e.g., sociodemographic characteristics of the population); secondly, because GPs can choose the number of patients they want to see on a day, and thirdly, because the impact of the visits in patients' health mainly depends on how effective these consultations are. Despite the limitations of strictly quantitative approaches, research considering qualitative factors influencing the measurement of the health output are very limited (Salinas-Jiménez and Smith, 1996; MurilloZamorano and Petraglia, 2011; Cordero et al., 2014). Leaving the models strictly oriented to the health system activity, we contribute to the literature by incorporating quality into the measurement of the healthcare output. The objective of this paper is to evaluate the level of efficiency of primary healthcare in Extremadura considering several quality indicators that may affect the efficiency with which the system is offering its services when measured under a strictly quantitative perspective. To that end, we calculate a series of synthetic indices of output, input and costs, using information from a rich dataset for primary health system in Extremadura (APEX). We use these indices to examine the efficiency of the centres participating in the study, applying a two-stage cluster analysis (Yang et al., 2009; López-Sánchez and Santos-Vijande, 2015), which allows performing an efficiency analysis with a special incidence in the inclusion of quality indicators. To the best of our knowledge this is the first time that a two-stage cluster approach is used to analyse primary healthcare efficiency. By doing so, we are able to cluster the individual units not only based on their efficiency scores, but also on their level of activity performance.

Methods

We use cluster analysis to determine the efficiency and activity levels of the health units of the region of Extremadura, retrieving data from APEX08 (Murillo-Zamorano et al., 2011), an information system for the study of primary care in the region of Extremadura. Extremadura is one of the largest and less populated regions in Spain, being therefore, sparsely populated. Because of that, its primary healthcare is structured around two levels of aggregation: Health Areas and Health Zones. The system consists of eight Health Areas, each of them divided into different Health Zones (organised around a primary care centre as the main provider). We use data from 104 Health Zones to build, using Principal Component Analysis, three main synthetic indices associated with the health output or output adjusted by quality, with the health inputs and with the costs of the system. The output adjusted by quality is obtained from the use of two indices, one related to activity and one related to quality. The activity output index is built using as variables the number of consultations (in per capita terms) by GP, paediatrician, nurse, and emergency unit. The qualitative output is obtained from two intermediate indices of quality. The first one contains information about the daily caseload by GP, paediatrician and nurse. The daily caseloads are assumed to be negatively associated with quality, because the staff with fewer consultations would be able to devote more time to each patient, offering a high quality service. Therefore, we accordingly adjust such index, so that it is directly associated with quality. The second index of qualitative output is built using the experience of the health staff (in days worked), the number (by patient) of diagnostics tests, and the inverse number of referrals (by patient) from primary to secondary care. We assume that the number of diagnostics test is positively associated with quality, as those tests help medical staff to better understand the nature of the problems. For the referrals, we consider that it is negatively associated with quality. Consequently, we express this indicator by its inverse value, so that the three variables have the same positive orientation towards quality. The index of input is obtained using (in per capita terms) the health staff in each centre (GPs, paediatricians, nurses), the non-health staff (admin staff, social workers), the number of prescriptions, and the centre area. The index of cost is built from the costs of the health staff, non-health staff, prescriptions, and area (all of them in per capita terms). The descriptive statistics of all variables and the indices are presented in Table 1. Once we obtain the indices we proceed to the application of cluster analysis, an exploratory data analysis tool that classifies observations in relatively homogenous groups, called clusters (Jobson, 1992). With that, similar individuals will belong to the same groups while different observations will belong to different clusters. Three main purposes exist when applying data clustering (Jain, 2010): (1) 'to gain insight into data, generate hypotheses, detect anomalies and identify salient features', (2) to identify similarities between individuals, and (3) to arrange and summarise the data. In this research and following previous literature (Yang et al., 2009; López-Sánchez and SantosVijande, 2015) we take a hybrid approach, known as two-stage cluster analysis, firstly applying the hierarchical method of Ward (Ward, 1963) and then the non-hierarchical k-means clustering method (Ball and Hall, 1965; MacQueen, 1967; Lloyd, 1982). The hierarchical method of Ward forms 'hierarchical groups of mutually exclusive subsets on the basis of their similarity with respect to specified characteristics' (Ward, 1963). It is designed to obtain the groups in order to minimise the within-cluster variance (Punj and Stewart, 1983). In the k-means procedure, individuals are reassigned by moving them to the group whose centroid is closest to that particular individual being reassigned (Punj and Stewart, 1983). This method minimises the mean squared distance from each of the observations to its closest centre. The kmeans methodology produces exceptional results if given a reasonable starting solution (Milligan and Cooper, 1987; Stock and Zacharias, 2011). Finally, and after both methods are applied, we use discriminant analysis to inform the number of cluster that should be considered (Greenly et al., 2005). Discriminant analysis allows for the identification of dissimilarities between two or more clusters in relation to several variables at the same time (Klecka, 1980).

Results

We run two cluster analyses. The first one is strictly quantitative, so we use the indices of quantitative output, inputs and costs. The second one incorporates the quality indicators, therefore, we use the output adjusted by quality or total output, together with inputs and costs. For the first stage of the cluster analysis, we consider a range of solutions between three and four groups. After looking at the composition of these clusters we decide to perform the nonhierarchical k-means methods using the same range of solutions. The existence of two clusters does not seem to be very sensible as the units contained in each of them may be very heterogeneous. Nevertheless, we decide to also include that option in the second stage of our analysis in case the k-means method provides with additional information. After analysing the solutions obtained, we consider that the most appropriate one is the four-cluster option, indicating the existence of four levels of efficiency and activity in primary healthcare in Extremadura, described later on. We determine the stability of these results by means of a discriminant analysis (Greenly et al., 2005). The discriminant functions provide a significant value of the Wilks' lambda for both analyses performed (Wilks' lambda = 0.354; $p = 0.000$ and Wilks' lambda = 0.494; $p = 0.000$, respectively); indicating that the four cluster obtained are statistically significantly different. Additionally, the discriminatory models correctly classify 99.00% and 99.04% of the cases, respectively. Table 2 shows the mean values of the indices of quantitative output, total output, input, and costs in each of the clusters built, for both analyses. Considering these values we characterise the groups based on their activity and efficiency levels as follows: efficient-active, efficient-inactive, inefficient-active, and inefficient-inactive. In the efficient-active clusters is the Health Zones present an index of output greater than the values of inputs and costs. Furthermore, they are active because their levels of outputs are high, compared to the Health Zones classified in the inactive clusters. The efficient-inactive cluster is that one that can be considered as efficient because the levels of output are greater than their inputs and costs. However, this group receives the characterisation of inactive because of its reduced outputs. The group categorised as inefficient-active, present high levels of activity and activity adjusted by quality, but their inputs and costs are greater than its output being, therefore, inefficient. The final cluster, despite using more resources than other groups, its level of output is very reduced and smaller to the corresponding inputs and costs, being inefficient-active. As seen from Table 2 and focussing first on the strictly quantitative analysis almost half of the Health Zones are inefficient (active and inactive), indicating that, efforts need to be done to reduce the resources used or to increase the activity of the centres so they can be considered efficient Health Zones. In relation to the analysis performed with the output adjusted by quality, although the number of Health Zones classified in the inefficient clusters decreases compared to the previous analysis, there are still a lot of Health Zones in those groups, needing to either reduce the resources they use or increase their activity and quality. Additionally, within the efficient Health Zones, there is a reduction in the number of efficient-active ones and an increase in the number of efficientinactive Zones, as a consequence of the inclusion of quality, indicating that the quality with which they are offering healthcare could not be the most appropriate. Therefore, claiming for an improvement in the quality with which healthcare is delivered. Looking at the Health Zones in which these differences are taking place, our results indicate that almost a quarter (21 Health Zones – 20.19%) are influenced by the incorporation of quality indicators (Table 3). In some of them the inclusion of qualitative indicators have a positive impact improving the total output, compared to the quantitative output. The changes with the maximum impact refer to six Health Zones that are ineffient-inactive under the first analysis and efficient-active under the second one. They present high levels of inputs and costs and low levels of quatitative output (inefficient-inactive). However, they are characterised by having a good quality, and because of that, the total output improves in such a way that they become efficientactive when the output adjusted by quality is considered. There are also Health Zones, which do not perform adequately in terms of quality so that, when they are evaluated with an output that includes those qualitative aspects, their efficiency reduces considerably. The most relevant results refer to nine Health Zones that, despite being efficient in both analyses, the incorporation of quality indicators worsens the activity levels so they become inactive. Actions in relation to these Health Zones should be taken in order to be able to increase the quality with which health services are delivered. A similar situation is represented by four Health Zones for which the incorporation of quality indicators worsens their outputs changing from active to inactive. Apart from that, in this case, the Health Zones are also inefficient. Consequently, actions are also needed to reduce inputs and costs so that an increase in the output together with a reduction in these two latter variables can have a positive impact and they become efficient-active. As seen from these results, the inclusion of qualitative indicators affects the efficiency of the health system, which higlights, the importance of the consideration of quality in primary healthcare. Furthermore, a high proportion of Health Zones are inefficient indicating that efforts need to be done to use the resources properly, to increase the output or to keep producing the similar outputs but with a considerable reduction in inputs and costs.

Discussion

Our study has analysed whether different efficiency levels exists in the provision of primary healthcare services in Extremadura. To that end that we have extracted data from APEX08, an information system specifically elaborated for the study of primary healthcare in the region and that contains detailed information about the Health Areas and Health Zones in which primary healthcare is organised. Using this dataset and applying Principal Component Analysis we have calculated synthetic indices of output, input and costs. For the output, we have not only considered the activity, we have also considered quality variables to adjust the quantitative output and obtain an output adjusted by quality. Such a consideration obeys to the idea that in order to be able to adequately measure health output, we need to take into account both quantitative (activity) and qualitative (quality) indicators. We have used these indices to perform an efficiency analysis, utilising a two-stage cluster analysis, applying the method of Ward (Ward, 1983) and the non-hierarchical k-means methodology (Ball and Hall, 1965; MacQueen, 1967; Lloyd, 1982). After studying the different solutions and using a discriminant analysis (Greenly et al., 2005), we have selected four clusters as the most appropriate solution, defining four different levels of efficiency and quality: efficient-active, efficient-inactive, inefficient-active and inefficient inactive. The cluster methodology applied in this research present the potentiality of allowing grouping the individual units not only based on their efficiency, but also on their activity. These four groups were the most appropriate solution for the two cluster analyses performed. In the first of them we only considered activity indicators, whilst in the second one we incorporated the quality variables into the definition of the output, using, therefore, the output adjusted by quality. In relation to the first analysis, results indicated the need for an increase in efficiency, either with a reduction in the resources used or an improvement in the activity delivered, given that 44.23% of the Health Zones participating in the study were classified in the inefficient clusters. Within the 55.77% Health Zones classified as efficient, there is a need to increase the activity in almost half, classified as inactive. The incorporation of quality indicators considerably affected the results, both positively and negatively. Less Health Zones were classified in the inefficient clusters, something which could be due to an increase in the output resulting by the incorporation of good quality indicators into the activity. However, the incorporation of quality also produced a decrease in the number of efficient-active Health Zones. Based on all the aforementioned, it becomes essential the consideration of quality in the analysis of the different health units in any system, given its influence in the efficiency and activity levels. The development of any piece of research that does not include these distinctive characteristics could lead to misleading results and to an inadequate assessment of the activity of the health units operating in a particular system. Similarly, and in terms of policy implications, these conclusions should be taken into account in the decision making process about where to devote the scarce economic resources as, it does not always occur that those health units with the highest production, in strictly quantitative terms, are the ones using the resources in the most efficient way. Finally, and in terms of future research, it would be very important the consideration of quality as perceived by the patients. Patients are the actual users of the healthcare and towards them it should be organised. Therefore, an analysis of their perception of the service that it is being delivered will facilitate the decision of where to assign the scarce resources.

Conflicts of interest

The authors declare no conflicts of interest

References

[1]. Amado D, Dyson R (2008) on comparing the performance primary care providers. *European Journal of Operational Research* 185 (3), 1469-1482.

[2]. Amado D, Dyson R (2009) Exploring the use of DEA for formative evaluation in primary diabetes care: an application to compare English practices. *Journal of Operational Research* 60 (11), 1469-1482.

[3]. Ball G, Hall D (1965) ISODATA, a novel method of data anlysis and pattern classification. *Technical report NTIS AD 699616*. Stanford Research Institute, Stanford, CA.

[4]. Chilingerian J, Sherman H (1997) DEA and primary care physician's report cards: Deriving preferred practice cones from managed care service concepts and operating strategies. *Annual Operational Research* 73, 36-66

[5]. Cordero JM, Crespo E, Murillo-Zamorano, LR (2014) the effect of quality and sociodemographic variables on efficiency measures in primary health care. *The European Journal of Health Economics* 15 (3), 289-302.

[6]. Goñi S (1999) an analysis of the effectiveness of the Spanish primary health care teams. *Health Policy* 48 (2), 107-117.

[7]. Greenly GE, Hooley GJ, Rudd JM (2005) Market orientation in a multiple stakeholder orientation context: implications for marketing capabilities and assets. *Journal of Business Research* 58 (11), 1483-1494.

[8]. Jain KJ (2010) Data clustering: 50 years beyond K-means. *Pattern Recognistion Letters* 31 (8), 651-666.

[9]. Jain AK, Dubes RC (1988) Algorithms for clustering data. Prentice-Hall, Inc.

[10]. Jobson J (1992) Applied Multivariate Data Analysis. Volume II: Categorical and Multivariate

[11]. Methods*.* New York. *Editorial Springer-Verlag*. I.S.B.N.: 0-387-97804-6.

[12]. Klecka WR (1980) Discriminant analysis. *Sage*, 19.

[13]. Lloyd S (1982) Least squares quantization in PCM. *IEEE Transactions on Information Theory* 28 (2), 129–137.

[14]. López-Sánchez JA, Santos-Vijande ML (2015) how value creation and relationship quality coalignment affects a firms's performance: An empirical analysis. *Journal of Marketing Channels* 22 (3), 214-230.

[15]. MacQueen J (1967) some methods for classification and analysis of multivariate observations. In: *Fifth Berkeley Symposium on Mathematics. Statistics and Probability*. University of California Press, 281–297.

[16]. Milligan GW, Cooper MC (1987) Methodology review: Clustering methods. *Applied Psychoclogical Measurement* 11 (4), 329-354.

[17]. Murillo-Zamorano LR, Petraglia C (2011) Technical efficiency in Primary Health Care: Does quality matter? *The European Journal of Health Economics* 12 (2), 115-125.

[18]. Murillo-Zamorano LR, Vega J, Corrales N, Fernández Y, de Miguel FJ, Morillo J, Nogales L, Sánchez M (2011) APEX08. Sistema de Información de la Atención Primaria en la Comunidad Autónoma de Extremadura, 2008. Consejería de Sanidad y Consumo, Junta de Extremadura. Murillo and Vega eds. ISBN: 978-84-6147-541-4.

[19]. Puig-Junoy J, Ortún V (2004) Cost efficiency in primary care contracting: A stochastic frontier cost function approach. *Health Economics* 13 (12), 1149-1165.

[20]. Punj G, Stewart DW (1983) Cluster analysis in marketing research: Review and suggestions for application. *Journal of Marketing Research* 134-148.

[21]. Salinas-Jiménez J, Smith, PC (1996) Data Envelopment Analysis applied to quality in primary health care. *Annals of Operational Research* 67 (1), 141-161.

[22]. Stock RM, Zacharias NA (2011) Patterns and performance outcomes of innovation orientation. Journal of Academic Marketing Science 39 (6), 870-888.

[23]. Ward JHJr (1963) Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association* 58 (301) pp. 236-244.

[24]. Yang F, Sun T, Zhang C (2009) an efficient hybrid data clustering method based on K-harmonic means and particle swarm optimization. *Experts Systems with Applications* 36 (6), 9847-9852.

Table 1. Descriptive statistics

Variable	Role	Mean	Standard deviation (SD)	Max	Min
Consultations by GP	Quantitative output	10.28	3.92	20.82	2.74
Consultations by pediatrician	Quantitative output	6.38	3.13	15.57	1.05
Consultations by nurse	Quantitative output	6.98	3.33	19.46	1.71
Consultations by emergency unit	Quantitative output	1.51	0.69	3.08	0.18
Daily caseload GP	Qualitative output I	40.19	10.12	64.79	17.16
Daily caseload pediatrics	Qualitative output I	20.51	8.73	53.26	4.50
Daily caseload nurse	Qualitative output I	28.12	11.27	87.61	10.96
Experience	Qualitative output II	5.825	534	6.470	3.815
Diagnostic tests	Qualitative output II	0.52	0.22	1.17	0.07
Referrals	Qualitative output II	3.09	0.69	6.67	1.61
Health staff	Input	0.0031	0.0015	0.0097	0.0013
Non-health staff	Input	0.0013	0.0008	0.0046	0.0003
Prescriptions	Input	23.62	4.86	35.54	11.59
Area	Input	0.14	0.09	0.47	0.01
Cost of health staff	Cost	210.87	95.74	594.34	89.40
Cost of non-health staff	Cost	31.83	18.27	85.97	7.67
Cost of prescriptions	Cost	288.32	112.84	813.73	0.15
Cost of the area	Cost	0.84	0.97	7.05	0.0003
Synthetic indices					
Index of activity or quantitative output		49.95	31.10	98.89	4.37
Index of quality I or qualitative output I		50.67	26.66	95.96	3.68
Index of quality II or qualitative output II		50.33	22.55	93.79	7.22
Index of total quality or total quality output		49.35	27.07	93.46	8.88
Index of total output or output adjusted by quality		51.22	26.90	94.57	7.45
Index of inputs		43.82	37.25	100.00	1.46
Index of costs		43.63	36.76	100.00	1.49

Source: Prepared by authors

Table 2. Mean values of the indices in each of the clusters

ANALYSIS QUANTITATIVE OUTPUT-INPUT-COSTS				
Synthetic index	**Cluster 1 (N = 28)**	**Cluster 2 (N = 30)**	**Cluster 3 (N = 32)**	**Cluster 4 (N = 14)**
Quantitative output, mean (SD)	74.39 (12.01)	15.42 (8.65)	71.49 (22.84)	26.27 (14.41)
Input, mean (SD)	26.02 (18.30)	5.64 (4.56)	89.72 (11.63)	47.24 (23.33)
Cost, mean (SD)	22.45 (17.84)	9.32 (7.06)	89.52 (12.53)	45.08 (23.80)
Characterisation	**Efficient-active**	**Efficient-inactive**	**Inefficient-active**	**Inefficient-inactive**
ANALYSIS TOTAL OUTPUT-INPUT-COSTS				
Synthetic index	**Cluster 1 (N = 25)**	**Cluster 2 (N = 38)**	**Cluster 3 (N = 28)**	**Cluster 4 (N = 13)**
Total output, mean (SD)	69.92 (11.00)	30.29 (16.07)	74.89 (16.81)	25.04 (10.18)
Input , mean (SD)	39.06 (23.00)	7.99 (9.10)	91.86 (10.81)	58.10 (26.25)
Cost, mean (SD)	30.84 (18.71)	9.55 (6.64)	93.66 (6.47)	61.31 (20.78)
Characterisation	**Efficient-active**	**Efficient-inactive**	**Inefficient-active**	**Inefficient-inactive**

Source: Prepared by authors

Table 3. Health Zones that change in their efficiency and quality level

Health Area	Health Zone	Analysis quatitative outpuy-input-cost	Analysis total output-input-cost
Changes from efficient-active to efficient-inactive			
Mérida	Aceuchal	Efficient-active	Efficient-inactive
	Calamonte	Efficient-active	Efficient-inactive
	Guareña	Efficient-active	Efficient-inactive
	Villafranca de los Barros	Efficient-active	Efficient-inactive
Llerena-Zafra	Fuente del Maestre	Efficient-active	Efficient-inactive
	Santos de Maimona	Efficient-active	Efficient-inactive
	Zafra II	Efficient-active	Efficient-inactive
Plasencia	Jaraíz de la Vera	Efficient-active	Efficient-inactive
	Plasencia-Norte/La Data	Efficient-active	Efficient-inactive
Changes from efficient-active to inefficient-inactive			
Navalmoral de la Mata	Villanueva de la Vera	Efficient-active	Inefficient-inactive
Changes from efficient-inactive to efficient-active			
Mérida	Mérida - Norte	Efficient-inactive	Efficient-active
Changes from inefficient-active to inefficient-inactive			
Badajoz	Roca de la Sierra	Inefficient-active	Inefficient-inactive
Plasencia	Ahigal	Inefficient-active	Inefficient-inactive
	Casas del Castañar	Inefficient-active	Inefficient-inactive
	Pinofranqueado	Inefficient-active	Inefficient-inactive
Changes from inefficient-inactive to efficient-active			
Badajoz	Alburquerque	Inefficient-inactive	Efficient-active
	San Vicente de Alcántara	Inefficient-inactive	Efficient-active
	Villanueva del Fresno	Inefficient-inactive	Efficient-active
Mérida	Hornachos	Inefficient-inactive	Efficient-active
Don Benito	Talarrubias	Inefficient-inactive	Efficient-active
Plasencia	Hervás	Inefficient-inactive	Efficient-active

Source: Prepared by authors