UNIVERSIDAD DE EXTREMADURA

TESIS DOCTORAL

CUATRO ENSAYOS SOBRE CALIDAD EN LA PRESTACIÓN DE LOS SERVICIOS SANITARIOS

ANA LUISA GODOY CABALLERO

DEPARTAMENTO DE ECONOMÍA

2016

UNIVERSIDAD DE EXTREMADURA

TESIS DOCTORAL

CUATRO ENSAYOS SOBRE CALIDAD EN LA PRESTACIÓN DE LOS SERVICIOS SANITARIOS

ANA LUISA GODOY CABALLERO

DEPARTAMENTO DE ECONOMÍA

CONFORMIDAD DEL DIRECTOR DE LA TESIS

Fdo.: Dr. Luis Regino Murillo Zamorano
Titular de Universidad
Departamento de Economía
Facultad de Ciencias Económicas y Empresariales
Universidad de Extremadura

2016

El presente trabajo para obtener el grado de Doctor en Economía con Mención Internacional ha sido realizado por Dña. Ana Luisa Godoy Caballero en el Departamento de Economía de la Universidad de Extremadura, bajo la dirección del Dr. Luis Regino Murillo Zamorano.

Agradecimientos

- Consejería de Economía, Comercio e Innovación de la Junta de Extremadura y al Fondo Social Europeo, por la concesión de una Beca predoctoral FPI (DOE nº 130 de 08/07/2010), y que sirvió de financiación durante los primeros años de desarrollo de la presente Tesis Doctoral.

Además, me gustaría expresar mi agradecimiento a todas aquellas personas que han sido partícipes de esta trayectoria que ahora termina y cuyo resultado queda plasmado en el desarrollo de esta Tesis Doctoral.

En primer lugar, a mi director de Tesis, Dr. D. Luis R. Murillo Zamorano, sin el cual nada de lo que aquí se presenta hubiese visto la luz. Gracias no solamente por los consejos en el terreno profesional sino también en el personal y en el día a día; por animarme a salir de España y hacerme ver la oportunidad que suponía el trabajar y estudiar fuera; por ayudarme a ver más allá del nubarrón que suponían los momentos de agobio, y sobre todo el apoyo incondicional recibido durante estos años.

A los profesores Pedro Mora, Jesús Punzón y Juan Luis Corral, porque con ellos me inicié en el mundo de la investigación y sé que perciben el fin de este proyecto con la misma felicidad que lo percibo yo.

To Professor Andrew Street for giving me the chance of doing a research stay in the Centre for Health Economics at the University of York. Also, thanks to Dr. Nils Gutacker and Dr. Chris Bojke, for all their knowledge and support during the research stay.

A todos mis compañeros de despacho durante los primeros años de esta andadura, en los que todos nos iniciábamos en el mundo de la investigación y compartíamos las mismas preocupaciones, lo que quizá contribuyese al ambiente de trabajo que creamos.

To all my colleagues (and friends) from the University of Lincoln, because when you are abroad a simple issue can be seen as a big problem, but if you have such a fantastic group of people around, it is easier to deal with them; because you were the family I did not have with me and the sun I so much missed.

A todos mis amigos, con independencia de donde estéis, el tiempo que hace que no os vea o hable con vosotros, el idioma en el que nos comunicamos,… Todos en algún momento habéis sido partícipes de este proyecto, y simplemente por estar ahí, gracias.

A Pablo, porque aunque has llegado en los últimos años de este camino, has sido un gran apoyo, y sin el que muchas veces no hubiese podido seguir hacia adelante. Gracias por quererme tanto y estar ahí día a día, te quiero.

A toda mi familia, especialmente a mis padres porque desde pequeña han confiado en mí, porque en todo momento he encontrado apoyo en ellos, y porque sin ellos no estaría donde estoy actualmente. A mi hermana Pilar, porque hace las veces de amiga, de hermana y ¡hasta de madre!, porque nada de lo que hago sería lo mismo si ella no estuviese siempre a mi lado, y porque tú y yo sabemos de sobra lo que nos queremos. Gracias.

“Continuous improvement is better than delayed perfection”

(Mark Twain)

Index

Declaration

An extended version of Chapter 1 entitled "Análisis Económico de la Asistencia Sanitaria en Atención Primaria. El caso de la Comunidad Autónoma de Extremadura" was published in 2012 as an international book with ISBN: 978-3-8473-6526-6. The research presented in this Thesis has been submitted to ***The Service Industries Journal***.

An earlier excerpt of the research presented in Chapter 2 has been published in ***Family Practice***, 33 (2) 179-185 in 2016 with title "Comparing importance and performance from a patient perspective in English general practice: a cross-sectional survey". In its final format, Chapter 2 is currently under revision in ***Health Services Research***.

Chapter 3 has been submitted to ***Health Economics*****,** under the title "Explaining patients' satisfaction with primary health care: a multilevel analysis".

Chapter 4, to be submitted to ***The European Journal of Health Economics***, was presented in a workshop at the Centre for Health Economics (CHE), University of York (UK) in September 2012, in the 20th Public Economics Meeting, held in Seville in January 2013, and in the 33rd Health Economics Association Conference, held in Santander in June 2013.

Abstract

This Doctoral Thesis analyses the quality with which health care services are being provided. The main purpose of this Thesis is to extend the literature of the economic analysis of quality in the health care field. From a methodological point of view we contribute to that aspect using the latest approaches of multilevel and cluster analyses, and supporting them with the use of empirical data extracted from strong databases, developed *ad hoc* for this research.

We have developed four studies at the regional level of the Spanish Autonomous Community of Extremadura and at the international level using information about the English National Health Service (NHS), utilising not only objective indicators that describe quality but also considering patients' opinions about the health care they receive, based on their satisfaction or on how they benefit from certain medical interventions. Furthermore, we have study the possible source of variation in users' levels of satisfaction.

Our results highlight the need to take into account not only the quality with which health care services are being delivered, but also the importance of considering quality as perceived by the patient in the identification of the sources of improvement in the provision of health care. Considering all that, we will be able to offer an excellent health care and with high quality standards, perceived as such by users of the system, and that, in the last instance, will be reflected in patients' levels of satisfaction.

KEY WORDS: health care, objective quality, subjective quality, patient's satisfaction.

UNESCO CODE: 5312.07, 5307.14, 5302.01

Resumen

La presente Tesis Doctoral desarrolla un análisis exhaustivo acerca de la calidad en la provisión de los servicios sanitarios. El objetivo principal es el de ampliar y extender de forma relevante la literatura existente sobre el análisis económico de la calidad en el ámbito de la atención sanitaria. Desde el punto de vista metodológico, contribuimos a este aspecto mediante la aplicación de los últimos desarrollos de metodología multinivel y análisis clúster, respaldando todo ello con el uso de datos empíricos extraídos de potentes bases de datos desarrolladas *ad hoc* para esta investigación.

Hemos elaborado cuatro estudios a nivel regional en la Comunidad Autónoma de Extremadura y a nivel internacional mediante la utilización de información referida al Servicio Nacional de Salud de Inglaterra, considerando no sólo indicadores objetivos de la calidad, sino también opiniones de los pacientes evaluando la actuación de las diferentes unidades sanitarias, en función su satisfacción o de cómo se benefician de determinados procedimientos médicos. Además, se han estudiado las posibles fuentes de la variación en los niveles de satisfacción de los usuarios del sistema.

Nuestros resultados ponen de manifiesto la necesidad no sólo de tener en cuenta la calidad con la que se prestan los servicios sanitarios, sino también la importancia de considerar la calidad desde el punto de vista del paciente a la hora de identificar las posibles fuentes de mejora en la provisión de los servicios sanitarios. Teniendo en cuenta todo esto, seremos capaces de ofrecer una atención sanitaria excelente y con elevados estándares de calidad, vista como tal por los usuarios del sistema y, que en última instancia, quedará reflejada en los niveles de satisfacción del paciente.

PALABRAS CLAVE: atención sanitaria, calidad objetiva, calidad subjetiva, satisfacción.

CÓDIGO UNESCO: 5312.07, 5307.14, 5302.01

INTRODUCTION

This Doctoral Thesis is organised in the form of four essays about quality in health care. The main purpose of this Thesis is to extent the literature of the economic analysis of quality in the health care field. From a methodological point of view we contribute to that aspect using the latest approaches of multilevel and cluster analyses, and supporting them with the use of empirical data extracted from strong databases, developed *ad hoc* for this research.

From the economic point of view, the ability to identify where resources can be saved or where quality can be improved allows us to differentiate health units that adequately manage their resources from those ones that do not achieve a certain standard and that, therefore, do not administer their resources in the best possible way. A great deal of research in Health Economics has focused on the analysis of the efficiency of the health care. However, it is not enough with being efficient, it is also essential to orientate the health system to the real users, patients, given that the ultimate objective of the health care is patients' satifaction, for which a high quality service needs to be offered, a quality that needs to be perceived by the user.

As Adil et al. (2013) establishes, "quality does not improve unless it is measured"; therefore, if we want to have a health system that it is able to provide a good quality health care, we need to measure its current status so that we can identify where actions are most needed, to improve the quality of the system, and with that, patients' satisfaction. Considering all this, we have developed four pieces of research, presented in four separated chapters, so that, as previously mentioned, we can study in detail quality in health care.

The first chapter of the Thesis presents an analysis of the current state of the primary health care service in the Spanish Autonomous Community of Extremadura. An extended version of this chapter entitled "Análisis Económico de la Asistencia Sanitaria en Atención Primaria. El caso de la Comunidad Autónoma de Extremadura" was published in 2012 as an international book with ISBN: 978-3-8473-6526-6". The research presented in this Thesis has been submitted to *The Service Industries Journal*.

The aim of this first piece of research is to evaluate the efficiency of the primary health care services in the Spanish Autonomous Community of Extremadura. This analysis, and differently to previous efficiency literature, has been performed with the consideration of a series of quality indicators that may be affecting the efficiency and activity levels, obtained by means of a two-stage cluster analysis. To the best of our knowledge, previous literature that has applied the cluster methodology in the analysis of health care has only considered one stage, incurring with that, in the possibility of introducing biases in the analyses.

Traditionally, health output has mainly been linked with quantitative aspects, related to the activity of the different centres, such as number of consultations or visits (Chilingerian and Sherman, 1996; Goñi, 1999), producing models that evaluate the health service from a strictly quantitative point of view. In our case, abandoning these traditional models strictly orientated to the activity of the health system, in this first chapter of the Thesis, we incorporate qualitative aspects into the measurement of output of the health care, so that we can analyse how these quality indicators influence the levels of efficiency with which health services are being delivered

We build a series of synthetic indices of quantitative output, output adjusted by quality, input and costs applying Principal Component Analysis (PCA). Using those indices we run a two-stage cluster analysis. The application of this methodology obtains more refined results than a single stage approach.

We perform two cluster analyses with the indices of output, input and costs. In the first of them, the output of the health system is obtained from a strictly quantitative point of view and compared to the levels of inputs and costs. For the second analysis the output included in it is an output that contains quality indicators; i.e. it is an output adjusted by quality. In both analyses, we obtain that the different health units in which the region is organised can be clustered in four levels of efficiency and activity, categorised as: efficient-active, efficient-inactive, inefficient-active and inefficient-inactive.

The comparison of both analyses allows us to recognise the importance of the consideration of the qualitative indicators as they considerably influence the levels of efficiency and activity of the different health units.

Additionally to the importance of considering indicators of quality in health care we need to take into consideration the user's point of view, because as indicated by Donabedian (1984), if we do not account for patients' opinions about the health service that they are receiving we cannot completely describe quality in health care. Furthermore, if the quality with which health care is provided is perceived as high quality by patients, that will be reflected in their levels of satisfaction with the service received.

Considering this, Chapter 2 of the Thesis analyses patients' satisfaction with the health care. An earlier excerpt of this research has been published in *Family Practice*, 33 (2) 179-185 in 2016 with title "Comparing importance and performance from a patient perspective in English general practice: a cross-sectional survey". In its final format, Chapter 2 is currently under revision in *Health Services Research*.

The work presented in Chapter 2 develops a new scale that adequately measures patients' satisfaction with the health care from the user's perspective, the w-HEALTHQUAL, which includes patients' preferences with certain aspects of the health system together with the importance that users assign to those aspects. To the best of our knowledge, there is not such a weighted scale in the literature that allows for the analysis of patients' satisfaction with health care.

Using the w-HEALTHQUAL, Chapter 2 also develops an application of it, with which we contribute to the literature about the analysis of satisfaction with health care. Specifically, we analyse how patients' level of satisfaction with the health care varies within the Spanish Autonomous Community of Extremadura. We have not found any previous research that aims to identify differences in satisfaction within settings, in spite of the fact that users' preferences with certain services have previously been measured across countries and systems (Schäfer et al., 2015).

We identify three groups of patients that differ between them in patients' level of satisfaction with a series of aspects of the health system: satisfied patients, dissatisfied patients and patient that are not satisfied or dissatisfied with the health care service received. This classification of patients also allows for the categorisation of each the centres in the region based on the satisfaction levels of the patients they treat.

The categorisation of the different health units based on the levels of satisfaction of the patients they treat suggests that a different distribution of health care may be needed so that a more effective health care can be delivered, that ultimately increases patients' satisfaction with primary care.

The identification of different levels of patients' satisfaction distributed across the region of Extremadura lead to the need of recognising what may be driving those levels of satisfaction that are indicating the reconsideration of the organisation of the health system, i.e. which aspects are influencing the different levels of satisfaction? Are they driven by patients' characteristics? Are they explained by particular features that characterise the different centres?

In view to further explore those questions, we have developed another piece of research presented in Chapter 3 and that has been submitted to *Health Economics*, under the title "Explaining patients' satisfaction with primary health care: a multilevel analysis".

The aim of the research developed in Chapter 3 is to explore what patients' characteristics and characteristics at the level of the primary care centres most influence satisfaction with primary health care, obtained as a combination of experiences and importance. Using data from patients attending primary care centres in Extremadura and information about the characteristic of each of the centres, we apply a two-level multilevel analysis, following an additive approach. This methodology lead us to identify what variables are influencing patients' general satisfaction and patients' satisfaction with a series of elements that build a primary care system, such as facilities, health staff, non-health staff and efficiency.

The aforementioned supposes two contributions to the existing literature. First of all, we do not solely consider general satisfaction with primary health care, but also with a series of particular aspects of it, such as the facilities, the health staff, the non-health staff and the efficiency of the centre. Secondly, our measure of satisfaction is obtained as a combination of experiences with a series of elements of the primary care and the importance patients attain to each them, gathering more information than a single question asking patients to report their general satisfaction.

A third contribution of this chapter consists of an extension of the multilevel literature. In the research presented in this chapter, we apply a multilevel model in primary care to study the determinants of variations in patients' levels of satisfaction, obtained as a combination of experiences and importance. In spite of the existence of research applying multilevel analyses in health care, a lot of this research has been centred on the analysis of hospitals rather than primary health care centres. Additionally, the scarce literature applying multilevel models in primary care has measured satisfaction using the responses given by patients to an overall satisfaction question, rather than asking about their experiences and importance judgements of a series of elements of the service.

Our study reveals that most of the variation in patients' satisfaction is attributable to characteristics related to the patients attending primary care, being age, education and waiting times the most influencing aspects. Furthermore, the are also aspects of the health care at the level of the centre that influence how satisfied patients are with the service received, being the daily caseload in paediatrics or the staff working in the centre some of them.

Chapter 4 of the Thesis, to be submitted to *The European Journal of Health Economics,* exhibits an international example of quality in health care. It was presented in a workshop at the Centre for Health Economics (CHE), University of York (UK) in September 2012, in the 20th Public Economics Meeting, held in Seville (Spain) in January 2013, and in the 33rd Health Economics Association Conference, held in Santander (Spain) in June 2013.

This chapter develops a novel multilevel approach that allows us to estimate the effect of the particular centre where health care is being delivered and that can be used as an indicator of hospital performance.

The aim of this last piece of research is to analyse the quality in the provision of health services in the English National Health Service (NHS) according to different patient-reported outcome measures (PROMs). The English Department of Health defines PROMs as those questionnaires filled by patients to evaluate their own health status before and after certain interventions (Department of Health, 2008).

In doing so we contribute to the literature in three ways: (1) we measure the quality of the different health units based on patients' changes in their health status as a consequence of receiving treatment; (2) we perform a comparative analysis of PROMs for the study of quality of the health services, and (3) we obtain an estimation of hospital performance which is used in the identification of differences between centres.

Using PROMs we estimate a series of multilevel models with fixed effects, which allows for the analysis of variation in the measures at provider level, as the method obtains an estimation of the performance of each of the health units, which can be used as an indicator of hospital performance (Jacobs et al., 2006). By analysing in detail the performance of the different hospitals we identify a series of health units that do not present an average behaviour, and that we categorise as outliers.

The study reveals the importance of paying attention to those health units presenting an unusual performance, as they refer to hospitals treating a substantial proportion of patients. The lack of consideration of those unusual observations could introduce a significant bias which could have important implications in policy terms.

All the above has manifested the implications of the consideration of quality in the analysis of health care. A quality that cannot only be measured using objective indicators, but also considering patients' views about the service they receive, the value they give to each of the aspects of the health care or how they benefit from the treatments or interventions. Considering all that, we will be able to offer an excellent health care, with a

good quality, that will be considered as such by patients and that will be reflected in their levels of satisfaction. This is the main concern of this doctoral thesis.

REFERENCES

Adil, D., Al Ghaswyneh, O. F. M., and Albkour, A. M. (2013). SERVQUAL and SERVPERF: A review of measures in services marketing research. *Global Journal of Management And Business Research*, 13 (6).

Chilingerian, J., and Sherman, H. (1996). Benchmarking physician practice patterns with DEA: a multi-stage approach for cost containment. *Annual Operational Research,* 67 (1), pp. 86-116.

Department of Health (2008). Guidance on the routine collection of Patient Reported Outcome Measures (PROMs), The Stationary Office, London.

Donabedian, A. (1984). La Calidad de la Atención Médica. *La Prensa Mexicana*, México.

Godoy-Caballero, A.L. and Murillo-Zamorano, L.R. (2012). Análisis Económico de la Asistencia Sanitaria en Atención Primaria. El caso de la Comunidad Autónoma de Extremadura. *EAE*. I.S.B.N.: 978-3-8473-6526-6.

Godoy Caballero, A. L., Bojke, C. and Gutacker, N., (2013). Comparing hospital quality performance estimates based on different patient-reported outcome measures. In *20th Encuentro Economía Pública: Estado del bienestar: sostenibilidad y reformas*. Universidad de Sevilla.

Godoy Caballero, A. L., Bojke, C. and Gutacker, N., (2013). Using patient-reported outcome measures (PROMs) in the comparison of hospital quality performance. In *33rd Health Economics Association Conference*. Universidad de Santander.

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

Jacobs, R., Smith, P.C. and Street, A., (2006). Measuring efficiency in health care: analytic techniques and health policy. *Cambridge University Press*, ISBN: 978-0-521-85144-2.

Schäfer, W. L., Boerma, W. G. W., Murante, A. M., Sixma, J. J. M., Schellevis, F. G., and Groenewegen, P. P. (2015). Assessing the potential for improvement in primary care in 34 countries: a cross-sectional survey. *Bulletin of the World Health Organization*, 93 (3), pp. 161-168.

Sirdifield, C., Godoy-Caballero, A., Windle, K., Jackson, C., McKay, S., Schäfer, W. and Siriwardena, A.N. (2016). Comparing importance and performance from a patient perspective in English general practice: a cross-sectional survey. *Family practice*, 33 (2), 179-185.

CHAPTER 1:
THE IMPORTANCE OF QUALITY INDICATORS IN PRIMARY HEALTH CARE

1. INTRODUCTION

The primary objective of health policy is to maintain and improve the health of its users increasing patients' satisfaction with the whole health system. In order to do so it is necessary to efficiently assign the scarce available economic resources so they are used where they are more needed and where better results are obtained.

Since Nunamaker (1983), a huge stream of research has focused on measuring the efficiency of the health systems. However, most of this research has mainly dealt with hospitals, providing less attention to primary health care. As Amado and Dyson (2008) point out, the consideration of this level of the health services instead of primary care obeys to the fact that hospitals are organisations with clear boundaries, where patients are admitted and discharged. On the contrary, the bounds of primary care are not that explicit, and health is delivered in an open community-based system. These particular characteristics of primary health care provision 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 with quantitative aspects, related to the activity of the different centres, such as number of consultations or visits (Chilingerian and Sherman, 1996; Chilingerian and Sherman, 1997; Goñi, 1999; Ozcan, 1998, and Pina and Torres, 1992), producing models that evaluate the health service from a strictly quantitative point of view. This quantitative orientation has, in most of the cases, been motivated by the lack of information related to other aspects of the health services and it has been widely criticised (Puig-Junoy and Ortún, 2004 and Amado and Dyson, 2009). Firstly, because the number of consultations may be affected by elements that cannot by controlled by providers, such as the sociodemographic characteristics of the population; secondly, because the General Practitioners (GPs) can choose the number of patients they want to see on a day, i.e. the number of consultations, and thirdly, because the impact of the visits in the improvement of patients' health depends mainly on how effective these consultations are. Despite these limitations that these strictly quantitative

approaches present, research considering qualitative factors that influence the measurement of the health output has been very limited. Among them and in relation to primary health care, we can mention Salinas-Jiménez and Smith (1996), García et al. (1999), and more recently Murillo-Zamorano and Petraglia (2011) and Cordero et al. (2014).

In this research, and leaving the models strictly orientated to the activity of the health system, we contribute to the literature that incorporates qualitative aspects into the measurement of output of the health care. The objective of the research is to evaluate the level of efficiency of primary health care in the Autonomous Community of Extremadura but with the consideration of a series of quality indicators that may be affecting the efficiency with which the health system is offering its services when measured under a strictly quantitative point of view. To that end, we calculate a series of synthetic indices of output, input and costs, using information from a rich dataset for the primary health system in Extremadura in 2008 (APEX08) (Murillo-Zamorano et al., 2011). For the index of output, we obtain two different indicators, one strictly quantitative and a second one that incorporates quality indicators, being therefore, an output adjusted by quality. By comparing these two, we can identify the extent to which qualitative aspects of the health care are actually influencing the total output of it.

These indices are used in the analysis of the efficiency of each of the centres participating in the study, for which we apply an analysis cluster. Specifically, we adopt a two-stage cluster approach (Wong, 1982; Yang et al., 2009; López-Sánchez and Santos-Vijande, 2015), so that the results of one of the methods are corrected or confirmed by the other one. The application of this methodology will allow us to perform an efficiency analysis with a special incidence in the inclusion of qualitative indicators, by comparing the different outputs of the health system, with its inputs and costs. To the best of our knowledge this is the first time that a two-stage cluster approach is used to analyse the efficiency of primary health care. By doing so, we will be able to cluster the individual units not only based on their efficiency scores, but also on their level of activity performance.

In achieving our research objective, we have structured the chapter as follows. In the second section and after this introduction, we present the methodology used which, as just mentioned, refers to a two-way cluster analysis. Afterwards, we devote a section to the data used in the analysis. These data, which refer to primary health services in the Spanish Autonomous Community of Extremadura has been retrieved from a rich database containing additional information characterising the region, and consists of a series of variables measuring the activity, quality, inputs and costs in every Health Zone in which the primary health care system in Extremadura is organised. These data are used to build four synthetic indices: (1) an index to measure output from a strictly quantitative point of view, (2) an index that considers the output as a combination of activity as well as quality variables, or output adjusted by quality, (3) an index of inputs and (4) and an index of costs. Section four shows the results of the analysis, detailing the levels of efficiency and activity with which we can categorise each Health Zone in the region and highlighting the influence of the inclusion of quality indicators. Finally, we present the discussion and conclusion of the research.

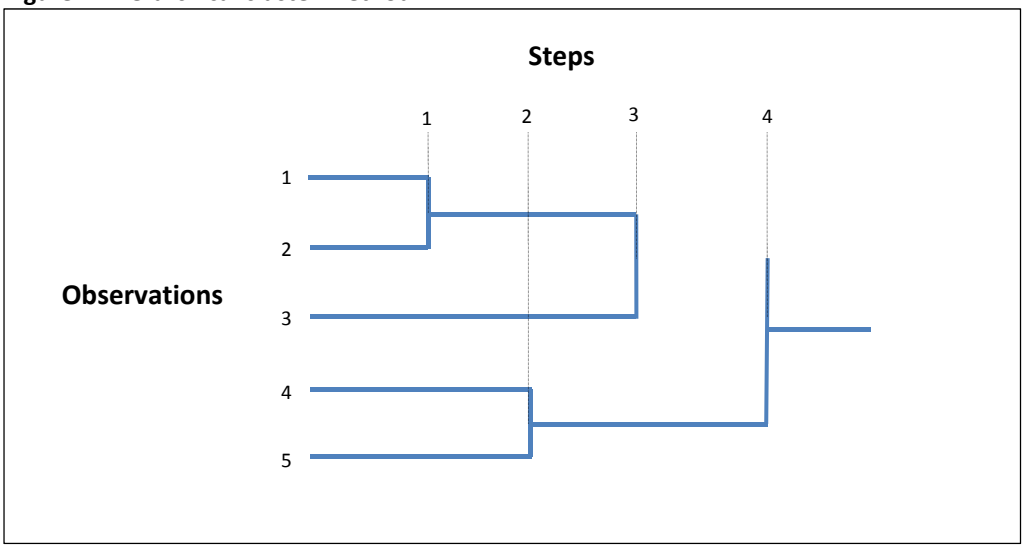
2. METHODS

In this section we describe the methodology used to determine the efficiency and activity levels of the different Health Zones in which the region of Extremadura is organised, by means of a cluster analysis. Cluster analysis is an exploratory data analysis tool that classifies observations in relatively homogenous groups, called clusters (Jobson, 1992). Its main objective is to identify groups of objects that show the highest possible degree of homogeneity within the groups and the highest possible degree of heterogeneity between groups. Following this approach, similar individuals will belong to the same groups while different observations will belong to different clusters. There are three main purposes 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.

Several techniques exist within the cluster methodology. In this piece of research and following previous literature (Wong, 1982; Yang et al., 2009; López-Sánchez and Santos-Vijande, 2015) we take a hybrid approach, also known as two-stage approach. In the first of the stages we apply the hierarchical method of Ward (Ward, 1963) and then we employ the non-hierarchical k-means clustering method (Lloyd, 1982; Ball and Hall, 1965, and MacQueen, 1967).

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). This method can be represented in a graph like the one shown in Figure 1, where, to simplify the explanation, we have only represented five observations and four steps. As we move from one step to the following, similar observations will become part of the same group while those which are different will be part of a different cluster. In each of the steps only one observation changes its belonging to a group, and once a particular individual has been assigned to a group it will do it until the end of the process. New observations could be added to that cluster, but that particular individual will not change to another one once it has been assigned to one group. The lower the number of groups we want, the greater the number of steps we need to take. Following this procedure, and depending on the number of groups we want, we will stop at that particular point of the procedure, as each of the steps can be considered as a solution of the method.

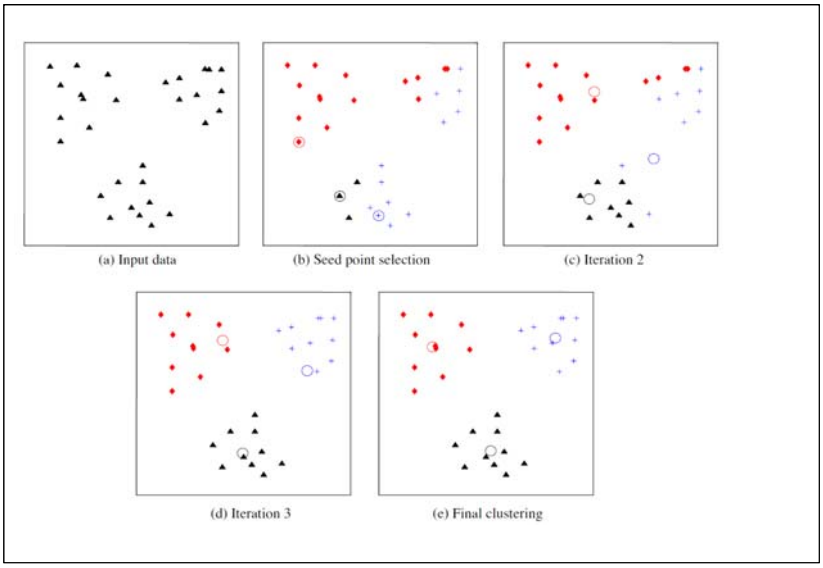
Figure 1. Hierarchical cluster method



Source: Adapted from Jobson (1992)

In the k-means procedure, individuals are reassigned by moving them to the group whose centroid is closest to that particular individual which is being reassigned (Punj and Stewart, 1983). It minimises the mean squared distance from each of the observations to its closest centre. The k-means method produces exceptional results if given a reasonable starting solution (Milligan and Cooper, 1987; Stock and Zacharias, 2011). This procedure, which can be represented as shown in Figure 2, has three main steps (Jain and Dubes, 1988; Jane, 2010): (1) select an initial distribution with a determined number of clusters and repeat the following steps until the group stabilises; (2) produce a new distribution by "assigning each pattern to its closest cluster centre", and (3) obtain new cluster centres.

Figure 2. K-means algorithm[1]



Source: Adapted from Jane (2010)

Finally, and after both methods are applied, we use a discriminant analysis in order 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).

3. DATA

We use data from APEX08[2] (Murillo-Zamorano et al., 2011), an information system for the study of primary care in the Autonomous Community of Extremadura and that contains detailed information for each of the Health Zones and Health Areas in which the region is organised.

Extremadura is one of the largest Autonomous Communities in Spain but also one of the less populated regions, being sparsely populated. Because of these particular characteristics, its primary health care is structured around two territorial levels of aggregation: Health Areas and Health Zones. The system consists of a total of eight Health Areas: Badajoz (I), Mérida (II), Don-Benito-Villanueva de la Serena (III), Llerena-Zafra (IV),

[1] (a) Input data with three clusters; (b) three seed points which are the cluster centres and the initial distribution of individuals in clusters; (c) and (d) intermediate interactions updating the label of the groups and their centres, and (e) final clusters obtained by the k-means procedure.
[2] APEX is the acronym used in Spanish to specify that the data refers to Primary Health Care in Extremadura.

Cáceres (V), Coria (VI), Navalmoral de la Mata (VII) and Plasencia (VIII). Each of them is, at the same time, divided into different Health Zones, each one of them organised around a primary care centre as the main provider. In the year 2008, there were a total of 109 operating Health Zones. The research presented in this chapter is developed for 104 of them for which we have complete information in all the variables used in the analysis. Figure 3 shows the distribution of the Health Zones across the Health Areas and across the regional map[3]. The existence of an adequate number of Health Zones and an appropriate provision of equipment is of predominant importance in order to guarantee that health services are delivered correctly and to reduce the inconveniences derived from having a sparsely populated region.

Figure 3. Map of the distribution of Health Zones and Health Areas in Extremadura

Source: Murillo-Zamorano et al. (2011)

In term of variables, we build three main synthetic indices associated with the health output or output adjusted by quality (INDOUT), with the health inputs (INDINP)

[3] A detailed distribution of the Health Zones across the Health Areas is presented in Appendix 1.

and with the costs of the primary health care system (INDCOSTS). INDOUT is obtained from the use of other two indices, one related to activity (INDACT) and another one related to quality (INDQUAT[4]).

In what comes next, we explain in detail the construction of the different indices[5]. Figure 4 shows the variables used as well as the procedure followed to obtain the three final indices.

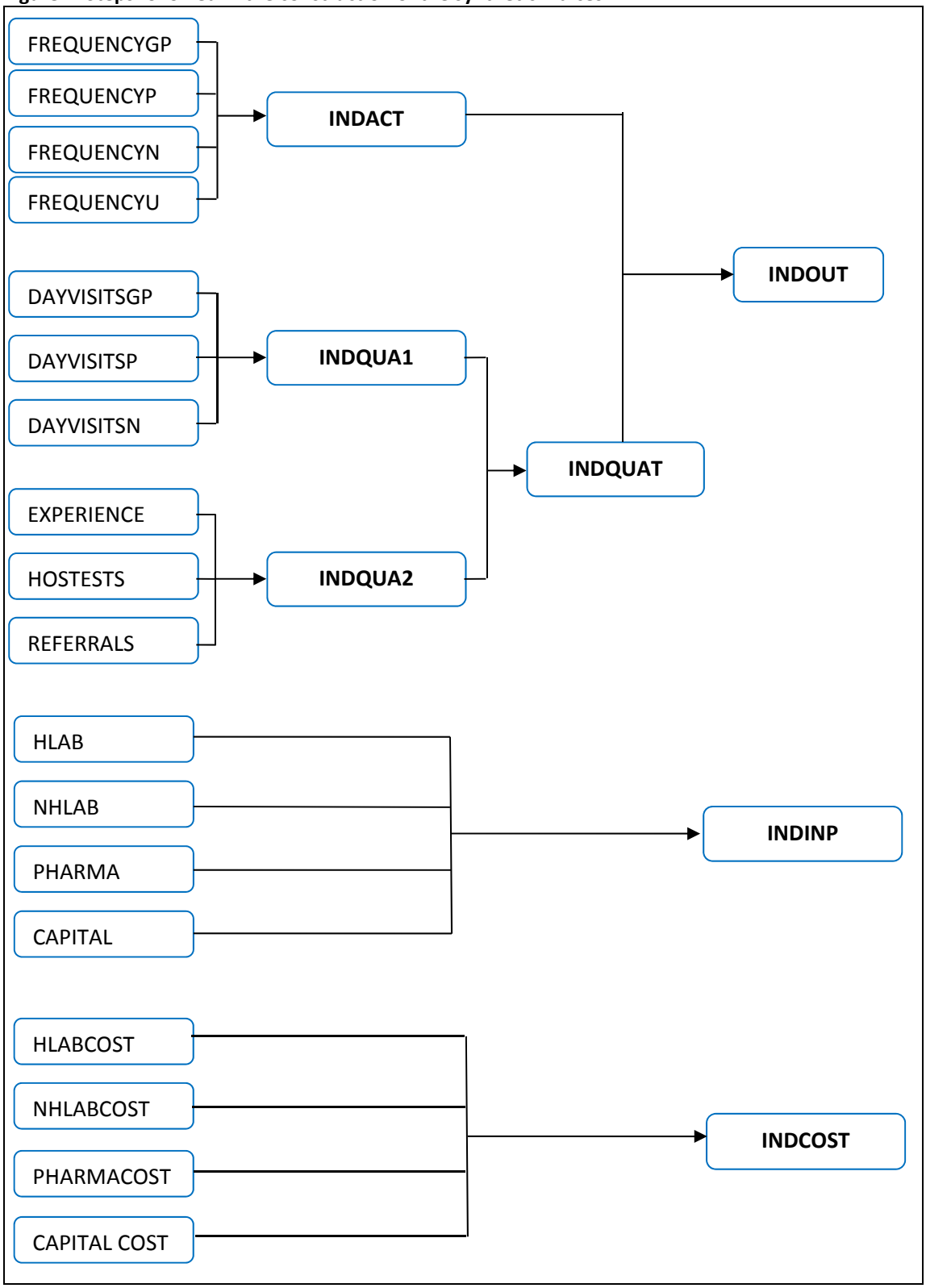
INDACT relates to the quantitative output and it is built from the number of visits or consultations by primary care professionals (in per capita terms), specifically, by the General Practitioners (GPs) – FREQUENCYGP, the paediatricians – FREQUENCYP, the nurses – FREQUENCYN, and the emergency units – FREQUENCYU.

INDQUAT refers to the qualitative output and it is obtained from two intermediate indices of quality, INDQUA1 and INDQUA2. INDQUA1 contains information about the daily caseload by professional, i.e. the daily number of visits or consultations, per GP – DAYVISITGP, paediatrician – DAYVISITP and nurse – DAYVISITN. The daily caseloads are assumed to be negatively associated with quality, because we consider that staff with fewer consultations, would be able to devote more time to each patient, offering with that a high quality service. INDQUA1 has been accordingly adjusted so that it is directly associated with quality of the health care.

[4] For the construction of the quality output (INDQUAT) has been necessary to obtain two intermediate indicators. The procedure followed is explained later on.

[5] The score each of the synthetic indices for each of the Health Zones and its rank is presented in Appendix 2 and 3. The first of them contains the information with the indices that are used in the cluster analyses, whilst in Appendix 3 we show the rest of synthetic indices.

Figure 4. Steps followed in the construction of the synthetic indices



Source: prepared by authors

INDQUA2 is built using the variables EXPERIENCE, HOSTESTS AND REFFERALS. EXPERIENCE is a proxy for the experience of the health staff, indicating the number of days worked since December 1990 to December 2008. HOSTESTS represent the number (by patient) of diagnostics tests, such as blood test, urine analysis, etc. that the primary care centres request to the reference hospital. We assume that this variable is positively associated with quality, as those tests may help medical staff to better understand the nature of the problems of the patients. REFERRALS indicates the inverse of number of referrals (by patient) form primary to secondary care. It is considered that a higher amount of referrals is negatively associated with quality. Consequently, if we express this indicator by its inverse values, the three variables that are included in this latter qualitative category will have the same positive orientation towards quality.

INDINP is obtained using the number of health staff[6] (HLAB) in each of the health centres, which includes GPs, paediatricians, nurses, midwives, physiotherapists, dentists, and X-ray technicians; the number of non-health staff (NHLAB) including staff such as admin staff or social workers; the number of prescriptions (PHARMA) and the area of the centre (AREA).

INDCOST is built from the following variables: the costs of the health staff (HLABCOST) and the non-health staff (NHLABCOST), the costs of the prescriptions (PHARMACOST) and the cost of the area (AREACOST). This latter variable refers most of it to the depreciation of the building where the health centre is located and the rest of it is related to other costs such as, administrative costs.

The main descriptive statistics of all these variables and the different indices obtained from them are presented in Table 1.

[6] We use the equivalent personnel, the reason being that in many situations, staff works in more than one Health Zone. Therefore, we consider the number of hours that, proportionally, each of them work in each of the Health Zones.

Table 1. Descriptive statistics

Variable	Role	Mean	SD	Maximum	Minimum
FREQUENCYGP	Quantitative output	10.28	3.92	20.82	2.74
FREQUENCYP	Quantitative output	6.38	3.13	15.57	1.05
FREQUENCYN	Quantitative output	6.98	3.33	19.46	1.71
FREQUENCYU	Quantitative output	1.51	0.69	3.08	0.18
DAYVISITSGP	Qualitative output I	40.19	10.12	64.79	17.16
DAYVISITSP	Qualitative output I	20.51	8.73	53.26	4.50
DAYVISITSN	Qualitative output I	28.12	11.27	87.61	10.96
EXPERIENCE	Qualitative output II	5.825	534	6.470	3.815
HOSTESTS	Qualitative output II	0.52	0.22	1.17	0.07
REFERRALS	Qualitative output II	3.09	0.69	6.67	1.61
HLAB	Input	0.003185	0.001519	0.009789	0.001396
NHLAB	Input	0.001383	0.000865	0.004620	0.000359
PHARMA	Input	23.62	4.86	35.54	11.59
CAPITAL	Input	0.14	0.09	0.47	0.01
HLABCOST	Cost	210.87	95.74	594.34	89.40
NHLABCOST	Cost	31.83	18.27	85.97	7.67
PHARMACOST	Cost	288.32	112.84	813.73	0.15
CAPITALCOST	Cost	0.84	0.97	7.05	0.00039
INDACT	Index for activity	49.95	31.10	98.89	4.37
INDQUA1	Index for quality I	50.67	26.66	95.96	3.68
INDQUA2	Index for quality II	50.33	22.55	93.79	7.22
INDQUAT	Index for total quality	49.35	27.07	93.46	8.88
INDOUT	Index for total output	51.22	26.90	94.57	7.45
INDINP	Index for inputs	43.82	37.25	100.00	1.46
INDCOST	Index for costs	43.63	36.76	100.00	1.49

Source: Prepared by authors

To finish with the description of the data, it is necessary to mention that the different indices are generated by means of Principal Component Analysis (PCA), firstly developed by Pearson (1901) and Hotelling (1933). PCA lineally transform a set of original variables into another smaller set of uncorrelated variables, which receive the name of factors or principal components, and that have most of the information from the original set (Kuroda et al., 2011; Abdi, 2003). Each of the principal components is a linear combination of the standardised values of the variables used in the construction of each of the indices. The number of principal components retrieved depends on the correlation of the initial variables, so, if they are strongly correlated one factor will be enough, whilst if the correlation is week several factor will be required to explain the maximum information of the original variables as possible. In the latter situation, we need to obtain as many intermediate indicators as factors, obtaining the final index by calculating a weighting sum. Correlation levels and the factors needed in the construction of each of the indices are presented in Appendix 4.

4. RESULTS

In this section, we apply a two-way cluster analysis to be able to identify the efficiency with which primary health care services are being delivered in Extremadura, first applying the method of Ward (Ward, 1963) and then the k-means method (Lloyd, 1982; Ball and Hall, 1965; MacQueen, 1967). As we are also interested in whether quality indicators have an important role in efficiency and, therefore, whether they should be considered, we run two cluster models. The first of them is strictly quantitative and in it, we use the indices of activity or quantitative output (INDACT), inputs (INDINP) and costs (INDCOST). The second one incorporates the quality indicators, and because of that, we use the output adjusted by quality or total output (INDOUT), together with inputs (INDINP) and costs (INDCOST).

In the first stage of the cluster analysis, we consider a range of solutions between three and four clusters. After looking at the composition of these clusters we decide to perform the non-hierarchical k-means methods using also a range of solutions between two and four. 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 from the application of this latter method, we consider that the most appropriate solution is the four-cluster option, indicating the existence of four levels of efficiency in primary health care in Extremadura, and that are described later on.

We also determine the stability of the results of the cluster analysis 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 hypothesis that the groups have the same means needs to be rejected, having four cluster significantly different. Additionally, the discriminatory models correctly classify 99.00% and 99.04% of the cases, respectively.

Table 2 shows the mean values of INDACT (index of quantitative output), INDINP (index of input) and INDCOST (index of costs) and of the INDOUT (total output or output adjusted by quality), INDINP and INDCOST in each of the clusters built in both analyses. Considering these values we characterise the groups based on their level of activity and efficiency in the following way: efficient-active, efficient-inactive, inefficient-active, and inefficient-inactive.

The efficient-active cluster is categorised as such because the Health Zones included in it, present an index of output (quantitative as well as total) greater than the corresponding values of input and costs. Furthermore, they are active given that their levels of output are high, compared to the levels that have the Health Zones that are classified in the inactive clusters.

The efficient-inactive cluster is that one that, similarly to the previous case, it can be considered as efficient because the levels of output are greater than their corresponding levels of inputs and costs. However, this group receives the characterisation of inactive because of its reduced level of output.

The group categorised as inefficient active, present high levels of activity and activity adjusted by quality, but their corresponding levels of inputs and costs are greater than its output, therefore, its categorisation as inefficient.

The final cluster receives the name of inefficient-active. Despite this group is using more resources than other groups, its level of output is very reduced and smaller to the corresponding inputs and costs.

As seen from Table 2 and focussing first on the strictly quantitative analysis (analysis quantitative output-input-costs) 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 that they can be considered as efficient Health Zones.

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

ANALYSIS QUANTITATIVE OUTPUT-INPUT-COSTS				
	Cluster 1 (N = 28)	**Cluster 2 (N = 30)**	**Cluster 3 (N = 32)**	**Cluster 4 (N = 14)**
INDACT [mean (SD)]	74.39 (12.01)	15.42 (8.65)	71.49 (22.84)	26.27 (14.41)
INDINP [mean (SD)]	26.02 (18.30)	5.64 (4.56)	89.72 (11.63)	47.24 (23.33)
INDCOST [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				
	Cluster 1 (N = 25)	**Cluster 2 (N = 38)**	**Cluster 3 (N = 28)**	**Cluster 4 (N = 13)**
INDOUT [mean (SD)]	69.92 (11.00)	30.29 (16.07)	74.89 (16.81)	25.04 (10.18)
INDINP [mean (SD)]	39.06 (23.00)	7.99 (9.10)	91.86 (10.81)	58.10 (26.25)
INDCOST [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

In relation to the analysis performed with the output adjusted by quality (analysis total output-input-costs), 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 also to either reduce the resources they use in the provision of their services or increase their activity and quality with which they are offering health care. Additionally, within the efficient Health Zones, there is a reduction in the number of efficient-active ones and an increase in the number of efficient-inactive Zones, as a consequence of the inclusion of the quality variables. This is indicating that the inclusion of qualitative indicators is reducing the total output of certain Health Zones, provided that the quality with which they are offering health care could not be the most appropriate, thereofore, claiming for an improvement in the quality with which health care services are being delivered in Extremadura.

Looking at the Health Zones in which these differences are taking place[7], our results indicate that almost a quarter of them (21 Health Zones – 20.19%) change the cluster where they belong to after the inclusion of the quality indicators, i.e. they are influenced, positively or negatively, by the incorporation of quality indicators. These Health Zones and the clusters they belong to in each of the analyses are presented in Table 3.

[7] The characterisation of each of the Health Zones in one of the clusters for both analyses is presented in Appendix 5 and 6.

Table 3. Health Zones that change in their efficiency level[8]

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
Mérida	Calamonte	Efficient-active	Efficient-inactive
Mérida	Guareña	Efficient-active	Efficient-inactive
Mérida	Villafranca de los Barros	Efficient-active	Efficient-inactive
Llerena-Zafra	Fuente del Maestre	Efficient-active	Efficient-inactive
Llerena-Zafra	Santos de Maimona	Efficient-active	Efficient-inactive
Llerena-Zafra	Zafra II	Efficient-active	Efficient-inactive
Plasencia	Jaraíz de la Vera	Efficient-active	Efficient-inactive
Plasencia	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
Plasencia	Casas del Castañar	Inefficient-active	Inefficient-inactive
Plasencia	Pinofranqueado	Inefficient-active	Inefficient-inactive
Changes from inefficient-inactive to efficient-active			
Badajoz	Alburquerque	Inefficient-inactive	Efficient-active
Badajoz	San Vicente de Alcántara	Inefficient-inactive	Efficient-active
Badajoz	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

As observed from the previous table, there are Health Zones in which the inclusion of the qualitative indicators have a positive impact improving with that the total output compared to the quantitative output. 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.

For the first case, improvement of the efficiency by the incorporation of quality indicators, 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 analysis. These six Health Zones present high indices of input and cost and low scores in the quatitative output (being therefore qualified as inefficient-inactive). However, they are characterised by having a good quality, and because of that, the total output improves in

[8] A table with the scores of the indices for each of these Health Zones in presented in Appendix 7.

such a way that they become efficient-active when the output adjusted by quality is considered.

For the second case, negative effects of the inclusion of quality indicators, relevant results are the ones provided by the nine Health Zones that, despite being efficient in both analyses because their levels of input and costs are very low, the incorporation of quality indicators worsens the level of activity so they become inactive when the quality indicators are incorporated. Actions in relation to these Health Zones should be taken in order to be able to increase the quality with which health services are being delivered. A similar situation is represented by four Health Zones for which the incorporation of quality indicators worsens their outputs changing from being active to inactive. Apart from that, in this case, the Health Zones are also inefficient, and consquently, actions are not only needed to increase quality so they become active, but also 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 Health Zones.

As seen from the previous results, the inclusion of quality indicators affects the efficiency of the health system, which higlights, how important it is the consideration of this type of variables in the analysis of primary health care. Furthermore, a high proportion of Health Zones are inefficient which is indicating that efforts need to be done to propertly use the existing resources, to increase the output or to keep producing the similar outputs but with a considerable reduction of their inputs and costs so that those Health Zones can be efficient.

5. DISCUSSION AND CONCLUSION

The study developed in this research has analysed whether different efficiency levels exists in the provision of primary health care services in the Autonomous Community of Extremadura.

To that end that we have extracted data from APEX08 (Murillo-Zamorano et al., 2011), an information system specifically elaborated for the study of primary health care

in the region and that contains detailed information about the different Health Areas and Health Zones in which primary health care is organised.

Using this dataset and applying Principal Component Analysis (PCA) we have calculated a series of synthetic indices of output, input and costs. A consideration that needs to be highlighted is the fact that for the case of the outputs, we have not only considered the activity (measured using the number of consultation by patient) we have also taken into account certain quality indicators to adjust the quantitative output and obtain an output adjusted by quality.

For the quality indicators, we have considered variables such as the daily caseload in different health specialities, the experience of health staff, referrals and additional tests required to offer a better and more informed health care. The consideration of these latter variables obeys to the idea that in order to be able to adequately measure health output, we need to take into consideration both quantitative (activity) and qualitative (quality) indicators.

In relation to the synthetic index of input we have included the health and non-health staff, the number of prescriptions and the area of the centres where health care services are being delivered. For the last of the indices, costs, we have considered the costs of these input variables.

We have used these indices to perform an efficiency analysis, utilising a cluster methodology. We have followed a two-stage approach, first applying the method of Ward (Ward, 1963) and the proceeding with the non-hierarchical k-means methodology (Lloyd, 1982; Ball and Hall, 1965, and MacQueen, 1967). After studying the different solutions obtained from both methods and using a discriminant analysis (Greenly et al., 2005) to guide the decision, we have selected four clusters as the most appropriate solution, defining four different levels of efficiency for every Health Zone in the region, based on their level of output in comparison with their levels of inputs and costs: 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 level of activity.

These four groups were the most appropriate solution for the two cluster analyses that we performed, depending on whether or not qualitative indicators were taken into consideration. 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 total output or output adjusted by quality.

In relation to the first analysis, that used the quantitative output, input and costs, the results indicated the need of an increase in efficiency, either with a reduction in the resources used or an improvement in the activity delivered , given that 46 (44.23%) out of the 104 Health Zones participating in the study were classified in the inefficient clusters. Within the 58 (55.77%) Health Zones classified as efficient, there is a need to increase the activity in 30 of them that were classified as inactive.

The incorporation of quality indicators considerably affected the results, both positively and negatively. The analysis considering the indices of total output, input and costs provided less Health Zones classified within the inefficient clusters (41 out the 104 Zones were in those two groups) than the analysis that did not include quality indicators, which could be due to an increase in the output resulting by the incorporation of good quality indicators into the activity variable. However, the incorporation of quality variables affected negatively to the number of efficient-active Health Zones, indicating the existence of health units that are affected negatively by the consideration of the variables that describe the quality with which they provide health services. The differences in the classification of the different health units affected a total of 21 Health Zones (20.19%).

The need to consider the inclusion of quality indicators in the analysis of the performance of the different health units is evident, given its influence in the levels of efficiency and activity. The development of research that does not include these distinctive characteristics could lead to misleading results and to the inadequate

assessment of the activity developed by the different 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 those 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 health care and towards them it should be organised. As indicated by Donabedian (1984), if we want to fully describe the quality in health care, we need to consider patients' views. 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. Accordingly, the next necessary step should be the analysis of the quality of the health care under the patients' point of view, so that we can, in the end, offer a health service with the highest possible quality and with which patients are highly satisfied.

REFERENCES

Abdi, H. (2003). Multivariate analysis. In: Lewis-Beck, M., Bryman, A. and Futing, T. (eds.). Encyclopedia for Research Methos for Social Sciences. Sage, Thousand Oaks, CA.

Amado, C., and Dyson, R. (2008). On comparing the performance primary care providers. *European Journal of Operational Research*, 185 (3), pp. 1469-1482.

Amado, C., and 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), pp. 1469-1482.

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

Chilingerian, J., and Sherman, H. (1996). Benchmarking physician practice patterns with DEA: a multi-stage approach for cost containment. *Annual Operational Research,* 67 (1), pp. 86-116.

Chilingerian, J., and Sherman, H. (1997). DEA and primary care physicians report cards: Deriving preferred practice cones from managed care service concepts and operating stratergies. *Annual Operational Research*, 73, pp. 36-66.

Cordero, J. M., Crespo, E. and Murillo-Zamorano, L. R. (2014). The effect of quality and socio-demographic variables on efficiency measures in primary health care. *The European Journal of Health Economics*, 15 (3), pp. 289-302.

Donabedian, A. (1984). La Calidad de la Atención Médica. *La Prensa Mexicana*, México.

García, F., Marcuello, C., Serrano, D., and Urbina, O. (1999). Evaluation of efficiency in primary health care centres: an application of Data Envelopment Analysis. *Financial Accountability & Management*, 15 (1), pp.67-83.

Godoy Caballero, A.L. and Murillo Zamorano, L.R. (2012). ANÁLISIS ECONÓMICO DE LA ASISTENCIA SANITARIA EN LA ATENCIÓN PRIMARIA. El caso de la Comunidad Autónoma de Extremadura. *Editorial Académica Española.* I.S.B.N. 978-3-8473-6526-6

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

Greenly, G. E., Hooley, G. J. and Rudd, J. M. (2005). Market orientation in a multiple stakeholder orientation context: implications for marketing capabilities and assets. Journal of Business Research, 58 (11), pp. 1483-1494.

Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24 (6), pp. 417-441.

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

Jobson, J. (1992). Applied Multivariate Data Analysis. Volume II: Categorical and Multivariate Methods. New York. *Editorial Springer-Verlag*. I.S.B.N.: 0-387-97804-6.

Klecka, W .R. (1980). Discriminant analysis. *Sage*, 19.

Kuroda, M., Mori, Y., Iizuka, M., and Sakakihara, M. (2011). Acceleration of the alternating least squares algorithm for principal components analysis. *Computational Statistics and Data Analysis*, 55 (1), pp. 143-153.

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

López-Sánchez, J. A. and Santos-Vijande, M. L. (2015). How value creation and relationship quality coalignment affects a firms's performance: An empirical analysis. *Journal of Marketing Channels*, 22 (3), pp. 214-230.

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, pp. 281–297.

Murillo-Zamorano, L. R., Vega-Cervera, J. A., Chamorro, A., Cordero, J. M., Crespo, E., de Miguel, F. J., Miranda, F. J., Nogales, L., and Sánchez, M. (2012). Eficiencia y Calidad de la atención primaria en Extremadura. Murillo and Vega eds. ISBN:978-84-614-7542-1.

Murillo-Zamorano, L.R. and Petraglia, C. (2011). Technical efficiency in Primary Health Care: Does quality matter? *The European Journal of Health Economics*, 12 (2). pp. 115-125.

Murillo-Zamorano, L.R., Vega, J., Corrales, N., Fernández, Y., de Miguel, F.J., 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.

Nunamaker, T.R. (1983). Measuring routine nursing service efficiency: A comparison of cost per patient day and data envelopment analysis models. *Health Services Research,* 18 (2 Pt 1), pp. 183-205.

Ozcan, Y. A. (1998). Physician benchmarking: measuring variation in practice behaviour in treatment of otitis media. *Health Care Management Sciences*, 1 (1), pp. 5-17.

Pearson, K (1901b) On Lines and Planes of Closest Fit to Systems of Points in Space. *Philosophical Magazine*, 2, pp. 559–72.

Pina, V., and Torres, L. (1992). Evaluating the efficiency of non-profit organizations: An application of data envelopment analysis to the public health services. *Financial Account*, 8 (3), pp. 213-225.

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

Salinas-Jiménez, J., and Smith, P.C. (1996). Data Envelopment Analysis applied to quality in primary health care. *Annals of Operational Research*, 67 (1), pp. 141-161.

Ward, J. H. Jr. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58 (301), pp. 236-244

Wong, M. A. (1982). A hybrid clustering method for identifying high-density clusters. Journal of the American Statistical Association, 77 (380), pp. 841-847.

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

APPENDICES

Appendix 1. Health Zones by Health Area

Health Zone of Badajoz (I)		
Alburquerque	Badajoz-Valdepasillas	Pueblonuevo del Guadiana
Alconchel	Badajoz-Zona Centro	Roca de la Sierra (La)
Badajoz-Ciudad Jardín	Barcarrota	San Vicente de Alcántara
Badajoz-La Paz	Jerez de los Caballeros	Santa Marta
Badajoz-Progreso	Montijo	Talavera la Real
Badajoz-San Fernando	Oliva de la Frontera	Villanueva del Fresno
Badajoz-San Roque	Olivenza	
Health Zone of Mérida (II)		
Aceuchal	Guareña	Mérida-Polígono Nueva Ciudad
Almendralejo-San Roque	Hornachos	Mérida-San Luis
Calamonte	Mérida-Norte	Villafranca de los Barros
Cordobilla de Lácara	Mérida-Obispo Paulo	Zarza de Alange
Health Zone of Don Benito-Villanueva (III)		
Cabeza del Buey	Navalvillar de Pela	Villanueva de la Serena-Norte
Campanario	Orellana la Vieja	Villanueva de la Serena-Sur
Castuera	Santa Amalia	Zalamea de la Serena
Don Benito-Este	Siruela	
Herrera del Duque	Talarrubias	
Health Zone of Llerena-Zafra (IV)		
Azuaga	Fuente del Maestre	Santos de Maimona (Los)
Fregenal de la Sierra	Llerena	Zafra I
Fuente de Cantos	Monesterio	Zafra II
Health Zone of Cáceres (V)		
Alcántara	Cáceres-Plaza de Toros	Talaván
Alcuéscar	Cáceres-Sur	Trujillo-Rural
Arroyo de la Luz	Guadalupe	Trujillo-Urbano
Berzocana	Miajadas	Valdefuentes
Cáceres-Aldea Moret	Navas del Madroño	Valencia de Alcántara
Cáceres-Centro	Salorino	Zorita
Cáceres-Norte	Santiago de Alcántara	
Health Zone of Coria (VI)		
Ceclavín	Moraleja	Valverde del Fresno
Coria	Torre de Don Miguel	
Hoyos	Torrejoncillo	
Health Zone of Navalmoral de la Mata (VII)		
Almaraz	Losar de la Vera	Villanueva de la Vera
Bohonal de Ibor	Navalmoral de la Mata	Villar del Pedroso
Castañar de Ibor	Talayuela	
Health Zone of Plasencia (VIII)		
Ahigal	Jaraíz de la Vera	Plasencia-Luis de Toro
Aldeanueva del Camino	Mohedas de Granadilla	Plasencia-Norte/Plasencia-La Data
Cabezuela del Valle	Montehermoso	Plasencia-Sur/Plasencia-San Miguel
Casas del Castañar	Nuñomoral	Serradilla
Hervás	Pinofranqueado	

Source: Prepared by authors

Appendix 2. Synthetic indices of quantitative output (INDACT), total output (INDOUT), input (INDINP) and cost (INDCOST) and rank by Health Zone

Health Zone	INDACT		INDOUT		INDINP		INDCOST	
	Index	Rank	Index	Rank	Index	Rank	Index	Rank
Health Area of Badajoz								
Alburquerque	25.00	75	65.50	37	23.38	59	35.48	50
Alconchel	57.52	47	84.05	14	93.02	17	84.62	27
Badajoz - Ciudad Jardín	6.64	98	17.89	84	2.19	102	1.86	103
Badajoz - La Paz	4.77	103	56.13	51	2.88	95	5.48	90
Badajoz - Progreso	5.09	102	52.69	56	6.32	83	6.44	87
Badajoz - San Fernando	14.03	87	12.67	91	6.38	82	11.58	73
Badajoz - San Roque	8.01	95	49.59	60	2.55	98	4.03	94
Badajoz - Valdepasillas	4.37	104	52.27	59	1.46	104	1.49	104
Badajoz - Zona Centro	7.07	96	52.68	58	5.25	85	7.53	83
Barcarrota	63.43	44	64.89	40	24.82	57	45.25	45
Jerez de los Caballeros	23.64	77	17.21	87	53.07	44	59.97	38
Montijo	25.43	73	16.90	89	3.92	89	7.40	84
Oliva de la Frontera	37.18	60	34.43	73	34.00	53	38.37	49
Olivenza	21.76	78	36.38	72	6.70	81	14.06	70
Pueblonuevo del Guadiana	28.19	68	37.07	69	23.79	58	25.96	54
Roca de la Sierra	39.76	58	17.11	88	89.90	22	83.42	29
San Vicente de Alcántara	15.60	83	65.25	38	87.31	23	28.45	53
Santa Marta	14.29	85	36.63	71	13.84	66	21.39	60
Talavera la Real	11.63	89	54.59	54	6.74	80	13.10	71
Villanueva del Fresno	28.60	66	61.78	42	60.20	39	15.51	64
Health Area of Mérida								
Aceuchal	79.42	25	58.88	48	6.77	79	9.24	76
Almendralejo - San José	6.20	100	7.45	104	51.85	45	12.61	72
Calamonte	88.12	15	44.28	62	27.50	55	11.15	74
Cordobilla de Lácara	84.54	19	81.47	18	99.00	11	88.78	23
Guareña	50.73	54	20.19	82	12.54	67	20.17	62
Hornachos	27.51	70	67.99	33	67.61	35	22.10	59
Mérida - Norte	25.41	74	69.79	31	7.23	77	8.41	79
Mérida - Obispo Paulo	8.84	92	8.28	103	4.31	86	1.91	102
Mérida - Polígono Nueva Ciudad	6.79	97	12.25	93	2.28	100	5.88	89
Mérida - San Luis	13.63	88	9.54	101	3.32	92	25.57	55
Villafranca de los Barros	70.85	37	38.19	68	8.78	74	8.08	81
Zarza de Alange	68.76	40	72.30	27	8.47	75	9.83	75
Health Area of Don Benito-Villanueva								
Cabeza del Buey	90.55	11	93.34	2	92.50	19	96.39	13
Campanario	80.66	23	68.46	32	14.33	64	14.95	67
Castuera	97.52	4	70.55	29	57.19	42	88.43	24
Don Benito - Este	29.60	65	12.03	95	4.14	87	89.18	22
Herrera del Duque	95.08	6	92.89	5	74.81	32	89.86	20
Navalvillar de Pela	87.36	17	79.04	24	49.28	46	47.11	42
Orellana la Vieja	73.78	35	41.04	65	92.95	18	91.96	18
Santa Amalia	98.52	2	78.57	25	16.27	63	3.99	95
Siruela	98.89	1	94.57	1	91.88	20	95.93	16
Talarrubias	6.50	99	55.14	52	58.90	41	47.38	41
Villanueva de la Serena - Norte	20.72	79	9.83	100	3.19	94	22.46	57
Villanueva de la Serena - Sur	5.79	101	12.94	90	8.44	76	6.51	85
Zalamea de la Serena	90.40	12	79.43	22	60.39	38	58.53	39

Source: Prepared by authors

Appendix 2. Synthetic indices of quantitative output (INDACT), total output (INDOUT), input (INDINP) and cost (INDCOST) and rank by Health Zone (Cont.)

Health Zone	INDACT		INDOUT		INDINP		INDCOST	
	Index	Rank	Index	Rank	Index	Rank	Index	Rank
Health Area of Llerena-Zafra								
Azuaga	85.25	18	80.23	20	45.38	48	28.74	52
Fregenal de la Sierra	80.06	24	66.21	35	31.41	54	22.21	58
Fuente de Cantos	28.42	67	22.02	81	11.40	69	15.09	66
Fuente del Maestre	70.09	38	29.56	75	11.88	68	4.77	91
Llerena	84.11	20	70.06	30	18.85	61	8.98	77
Monesterio	75.94	31	43.28	63	53.61	43	15.21	65
Santos de Maimona	66.22	41	38.39	67	9.43	72	7.83	82
Zafra I	14.04	86	8.92	102	3.45	90	4.48	92
Zafra II	56.87	50	60.98	45	8.89	73	8.69	78
Health Area of Cáceres								
Alcántara	97.57	3	84.66	11	99.98	5	100.00	1
Alcuéscar	76.08	30	59.11	46	81.64	28	89.27	21
Arroyo de la Luz	71.51	36	58.92	47	18.31	62	62.63	36
Berzocana	97.39	5	82.04	16	99.99	4	99.91	5
Cáceres - Aldea Moret	26.69	72	22.66	80	10.56	70	14.62	69
Cáceres - Centro	33.51	61	57.21	49	2.71	97	3.54	99
Cáceres - Norte	10.23	91	53.30	55	1.95	103	3.87	97
Cáceres - Plaza de Toros	14.52	84	10.05	99	2.86	96	3.14	101
Cáceres - Sur	8.23	94	27.48	77	61.73	37	29.58	51
Guadalupe	58.19	45	61.33	44	96.43	14	87.25	25
Logrosán	49.89	55	84.36	13	94.42	16	97.14	12
Miajadas	55.35	51	25.18	79	38.45	50	76.14	33
Navas del Madroño	87.80	16	63.66	41	41.67	49	44.44	46
Salorino	93.44	9	79.81	21	99.92	6	99.63	7
Santiago de Alcántara	94.22	8	78.56	26	100.00	2	99.84	6
Talaván	81.05	22	81.89	17	99.81	7	97.69	10
Trujillo - Rural	89.86	13	93.09	4	37.11	51	4.05	93
Trujillo - Urbano	27.22	71	39.45	66	75.47	31	99.99	2
Valdefuentes	94.54	7	92.05	6	71.55	33	90.88	19
Valencia de Alcántara	64.21	43	65.93	36	47.44	47	52.62	40
Zorita	37.39	59	67.20	34	96.95	13	96.17	15
Health Area of Coria								
Ceclavín	32.58	62	17.72	85	71.07	34	73.81	34
Coria	19.87	80	10.51	97	3.94	88	6.47	86
Hoyos	77.47	28	85.66	9	66.47	36	42.47	47
Moraleja	30.47	63	12.36	92	7.15	78	14.74	68
Torre de Don Miguel	75.13	32	42.81	64	99.26	10	98.45	9
Torrejoncillo	74.28	33	81.40	19	90.69	21	78.32	32
Valverde del Fresno	79.09	26	71.41	28	77.76	29	45.70	44

Source: Prepared by authors

Appendix 2. Synthetic indices of quantitative output (INDACT), total output (INDOUT), input (INDINP) and cost (INDCOST) and rank by Health Zone (Cont.)

Health Zone	INDACT		INDOUT		INDINP		INDCOST	
	Index	Rank	Index	Rank	Index	Rank	Index	Rank
Health Area of Navalmoral de la Mata								
Almaraz	42.08	57	54.98	53	98.45	12	96.29	14
Bohonal de Ibor	89.53	14	86.81	8	100.00	3	99.91	4
Castañar de Ibor	69.39	39	52.68	57	99.42	8	99.21	8
Losar de la Vera	24.71	76	19.92	83	35.71	52	40.87	48
Navalmoral de la Mata	8.82	93	12.18	94	3.28	93	3.84	98
Talayuela	16.73	81	25.28	78	5.41	84	6.40	88
Villanueva de la Vera	74.16	34	46.46	61	58.91	40	46.89	43
Villar del Pedroso	78.52	27	84.41	12	100.00	1	99.97	3
Health Area of Plasencia								
Ahigal	57.31	48	36.89	70	77.10	30	65.42	35
Aldeanueva del Camino	77.20	29	86.91	7	85.10	27	79.69	31
Cabezuela del Valle	81.83	21	56.73	50	25.59	56	24.56	56
Casas del Castañar	28.03	69	17.26	86	86.62	24	86.04	26
Hervás	47.12	56	79.17	23	14.00	65	61.66	37
Jaraíz de la Vera	65.69	42	33.08	74	9.61	71	15.63	63
Mohedas de Granadilla	50.76	53	83.68	15	86.15	26	84.33	28
Montehermoso	57.94	46	85.59	10	20.72	60	20.88	61
Nuñomoral	91.80	10	93.14	3	99.32	9	97.33	11
Pinofranqueado	52.54	52	28.75	76	86.46	25	82.03	30
Plasencia - Luis de Toro	11.08	90	10.47	98	2.23	101	3.91	96
Plasencia-Norte/Plasencia-La Data	57.15	49	65.11	39	2.46	99	3.22	100
Plasencia-Sur/Plasencia-San Miguel	15.71	82	11.92	96	3.44	91	8.39	80
Serradilla	29.89	64	61.76	43	96.28	15	95.18	17

Source: Prepared by authors

Appendix 3. Synthetic indices of qualitative output 1 (INDQUA1), qualitative output 2 (INDQUA2), total qualitative output (INDQUAT) and rank by Health Zone

Health Zone	INDQUA1		INDQUA2		INDQUAT	
	Index	Rank	Index	Rank	Index	Rank
Health Area of Badajoz						
Alburquerque	70.04	34	71.50	25	84.87	21
Alconchel	74.14	27	70.61	29	86.06	16
Badajoz - Ciudad Jardín	53.79	53	44.62	56	45.94	53
Badajoz - La Paz	83.59	11	85.82	4	93.30	2
Badajoz - Progreso	80.92	15	75.82	17	90.13	9
Badajoz - San Fernando	42.49	62	34.78	70	25.35	79
Badajoz - San Roque	69.26	37	72.72	21	85.14	19
Badajoz - Valdepasillas	79.27	20	78.13	12	90.42	8
Badajoz - Zona Centro	71.79	32	78.42	11	88.40	14
Barcarrota	27.04	80	70.48	30	50.75	49
Jerez de los Caballeros	73.94	28	13.23	103	29.56	72
Montijo	13.48	91	61.73	40	27.27	76
Oliva de la Frontera	27.57	78	66.47	33	45.46	55
Olivenza	36.24	68	70.13	32	61.05	35
Pueblonuevo del Guadiana	43.88	61	59.94	43	56.19	41
Roca de la Sierra	37.10	67	22.48	93	15.30	92
San Vicente de Alcántara	75.24	23	90.52	2	92.78	3
Santa Marta	79.94	17	39.20	62	67.82	31
Talavera la Real	72.96	30	71.49	26	86.00	17
Villanueva del Fresno	86.84	6	45.16	55	77.82	24
Health Area of Mérida						
Aceuchal	31.58	74	50.40	52	30.80	69
Almendralejo - San José	20.09	86	20.39	95	9.50	103
Calamonte	27.29	79	19.42	97	10.99	98
Cordobilla de Lácara	87.60	5	23.42	92	56.82	39
Guareña	30.57	75	20.90	94	12.43	96
Hornachos	79.80	18	64.70	35	85.56	18
Mérida - Norte	84.41	10	71.19	27	89.59	11
Mérida - Obispo Paulo	5.72	101	40.29	61	11.85	97
Mérida - Polígono Nueva Ciudad	26.62	81	54.03	49	30.27	71
Mérida - San Luis	19.88	87	33.46	73	13.77	94
Villafranca de los Barros	34.45	72	34.09	72	20.23	88
Zarza de Alange	41.60	64	61.15	42	55.18	42
Health Area of Don Benito-Villanueva						
Cabeza del Buey	68.31	38	88.34	3	90.69	7
Campanario	9.57	96	77.43	13	39.86	58
Castuera	3.68	104	70.66	28	27.81	75
Don Benito - Este	6.32	99	32.62	75	9.66	102
Herrera del Duque	55.74	50	82.32	8	84.25	22
Navalvillar de Pela	13.21	93	81.47	9	49.65	50
Orellana la Vieja	34.86	71	34.20	71	20.51	87
Santa Amalia	65.28	42	29.66	79	39.08	59
Siruela	77.60	21	82.54	7	91.28	4
Talarrubias	84.43	9	75.52	18	90.93	6
Villanueva de la Serena - Norte	9.32	97	27.10	83	8.88	104
Villanueva de la Serena - Sur	60.59	47	28.53	82	33.42	68
Zalamea de la Serena	62.96	45	38.10	63	47.73	51

Source: Prepared by authors

Appendix 3. Synthetic indices of qualitative output 1 (INDQUA1), qualitative output 2 (INDQUA2), total qualitative output (INDQUAT) and rank by Health Zone (Cont.)

Health Zone	INDQUA1		INDQUA2		INDQUAT	
	Index	Rank	Index	Rank	Index	Rank
Health Area of Llerena-Zafra						
Azuaga	36.20	69	64.69	36	53.71	43
Fregenal de la Sierra	8.20	98	76.82	14	37.74	62
Fuente de Cantos	28.95	77	57.12	47	35.34	64
Fuente del Maestre	5.22	102	37.46	66	10.80	99
Llerena	15.39	90	71.67	24	38.85	60
Monesterio	12.11	94	53.76	51	20.70	86
Santos de Maimona	6.25	100	64.26	37	24.46	82
Zafra I	3.81	103	37.74	64	10.52	100
Zafra II	20.92	85	76.76	15	52.43	45
Health Area of Cáceres						
Alcántara	67.44	40	37.58	65	52.70	44
Alcuéscar	45.48	59	41.80	59	33.92	67
Arroyo de la Luz	24.91	83	62.66	38	37.72	63
Berzocana	56.16	49	43.24	58	46.83	52
Cáceres - Aldea Moret	45.61	58	45.39	54	38.00	61
Cáceres - Centro	54.16	51	62.44	39	69.23	29
Cáceres - Norte	80.27	16	65.65	34	86.16	15
Cáceres - Plaza de Toros	33.19	73	25.62	86	15.22	93
Cáceres - Sur	89.44	3	25.69	85	61.74	34
Guadalupe	25.46	82	72.38	22	51.63	46
Logrosán	84.79	8	85.55	5	93.46	1
Miajadas	16.36	88	44.11	57	17.27	91
Navas del Madroño	44.10	60	36.94	68	28.22	74
Salorino	72.30	31	28.87	81	45.80	54
Santiago de Alcántara	82.50	12	18.06	99	42.81	57
Talaván	74.64	25	37.38	67	60.73	36
Trujillo - Rural	58.90	48	93.79	1	89.92	10
Trujillo - Urbano	29.19	76	74.84	19	59.47	38
Valdefuentes	75.82	22	58.52	45	80.48	23
Valencia de Alcántara	61.93	46	41.27	60	51.22	47
Zorita	67.49	39	59.05	44	76.02	27
Health Area of Coria						
Ceclavín	50.48	54	24.75	87	22.97	83
Coria	13.46	92	35.41	69	12.44	95
Hoyos	66.17	41	56.15	48	72.77	28
Moraleja	11.88	95	29.19	80	10.01	101
Torre de Don Miguel	40.54	65	30.16	78	20.98	85
Torrejoncillo	35.49	70	74.70	20	65.60	32
Valverde del Fresno	15.49	89	76.34	16	44.97	56

Source: Prepared by authors

Appendix 3. Synthetic indices of qualitative output 1 (INDQUA1), qualitative output 2 (INDQUA2), total qualitative output (INDQUAT) and rank by Health Zone (Cont.)

Health Zone	INDQUA1		INDQUA2		INDQUAT	
	Index	Rank	Index	Rank	Index	Rank
Health Area of Navalmoral de la Mata						
Almaraz	89.06	4	24.47	89	59.84	37
Bohonal de Ibor	94.87	2	24.33	90	65.46	33
Castañar de Ibor	86.62	7	7.22	104	34.16	66
Losar de la Vera	74.89	24	17.51	100	34.54	65
Navalmoral de la Mata	69.75	36	15.29	102	28.26	73
Talayuela	81.30	14	24.75	88	51.03	48
Villanueva de la Vera	63.82	43	16.31	101	25.03	80
Villar del Pedroso	95.96	1	26.48	84	68.66	30
Health Area of Plasencia						
Ahigal	53.85	52	31.40	76	30.65	70
Aldeanueva del Camino	69.99	35	57.56	46	76.48	26
Cabezuela del Valle	21.84	84	54.02	50	26.75	77
Casas del Castañar	63.00	44	18.09	98	25.87	78
Hervás	71.36	33	70.44	31	84.90	20
Jaraíz de la Vera	49.16	55	19.69	96	19.07	90
Mohedas de Granadilla	73.30	29	85.47	6	91.08	5
Montehermoso	74.59	26	79.44	10	89.58	12
Nuñomoral	79.65	19	72.04	23	88.50	13
Pinofranqueado	47.10	56	30.38	77	24.97	81
Plasencia - Luis de Toro	46.37	57	23.56	91	19.96	89
Plasencia-Norte/Plasencia-La Data	42.16	63	61.59	41	56.45	40
Plasencia-Sur/Plasencia-San Miguel	37.99	66	32.90	74	21.35	84
Serradilla	81.74	13	47.89	53	76.69	25

Source: Prepared by authors

Appendix 4. Synthetic indices, extracted factors and PCA information

Index	No. of factors	Rotation	% of variance	Factor weights
Quantitative output (INDACT)	2	Yes	71,880	54,8/45,2
Qualitative output 1 (INDQUA1)	2	Yes	74,821	53,1/46,9
Qualitative output 2 (INDQUA2)	2	Yes	73,467	54,6/45,4
Total qualitative output (INDQUAT)	1	No	50,760	100
Total output (INDOUT)	1	No	50,092	100
Inputs (INDINP)	1	No	76,067	100
Costs (INDCOST)	1	No	68,250	100

INDACT: 0.548*F1 + 0.452*F2

F1 = 0.925*FREQUENCYGP + 0.111*FREQUENCYP + 0.828*FREQUENCYN + 0.150*FREQUENCYU

F2 = 0.028*FREQUENCYGP + 0.765*FREQUENCYP + 0.334*FREQUENCYN + 0.776*FREQUENCYU

INDQUA1[9]: 100 – [0.531*F1 + 0.469*F2]

F1 = 0.892*DAYVISITSGP + 0.623*DAYVISITSP + 0.092*DAYVISITSN

F2 = 0.028*DAYVISITSGP + 0.373*DAYVISITSP + 0.955*DAYVISITSN

INDQUA2: 0.564*F1 + 0.454*F2

F1 = 0.029*EXPERIENCE + 0.774*HOSTESTS + 0.777* REFERRALS

F2 = 1.000*EXPERIENCE + 0.031*HOSTESTS + 0.013* REFERRALS

INDQUAT: 0.712* (INDQUA1 + INDQUA2)

INDOUT: 0.708*(INDOUT + INDQUAT)

INDINP: 0.933*HALB + 0.937*NHLAB + 0.777*PHARMA + 0.831*AREA

INDCOST: 0.881*HLABCOST + 0.889*NHLABCOST + 0.785*PHARMACOST + 0.741*CAPITALCOST

Source: Prepared by authors

[9] In the computation of INDQUA1 we first obtained an index for the daily visits in each of the specialities and then build the final qualitative output, the reason being that the daily caseload in negatively associated with quality, and INDQUA1 is positively associated with quality.

Appendix 5. Characterisation of the Health Zones in each of the clusters. Analysis activity-input-costs (INDACT-INDINP-INDCOST)

Efficient-active (N = 26)			
Health Zone of Badajoz			
Barcarrota			
Health Zone of Mérida			
Aceuchal	Calamonte	Guareña	Villafranca de los Barros
Zarza de Alange			
Health Zone of Don Benito-Villanueva de la Serena			
Campanario	Navalvillar de Pela	Santa Amalia	Zalamea de la Serena
Health Zone of Llerena-Zafra			
Azuaga	Fregenal de la Sierra	Fuente del Maestre	Llerena
Monesterio	Santos de Maimona	Zafra II	
Health Zone of Cáceres			
Arroyo de la Luz	Navas del Madroño	Trujillo - Rural	Valencia de Alcántara
Health Zone of Coria			
Hoyos	Valverde del Fresno		
Health Zone of Navalmoral de la Mata			
Villanueva de la Vera			
Health Zone of Plasencia			
Cabezuela del Valle	Jaraíz de la Vera	Montehermoso	Plasencia-Norte/La Data
Efficient-inactive (N = 30)			
Health Zone of Badajoz			
Badajoz - Ciudad Jardín	Badajoz - La Paz	Badajoz - Progreso	Badajoz - San Fernando
Badajoz - San Roque	Badajoz - Valdepasillas	Badajoz - Zona Centro	Montijo
Olivenza	Pueblonuevo del Guadiana	Santa Marta	Talavera la Real
Health Zone of Mérida			
Mérida - Norte	Mérida - Obispo Paulo	Mérida - Polígono Nueva Ciudad	Mérida - San Luis
Health Zone of Don Benito-Villanueva de la Serena			
Villanueva de la Serena - Norte	Villanueva de la Serena - Sur		
Health Zone of Llerena-Zafra			
Fuente de Cantos	Zafra I		
Health Zone of Cáceres			
Cáceres - Aldea Moret	Cáceres - Centro	Cáceres - Norte	Cáceres - Plaza de Toros
Health Zone of Coria			
Coria	Moraleja		
Health Zone of Navalmoral de la Mata			
Navalmoral de la Mata	Talayuela		
Health Zone of Plasencia			
Plasencia - Luis de Toro	Plasencia-Sur/San Miguel		

Source: Prepared by authors

Appendix 5. Characterisation of the Health Zones in each of the clusters. Analysis activity-input-costs (INDACT-INDINP-INDCOST) (Cont.)

Inefficient-active (N = 34)			
Health Zone of Badajoz			
Alconchel	Roca de la Sierra		
Health Zone of Mérida			
Cordobilla de Lácara			
Health Zone of Don Benito-Villanueva de la Serena			
Cabeza del Buey	Castuera	Herrera del Duque	Orellana la Vieja
Siruela			
Health Zone of Llerena-Zafra			
-			
Health Zone of Cáceres			
Alcántara	Alcuéscar	Berzocana	Guadalupe
Logrosán	Salorino	Santiago de Alcántara	Talaván
Trujillo - Urbano	Valdefuentes	Zorita	
Health Zone of Coria			
Torre de Don Miguel	Torrejoncillo		
Health Zone of Navalmoral de la Mata			
Almaraz	Bohonal de Ibor	Castañar de Ibor	Villar del Pedroso
Health Zone of Plasencia			
Ahigal	Aldeanueva del Camino	Casas del Castañar	Mohedas de Granadilla
Nuñomoral	Pinofranqueado	Serradilla	
Inefficient-inactive (N = 14)			
Health Zone of Badajoz			
Alburquerque	Jerez de los Caballeros	Oliva de la Frontera	San Vicente de Alcántara
Villanueva del Fresno			
Health Zone of Mérida			
Almendralejo - San José	Hornachos		
Health Zone of Don Benito-Villanueva de la Serena			
Don Benito - Este	Talarrubias		
Health Zone of Llerena-Zafra			
-			
Health Zone of Cáceres			
Cáceres - Sur	Miajadas		
Health Zone of Coria			
Ceclavín			
Health Zone of Navalmoral de la Mata			
Losar de la Vera			
Health Zone of Plasencia			
Hervás			

Source: Prepared by authors

Appendix 6. Characterisation of the Health Zones in each of the clusters. Analysis activity adjusted by quality-input-costs (INDOUT-INDINP-INDCOST)

Efficient-active (N = 26)			
Health Zone of Badajoz			
Alburquerque	Barcarrota	San Vicente de Alcántara	Villanueva del Fresno
Health Zone of Mérida			
Hornachos	Mérida - Norte	Zarza de Alange	
Health Zone of Don Benito-Villanueva de la Serena			
Campanario	Navalvillar de Pela	Santa Amalia	Talarrubias
Zalamea de la Serena			
Health Zone of Llerena-Zafra			
Azuaga	Fregenal de la Sierra	Llerena	Monesterio
Health Zone of Cáceres			
Arroyo de la Luz	Navas del Madroño	Trujillo - Rural	Valencia de Alcántara
Health Zone of Coria			
Hoyos	Valverde del Fresno		
Health Zone of Navalmoral de la Mata			
-			
Health Zone of Plasencia			
Cabezuela del Valle	Hervás	Montehermoso	
Efficient-inactive (N = 30)			
Health Zone of Badajoz			
Badajoz - Ciudad Jardín	Badajoz - La Paz	Badajoz - Progreso	Badajoz - San Fernando
Badajoz - San Roque	Badajoz - Valdepasillas	Badajoz - Zona Centro	Montijo
Olivenza	Pueblonuevo del Guadiana	Santa Marta	Talavera la Real
Health Zone of Mérida			
Aceuchal	Calamonte	Guareña	Mérida - Obispo Paulo
Mérida - Polígono Nueva Ciudad		Mérida - San Luis	Villafranca de los Barros
Health Zone of Don Benito-Villanueva de la Serena			
Villanueva de la Serena - Norte	Villanueva de la Serena - Sur		
Health Zone of Llerena-Zafra			
Fuente de Cantos	Fuente del Maestre	Santos de Maimona	Zafra I
Zafra II			
Health Zone of Cáceres			
Cáceres - Aldea Moret	Cáceres - Centro	Cáceres - Norte	Cáceres - Plaza de Toros
Health Zone of Coria			
Coria	Moraleja		
Health Zone of Navalmoral de la Mata			
Navalmoral de la Mata	Talayuela		
Health Zone of Plasencia			
Jaraíz de la Vera	Plasencia - Luis de Toro	Plasencia-Norte/La Data	Plasencia-Sur/San Miguel

Source: Prepared by authors

Appendix 6. Characterisation of the Health Zones in each of the clusters. Analysis activity adjusted by quality-input-costs (INDOUT-INDINP-INDCOST) (Cont.)

Inefficient-active (N = 34)			
Health Zone of Badajoz			
Alconchel			
Health Zone of Mérida			
Cordobilla de Lácara			
Health Zone of Don Benito-Villanueva de la Serena			
Cabeza del Buey	Castuera	Herrera del Duque	Orellana la Vieja
Siruela			
Health Zone of Llerena-Zafra			
-			
Health Zone of Cáceres			
Alcántara	Alcuéscar	Berzocana	Guadalupe
Logrosán	Salorino	Santiago de Alcántara	Talaván
Trujillo - Urbano	Valdefuentes	Zorita	
Health Zone of Coria			
Torre de Don Miguel	Torrejoncillo		
Health Zone of Navalmoral de la Mata			
Almaraz	Bohonal de Ibor	Castañar de Ibor	Villar del Pedroso
Health Zone of Plasencia			
Aldeanueva del Camino	Mohedas de Granadilla	Nuñomoral	Serradilla
Inefficient-inactive (N = 14)			
Health Zone of Badajoz			
Jerez de los Caballeros	Oliva de la Frontera	Roca de la Sierra	
Health Zone of Mérida			
Almendralejo - San José			
Health Zone of Don Benito-Villanueva de la Serena			
Don Benito - Este			
Health Zone of Llerena-Zafra			
-			
Health Zone of Cáceres			
Cáceres - Sur	Miajadas		
Health Zone of Coria			
Ceclavín			
Health Zone of Navalmoral de la Mata			
Losar de la Vera	Villanueva de la Vera		
Health Zone of Plasencia			
Ahigal	Casas del Castañar	Pinofranqueado	

Source: Prepared by authors

Appendix 7. Health Zones that change in their efficiency level after the inclusion of quality indicators

Health Area	Health Zone	INDACT	INDOUT	INDINP	COST	Analysis activity-input-cost	Analysis activity adjusted by quality-input-cost
Changes from efficient-active to efficient-inactive							
Mérida	Aceuchal	79.42	58.88	6.77	9.24	Efficient-active	Efficient-inactive
Mérida	Calamonte	88.12	44.28	27.5	11.15	Efficient-active	Efficient-inactive
Mérida	Guareña	50.73	20.19	12.54	20.17	Efficient-active	Efficient-inactive
Mérida	Villafranca de los Barros	70.85	38.19	8.78	8.08	Efficient-active	Efficient-inactive
Llerena-Zafra	Fuente del Maestre	70.09	29.56	11.88	4.77	Efficient-active	Efficient-inactive
Llerena-Zafra	Santos de Maimona	66.22	38.39	9.43	7.83	Efficient-active	Efficient-inactive
Llerena-Zafra	Zafra II	56.87	60.98	8.89	8.69	Efficient-active	Efficient-inactive
Plasencia	Jaraíz de la Vera	65.69	33.08	9.61	15.63	Efficient-active	Efficient-inactive
Plasencia	Plasencia-Norte/Plasencia-La Data	57.15	65.11	2.46	3.22	Efficient-active	Efficient-inactive
Changes from efficient-active to inefficient-inactive							
Navalmoral de la Mata	Villanueva de la Vera	74.16	46.46	58.91	46.89	Efficient-active	Inefficient-inactive
Changes from efficient-inactive to efficient-active							
Mérida	Mérida - Norte	25.41	69.79	7.23	8.41	Efficient-inactive	Efficient-active
Changes from inefficient-active to inefficient-inactive							
Badajoz	Roca de la Sierra	39.76	17.11	89.9	83.42	Inefficient-active	Inefficient-inactive
Plasencia	Ahigal	57.31	36.89	77.10	65.42	Inefficient-active	Inefficient-inactive
Plasencia	Casas del Castañar	28.03	17.26	86.62	86.04	Inefficient-active	Inefficient-inactive
Plasencia	Pinofranqueado	52.54	28.75	86.46	82.03	Inefficient-active	Inefficient-inactive
Changes from inefficient-inactive to efficient-active							
Badajoz	Alburquerque	25.00	65.50	23.38	35.48	Inefficient-inactive	Efficient-active
Badajoz	San Vicente de Alcántara	15.60	65.25	87.31	28.45	Inefficient-inactive	Efficient-active
Badajoz	Villanueva del Fresno	28.60	61.78	60.20	15.51	Inefficient-inactive	Efficient-active
Mérida	Hornachos	27.51	67.99	67.61	22.10	Inefficient-inactive	Efficient-active
Don Benito-Villlanueva	Talarrubias	6.50	55.14	58.90	47.38	Inefficient-inactive	Efficient-active
Plasencia	Hervás	47.12	79.17	14.00	61.66	Inefficient-inactive	Efficient-active

Source: Prepared by authors

CHAPTER 2: THE w-HEALTHQUAL: A MEASUREMENT SCALE FOR THE ANALYSIS OF PATIENTS' SATISFACTION WITH PRIMARY HEALTH CARE

1. INTRODUCTION

Primary health care is the first contact point with the health care system for patients. Therefore, it is essential to offer a high quality service, a quality that needs to be recognised as high quality by both, clinicians and patients. Service quality has been recognised as one of the values of central importance in organisations, in both manufacturing and service sectors (Berry et al., 1985; Bitner et al., 1990; LeBlanc and Nguyen, 1988).

Particularly for the health care field, quality can be understood as an objective measure related to how the health system provides its services. Indicators for this dimension of quality include referrals from primary to secondary care, number of consultations per doctor and per day, number of complementary tests, mortality rates, etc. Despite this, it is impossible to fully describe quality if we do not take into account the user's point of view (Donabedian, 1984). Considering patients' perceptions is important as patients provide reliable data and valid diagnosis about their health and about how they feel after seeing their doctor (Dawson et al., 1998). Furthermore, as indicated by Cronin and Taylor (1992), "service quality is an antecedent of consumer satisfaction". Considering this, efforts need to be focused on offering a high quality service as perceived by the patient, which will ultimately be translated into a greater patient's satisfaction. Martilla and James (1977) establish that the importance that users of a particular service give to the elements of that service will influence how satisfied they are with it. However, when measuring importance together with perception there are discrepancies in the literature, with authors rejecting the idea of including this indicator (Adil et al., 2013) and others supporting it as importance may increase the diagnosis power of the scale being used (Jain and Gupta, 2004)

With these ideas in mind, the aim of the present study is to develop a scale that allows us to adequately measure patients' satisfaction with primary care as perceived by the user, i.e. to obtain a measurement instrument that considers all the relevant aspects of primary health care and that precisely reflects patients' priorities.

In doing so, we, firstly, present a weighted scale specially designed to measure patients' satisfaction with primary health care that comprises not only patients' preferences, but also a weighting factor to incorporate the importance of the elements being assessed. Secondly, we analyse how patients' satisfaction varies within settings.

To the best of our knowledge there is not a scale in the literature specifically created for the study of satisfaction in primary health care that includes patient' preferences with the health care together with the importance patients attribute to those preferences. Furthermore, there is not any piece of research aiming to identify differences in satisfaction within settings, in spite of the fact that users' preferences with certain services have previously been measured across countries and systems (Schäfer et al., 2015).

In order to achieve the previously mentioned objective, the rest of the research is structured as follows. First, we analyse the different measures of service quality proposed in the literature and that have been the basis for the construction of our measurement instrument, the weighted HEALTHQUAL scale (w-HEALTHQUAL). Next, we describe the methods applied to build our scale, using information provided by primary health care users in the Spanish region of Extremadura. This new scale allows us to adequately measure patients' satisfaction with primary health care. Following the methods section, we present the results of our analysis, firstly, looking at the characteristics of the w-HEALTQUAL; secondly, performing a cluster analysis that allows us to identify groups of patients that differ among them based on their levels of satisfaction with the health care, and thirdly, representing the levels of satisfaction across the health map of the region of Extremadura. Finally, we present a discussion of the main findings and conclusions of the research.

2. CONCEPTUAL FRAMEWORK

The w-HEALTHQUAL is developed to measure patients' satisfaction with primary health care and it is originated on the basis of a previous scale, the HEALTHQUAL (Murillo-

Zamorano et al., 2012; Miranda et al., 2010), an adaptation of other measurement scales: the SERVQUAL (Parasuraman et al., 1985, 1988) and SERVPERF (Cronin and Taylor, 1992).

The SERVQUAL has traditionally been the most accepted instrument when measuring the perceived quality of a service. It is based on the idea that the perception of a service's quality should be assessed by analysing the gap between consumers' expectations of that service and their experience of the performance of that particular service. The SERVQUAL consists of a tool originally created to evaluate consumers' perceptions of quality in retail organisations. It has been used in a wide variety of sectors and countries (Badri et al., 2005; Kilbourne et al., 2004). However, despite the popularity of the scale, many authors have criticised the instrument designed by Parasuraman and colleagues, defending the idea that just performance-based measures, rather than a combination of the experience of service users and their expectations of the service being evaluated, can capture consumers' perception of the quality of a service (Mazis et al., 1975; Churchill and Surprenant, 1982; Woodruff et al., 1983; Brown et al., 1993; Peter et al., 1993, among others).

Following this approach, Cronin and Taylor (1992) modified the initial instrument and tested a performance-based measure in order to find an alternative approach to the original scale. In their analysis, they conclude that the "performance items adequately define the domain of service quality", producing the SERVPERF. This new scale (which only considers users' experiences with the service being evaluated) "explains more of the variation in service quality than does the SERVQUAL scale", i.e. it is an enhanced instrument for measuring service quality (Adil et al., 2013).

One of the recommendations when using the SERVQUAL and SERVPERF has been to adapt the measurement instruments to the research field where they are going to be used (Ladhari, 2009), so that, any biases that this aspect could introduce can be eliminated. Following this recommendation, Murillo-Zamorano et al., (2012) and Miranda et al. (2010) designed the HEALTHQUAL scale, an adaptation of the SERVPERF specifically created for the analysis of patients' perception of the quality of primary health care. In that adaptation some questions were eliminated given that they were closely related to

the profit orientated nature of the service where they were first used. Furthermore, the decision of what variables should be included in the scale was based on a detailed review of the literature about the dimensions of the quality of health services.

However, and despite the strengths of the HEALTHQUAL, it does not take into account any information about the importance that the different attributes being considered have for patients and that allows us to obtain a measure of satisfaction. The w-HEALTHQUAL scale, developed in this research, does incorporate this aspect, filling with that, the gap of the literature.

3. METHODS

3.1 Sample

We use a survey designed to explore patients' views with primary health care in the region of Extremadura[10]. This region is characterised for being very large in extension and sparsely populated[11]. Because of these particular characteristics, the primary health care system in the region is organised in two levels of aggregation: Health Areas and Health Zones. There are a total of eight Health Areas in the region: Badajoz (I), Mérida (II), Don Benito-Villanueva de la Serena (III), Llerena-Zafra (IV), Cáceres (V), Coria (VI), Navalmoral de la Mata (VII) and Plasencia (VIII). Each of them consists of different Health Zones (organised around a primary care centre as the main provider) in order to offer a more effective and operating primary health care[12].

The survey was delivered (between September and October 2008) to patients that had attended primary health care centres in Extremadura in May 2008. It was created in order to fulfil a series of requirements and organised in a series of set of questions.

Among the requirements, the survey was specifically designed to cover all the relevant dimensions of primary health care; to be reliable; to be a useful tool for

[10] The full survey is presented in Appendix 1.
[11] Specifically the extension of the region is 41,634.43 km^2, and in 2008 it had a population of 1,097,744 inhabitants, having a population density of 26.36 pop. per km^2.
[12] A map of the distribution of Health Zones across the region is shown in Appendix 2.

management activities that allows for the identification of those aspects of the service with which patients are more satisfied and, to be short, easily understood and easy to use by patients.

The organisation of the survey includes five blocks of questions: (1) background information about the patients, such as the health centre they go to, gender or age; (2) questions evaluating patients perceptions of the performance of the service in relation to health care facilities, health staff, non-health staff and efficiency elements; (3) questions asking patients about their level of overall satisfaction with the health service; (4) objectives measures such as time waited to be seen or number of years registered in the practice, and (5) questions asking patients to give an importance score to the same aspects of the service for which they have been asked to rate their experience, i.e. importance scores for the health care facilities, health staff, non-health staff and efficiency of the centre.

In order to validate the scale, surveys were piloted with a smaller sample consisting of thirty health care services users and a panel of ten experts in quality management and health economics. Experts were asked to examine the structure and the vocabulary employed in the development of the survey, resulting in the omission of three questions and rewording of some others.

Once we edited the questionnaires, we sent the surveys to a total of 20,271 primary health care users in the region, obtaining 2,402 completed questionnaires from 97 primary health care centres. The technical data of the study is presented in Table 1.

Table 1. Technical data of the study

POPULATION	People that have gone to a primary health care centre in the region of Extremadura during May 2008		
GEOGRAPHICAL AREA	Extremadura (Spain)		
SAMPLE SIZE	2,402 validated surveys from 97 Health Centres		
SAMPLE ERROR	1.24%		
CONFIDENCE LEVEL	95%	z= 1.96	p=q=0.5
FIELD WORK	September-October 2008		

Source: Prepared by authors

The demographic profile of respondents is presented in Table 2. The largest group of patients (37.81%) are aged 65 and over and the next largest group (28.05%) are aged

30-45. The majority of respondents refer to female patients (60.75% vs. 39.25% for males).

Table 2. Profile of surveyed users

Gender	Male	39.25%
	Female	60.75%
Age	< 30 years	9.67%
	30-45 years	28.05%
	45-64 years	24.46%
	≥ 65 years	37.81%

Source: Prepared by authors

3.2 Measurements

All variables are measured on seven-point Likert scales with 1 corresponding to *completely disagree* and 7 to *completely agree*. From all the aspects contained in the survey we utilise the questions asking patients to report their perception of the performance of the system based on their experiences about the facilities, health staff, non-health staff and efficiency of the centre (second block of questions) and the questions asking patients to rate how important these aspects of the services are for them (fifth block of questions) to build the w-HEALTHQUAL scale. All the questions contained in the scale refer to a total of twenty-five items, grouped in four categories of attributes (Table 3).

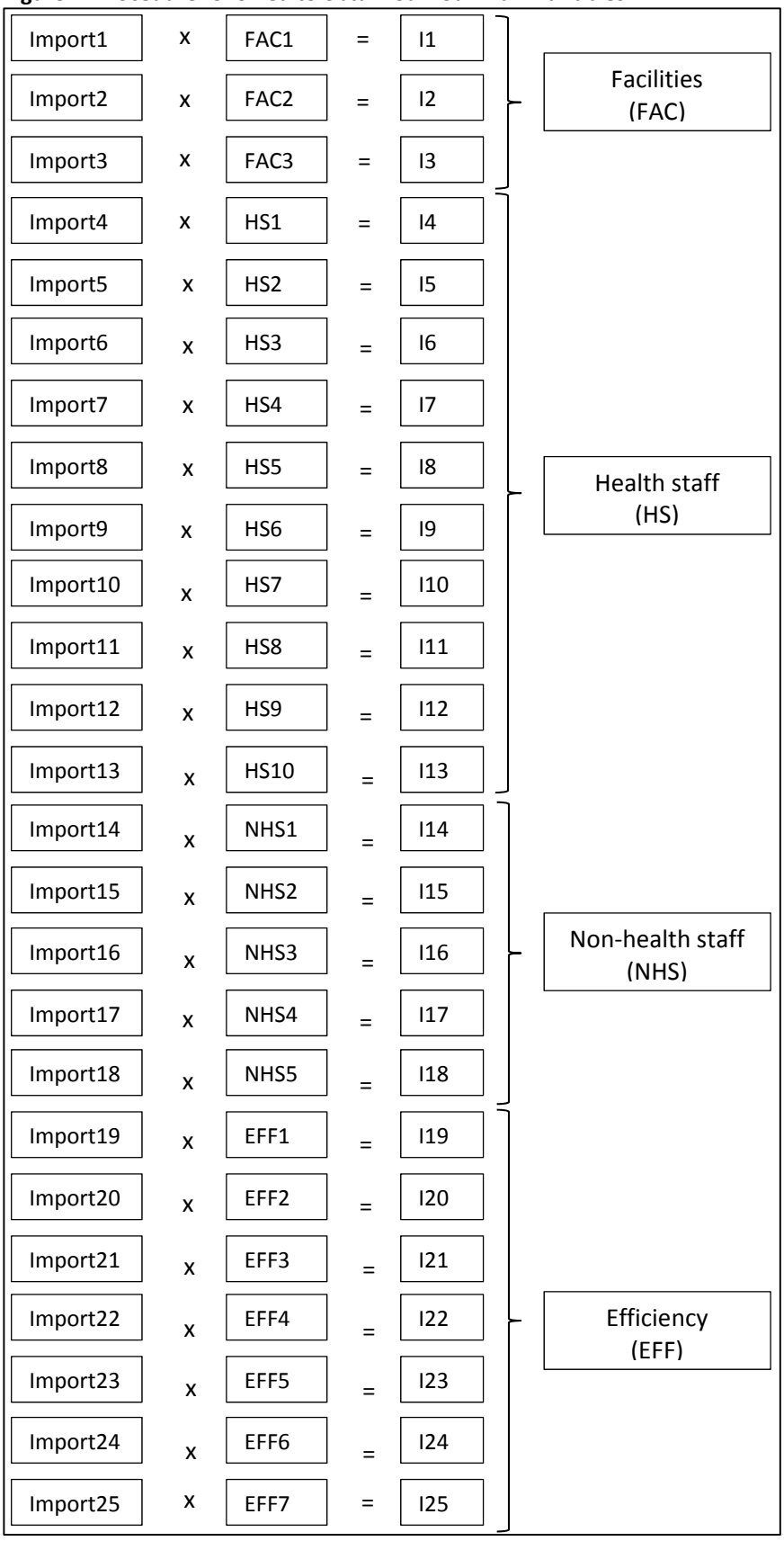
Table 3. Attributes measured by the w-HEALTHQUAL scale

Category	Definition
Facilities	1. - Cleanliness of facilities
	2.- Equipment of the health centre
	3.- Location of the health centre
Health Staff	4.- Health staff cleanliness
	5.- Health staff professionalism
	6.- Health staff kindness and politeness
	7.- Trust in health staff
	8.- Personalised service
	9.- Communication with health staff
	10.- Health staff attention to patients' problems
	11.- Health staff interest in solving patients' problems
	12.- Health staff understanding of patients' problems
	13.- Health staff prestige
Non-health staff	14.- Non-health staff cleanliness
	15.- Non-health staff professionalism
	16.- Non-health staff kindness and politeness
	17.- Non-health staff attention to patients' problems
	18.- Non-health staff interest in solving patients' problems
Efficiency	19.- Ease of making an appointment
	20.- Bureaucracy intensity
	21.- Waiting time in the health centre before entering the consulting room
	22.- Speed of complementary tests
	23.- Resolution of complaints
	24.- Time devoted to each patient
	25.- Health centre opening hours

Source: Prepared by authors

The satisfaction attributes that form the w-HEALTHQUAL, are built by multiplying the importance (Import1-Import25) and the performance (FAC1-FAC3, HS1-HS10, NHS1-NHS5, EFF1-EFF7) scores of each of the items (Fishbein and Ajzen, 1975; Paul, 2003; Jain and Gupta, 2004). In order to simplify the analysis by reducing the number of variables considered in it, the whole set of variables is reorganised in the following four main categories: facilities (FAC), health staff (HS), non-health staff (NHS) and efficiency (EFF). We build each of these four variables by calculating the average score of the items within every category of attributes, i.e. an average value of the three facility variables, a score for the ten variables describing characteristics of the health staff, the average value of the five items characterising aspects related to the non-health staff, and a mean score for the attributes showing patients' satisfaction with the efficiency of the system (Figure 1).

Figure 1. Procedure followed to obtain our four main variables



Source: Prepared by authors

The w-HEALTHQUAL adequately measures patients' satisfaction with the health care as it combines users' experiences with a series of elements of the system and the importance they give to those elements. It also improves previous measurement scales in different ways. First of all, it has been adapted to the research field for which is going to be used, eliminating with that any biases that the utilisation of other instruments (for example the SERVQUAL or the SERVPERF) developed for the study of satisfaction in other settings could introduce (Ladhari, 2009). Secondly, the w-HEALTHQUAL includes information about the importance that patients assign to the elements of the service being evaluated, building upon previous scales employed in health care (the HEALTHQUAL scale).

Furthermore, in relation to the Cronbach´s alpha of the scale, the literature indicates that this value has to be greater than 0.8 (Luque, 2000; Hair et al., 1999). The Cronbach´s alpha of the w-HEALTHQUAL is 0.97, supporting the reliability and internal consistency of the scale, and additionally, it is greater than the 0.96 coefficient obtained by Murillo-Zamorano et al. (2012) and Miranda et al., (2010) before the inclusion of the importance weights, improving with this the HEALTHQUAL scale.

4. RESULTS

The results of the study are shown in three stages. First, we present the main characteristics of our satisfaction indicators; it is, of the w-HEALTHQUAL scale. Second, we perform a cluster analysis to identify different groups that differ among them in their levels of satisfaction with the facilities, health staff, non-health staff and efficiency of primary health care. And third, we represent the different levels of satisfaction with the health care across the health map in the region to identify which actions are needed and where and to assess whether a different distribution of primary health care is necessary so that a better service can be delivered.

4.1 Characteristics of the w-HEALTHQUAL

As mentioned in the previous section, the w-HEALTHQUAL is built using patients performance and importance scores about a series of elements of the health care they receive. Table 4 shows the value of each of these indicators as well as the final satisfaction score for each of the twenty-five attributes initially considered and for the four final indicators of facilities (FAC), health staff (HS), non-health staff (NHS) and efficiency (EFF).

In general, aspects of the health care for which patients report their experience is worse are related to efficiency items (I19-I25), as indicated by their performance score. On the other side, elements of the health care for which patients have better experiences are related to facilities and health staff items. The non-health staff elements of the service being evaluated also receive high performance rating scores, but not as high as the other elements of the service.

All the aspects of primary health care studied in the research are seen as very important by patients with scores above 5 points (out of 7) in all the cases.

Among the satisfaction variables[13], the w-HEALTHQUAL, the aspect of the health service for which patients report the highest satisfaction is an item related to health staff – I4 (health staff cleanliness). In contrast, the items with the lowest satisfaction refer to efficiency items, being I19 (ease of making an appointment) the one for which patients assign the lowest scores, and therefore, being patients least satisfied with this aspect of health care.

By group of items (i.e. satisfaction by category), the indicators assessing the facilities of the centre are the ones for which patients have the highest satisfaction level, followed by the items related to staff, first health staff and then non-health staff, and finishing with the aspects of the health care related to efficiency, for which patients show the lowest satisfaction score.

[13] The satisfaction variables take values in a range from 1 to 49.

Table 4. The w-HEALTHQUAL. Performance, importance and satisfaction scores by item

Group of category	Item	Performance, mean (SD)	Importance score, mean (SD)	Satisfaction score, mean (SD)	Satisfaction score by category, mean (SD)
Facilities (FAC)	**I1**	5.74 (1.45)	6.17 (1.24)	36.11 (12.46)	33.43 (10.99)
	I2	5.01 (1.70)	5.94 (1.48)	30.59 (13.54)	
	I3	5.54 (1.80)	5.90 (1.36)	33.62 (14.12)	
Health staff (HS)	**I4**	6.02 (1.29)	6.27 (1.13)	38.41 (11.46)	32.79 (12.07)
	I5	5.43 (1.60)	6.20 (1.38)	34.63 (13.14)	
	I6	5.35 (1.71)	6.01 (1.44)	33.27 (13.77)	
	I7	5.16 (1.76)	5.73 (1.53)	30.69 (14.03)	
	I8	5.23 (1.78)	5.85 (1.54)	31.76 (14.22)	
	I9	5.39 (1.73)	5.89 (1.50)	32.75 (13.95)	
	I10	5.21 (1.81)	5.94 (1.53)	32.12 (14.35)	
	I11	5.16 (1.83)	5.95 (1.55)	31.99 (14.46)	
	I12	5.02 (1.83)	5.88 (1.54)	30.78 (14.36)	
	I13	5.16 (1.68)	5.88 (1.47)	31.48 (13.69)	
Non-health staff (NHS)	**I14**	5.73 (1.45)	5.97 (1.27)	35.03 (12.47)	29.47 (12.52)
	I15	5.02 (1.74)	5.76 (1.49)	29.99 (13.72)	
	I16	4.88 (1.84)	5.72 (1.57)	29.05 (14.19)	
	I17	4.70 (1.87)	5.46 (1.64)	26.90 (14.21)	
	I18	4.63 (1.91)	5.43 (1.68)	26.36 (14.37)	
Efficiency (EFF)	**I19**	3.51 (2.23)	5.25 (2.15)	20.03 (15.80)	24.02 (11.83)
	I20	4.31 (1.89)	5.42 (1.85)	24.45 (14.25)	
	I21	3.71 (1.98)	5.14 (2.09)	20.12 (14.01)	
	I22	4.00 (2.01)	5.46 (1.94)	23.03 (14.71)	
	I23	4.17 (1.90)	5.33 (1.91)	23.38 (14.15)	
	I24	4.74 (1.93)	5.65 (1.80)	27.99 (14.99)	
	I25	4.79 (1.95)	5.80 (1.52)	29.12 (14.61)	

Source: Prepared by authors

4.2 Cluster analysis

As mentioned in the introduction, to the best of our knowledge, there is not a study that analyses whether differences in the level of satisfaction with the health care within settings exits. In this research we fill that gap in the literature by means of a cluster analysis, looking at whether there are differences in patients' levels of satisfaction with primary health care in the Spanish region of Extremadura.

There are several techniques within the cluster methodology. In this piece of research we take a hybrid approach (also known as two-stage approach) (Wong, 1982; Yang et al., 2009; López-Sánchez and Santos-Vijande, 2015). We first employ the hierarchical method of Ward (Ward, 1963) and then proceed with non-hierarchical k-means clustering method (Lloyd, 1982; MacQueen, 1967), which builds on the solution obtained from the application of previous clustering methods, in our case, the method of Ward.

When applying the hierarchical method of Ward, we choose to analyse a wide range of solutions, between two and seven clusters, so we can have different ideas of how all the observations can be distributed into the different groups. Nevertheless, after looking at the composition of each of these clusters, we decide to perform the non-hierarchical k-means method with a range of solutions between two and five clusters. Using six or seven clusters does not seem to be a very sensible option as some of the clusters formed when applying the method of Ward show levels of satisfaction similar to the ones presented by other clusters, suggesting that the observations classified within these clusters should belong to the same group. By analysing the solutions resulting from the application of the k-means method, we consider a three-cluster solution as the most appropriate one, pointing to the existence of three groups of patients that differ according to their levels of satisfaction with the facilities, health staff, non-health staff and efficiency of the primary health care centre they normally attend.

We determine the stability of these results by means of a discriminant analysis (Greenly et al., 2005). The discriminant functions generate a significant value of the Wilks'

lambda in all the cases (Wilks' lambda = 0.148; p = 0.000). The discriminatory model correctly classifies 98.3% of the cases.

Patients' levels of satisfaction with our four main variables in each of the clusters are presented in Table 5. Cluster 1 is formed by patients whose levels of satisfaction with the health system are very high. For the four indicators used, the satisfaction scores are over 40 points, except for the efficiency ones, although they also present a high score. Therefore, we could categorise these users as "highly satisfied". Cluster 3 refers to patients whose levels of satisfaction are poor, as indicated by the mean values that the satisfaction variables in this group show. We can refer to this group of patients as "not satisfied". There is also a group of patients, cluster 2, comprising those individuals who are neither satisfied nor dissatisfied with the service they receive when attending a primary health care centre in the region. Patients in this cluster have satisfaction scores between 21 and 33 in all four indicators used in the analysis being, therefore, "not satisfied or dissatisfied".

Table 5. Patients' levels of satisfaction in each of the clusters*

Variables	Cluster 1 (N = 848)	Cluster 2 (N = 1042)	Cluster 3 (N = 512)
FAC [mean (SD)]	42.19 (6.31)	32.61 (7.73)	20.64 (9.34)
HS [mean (SD)]	43.85 (4.84)	32.07 (6.96)	15.92 (7.60)
NHS [mean (SD)]	41.76 (6.09)	27.33 (7.30)	13.46 (7.01)
EFF [mean (SD)]	35.74 (7.59)	21.02 (6.88)	10.70 (6.42)

Source: Prepared by authors

*Cluster 1: *"satisfied patients"*; cluster 2: *"patients not satisfied or dissatisfied"*; cluster 3: *"not satisfied patients"*

In the next section, we employ the information provided by the cluster analysis to identify whether there are differences between the way primary health care is administratively organised and how it should be distributed so that every patient reports high levels of satisfaction, i.e. to identify whether a different distribution of primary health care can be draw in the region of Extremadura in order to improve the health service finally delivered.

4.3 Geographical representation of the levels of satisfaction with the health care

Once we have our patients classified in one of the three clusters, we proceed to categorise each of the Health Zones as belonging to one of the three groups, so the

distribution of patients' levels of satisfaction can be drawn in the health map of the region of Extremadura. By doing this, we can identify whether there are differences between the administrative classification of Health Zones and the classification according to patients' needs.

The procedure we follow consists of analysing, for every Health Zone, the distribution of their patients in each of the clusters, so that, we can characterise the Health Zones according to the cluster that the majority of their patients belong to. For example if a Health Zone receives a total of seventy completed questionnaires and, according to their responses, fifty patients are classified in cluster 1, fifteen patients in cluster 2 and the remaining five patients in cluster 3, the Health Zone is characterised as belonging to cluster 1. After following this procedure we obtain that out of the 97 primary health care centres participating in the study, 35 can be classified in cluster 1, 52 are included in cluster 2, and the remaining 10 in cluster 3.

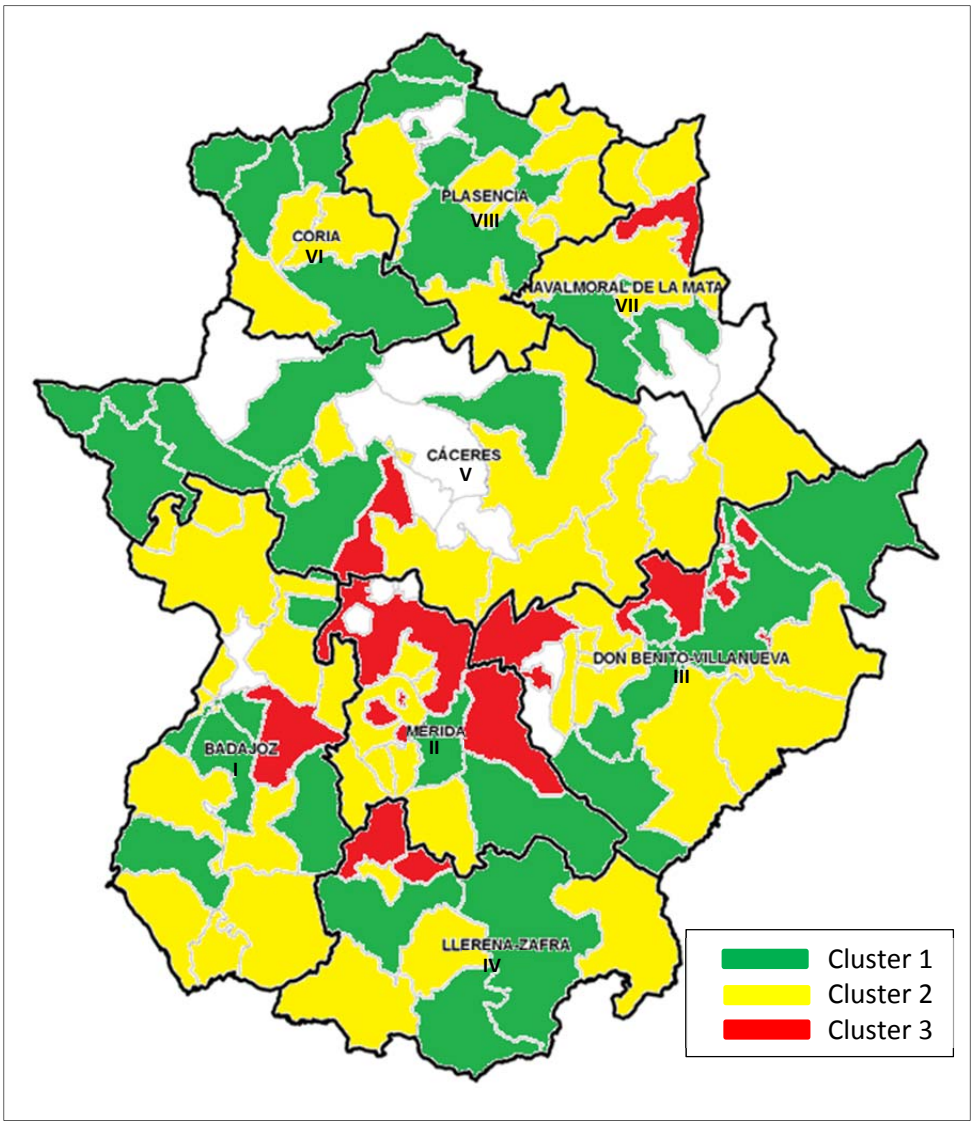
We represent the characterisation of each of the Health Zones across the map of primary health centres in the region in Figure 2. In this map, we can observe the delimitation of the eight Health Areas as well as the disaggregation of them in each of the Health Zones. We assign a colour to each of the Health Zones in the map to indicate the cluster where they belong following the procedure previously described[14].

All the three different clusters are present in six of the eight Health Areas. In Health Areas VI and VIII none of the Health Zones are catalogued in cluster 3, meaning that these two Health Areas do not assist very unsatisfied patients (the number of Health Zones in each of the cluster by Health Areas is presented in Table 6). These two Health Areas, VI and VIII, are also the ones with a higher proportion of Health Zones in cluster 1, i.e. Health Zones where patients are, in general, satisfied with the health care they receive. However, in the case of Health Area VIII, the majority of Health Zones are categorised as belonging to cluster 2, which means that some kind of improvement is needed in order to increase patients' satisfaction with the health care.

[14] Health Zones that appear in white, refer to those ones for which we do not receive any surveys and that, therefore, are not included in the analysis.

On the other hand, Health Area II is the area with the lowest proportion of Health Zones where patients are satisfied with the attention received; only two out of the twelve Health Zones present in the study are assisting satisfied patients.

Figure 2. Classification of Health Zones according to patients' levels of satisfaction



Source: Prepared by authors

Most of the Health Areas have a higher proportion of Health Zones where patients are not satisfied of dissatisfied – cluster 2. Moreover, further actions are needed to improve patients' satisfaction in some of the Health Areas that present Health Zones where patients are unsatisfied with the health care received. This is especially important in Health Areas II and IV, with around a quarter of the Health Zones being in cluster 3.

Table 6. Number of Health Zones in each of the clusters by Health Area*

Health Area	Cluster 1		Cluster 2		Cluster 3		Total	
	n	%	n	%	n	%	n	%
I	7	35.00	12	60.00	1	5.00	20	100
II	2	16.67	7	58.33	3	25.00	12	100
III	5	38.46	6	46.15	2	15.39	13	100
IV	3	33.33	4	44.45	2	22.22	9	100
V	6	35.29	10	58.82	1	5.89	17	100
VI	4	57.14	3	42.86	-	-	7	100
VII	6	46.15	7	53.85	-	-	13	100
VIII	2	33.33	3	50.00	1	16.67	6	100
Total	**35**	**36.08**	**52**	**53.61**	**10**	**10.31**	**97**	**100**

Source: Prepared by authors

*Cluster 1: *"satisfied patients"*; cluster 2: *"patients not satisfied or dissatisfied"*; cluster 3: *"not satisfied patients"*

Previous results highlight the idea that, not only within the region, but even within every Health Area, health care is unevenly distributed resulting in differences in the levels of satisfaction with the system reported by patients. Given that the health system is treated different patients, different policies are required so that different needs can be fulfilled adequately.

As previously mentioned, currently, in the context of the distribution of the health care in Extremadura, when a decision is made it affects the whole region or a particular Health Area, but it does not consider whether different requirements may be needed within every of these administrative units. Therefore, the key message for policy makers is the need to consider that a new organisation of the health care based on patients' needs rather than on an administrative distribution may be needed, so that, health care is delivered according to what patients really require. This, in the last instance, could contribute to an increase in patients' levels of satisfaction with the health care.

5. DISCUSSION AND CONCLUSION

Our research has developed a new measure of patient satisfaction with primary health care, the w-HEALTHQUAL, so that it can be studied with an accurate and non-biased indicator. The analysis has been performed with data provided by patients attending primary health care in the Spanish Autonomous Community of Extremadura. This approach has allowed us to identify whether a different distribution of primary

health care across the region can be drawn so that health care can be delivered according to patients' needs rather than based on where they live and the centre they go to, and consequently, being able to improve the quality with which service is delivered.

We have used a survey asking patients to report their experiences with a series of health attributes and the importance they assign to each of them to build the w-HEALTHQUAL scale, which allows us to construct a measure of satisfaction with primary health care. The fact that we consider patients' views is very important given that as mentioned by Crow et al., (2002), patients provide truthful data and are an essential aspect to take into consideration so that an improved health care can be delivered. The survey was originated on the basis of the HEALTHQUAL scale (Murillo-Zamorano et al., 2012; Miranda et al., 2010), and it consists of a measure specifically designed for the study of patients preferences with primary health care, and including the importance that users of the service attribute to the each of the elements being analysed. Previous literature has used existing scales in other fields without performing any adaptation of the instruments, and has concluded that, they were able to identify deficiencies and offer a starting point for the identification of underlying problems that may be interfering on the quality of the service being delivered (Babakus and Mangold, 1992; Yang et al., 2004). Therefore, the w-HEALTHQUAL, that improves existing measurement scales in terms of adaptation and reliability, will be a better reflection of patients' views and their needs.

The survey has been validated, reducing with that any source of misunderstanding it could generate and, therefore, producing accurate information from patients responding to it. The w-HEALTHQUAL, therefore and considering all the aforementioned, adequately measures patients' satisfaction with the health service and helps the system to identify shortcomings in the provision of health care. This, in the last instance, will allow for the delivery of a better health and the implementation of health policies adapted to patients' needs, which could be translated in a continuous improvement of service users' satisfaction.

Using the w-HEALTHQUAL, we have obtained four indicators related to four main aspects of the service: facilities, health staff, non-health staff and efficiency. These

variables have been used to perform a hybrid cluster analysis, first employing the hierarchical method of Ward (Ward, 1963) and then applying the k-means clustering method (Lloyd, 1982; MacQueen, 1967). The cluster analysis has allowed us to identify different satisfaction areas in the map of the distribution of Health Zones, so that, it can be reorganised based on patients' levels of satisfaction with primary care rather than considering an administrative classification of the health care.

Results have indicated the existence of three main clusters. Patients with high levels of satisfaction with the four indicators have been classified in cluster 1 – "highly satisfied" patients. Those service users whose satisfaction with the health care received was poor have been categorised as "not satisfied" and have been grouped in the third cluster. Finally, there has been an intermediate group – cluster 2 – including patients that were neither satisfied nor disappointed with the system, so that they could be described as "not satisfied or dissatisfied".

Using this classification of patients in one of the three clusters, we have been able to categorise each of the 97 Health Zones participating in the study in one of the levels of satisfaction reported by patients attending their primary health care centres. According to this, 35 Health Zones could be classified in cluster 1; 52 in cluster 2 and the remaining 10, in cluster 3. Therefore, a majority of patients were not satisfied or dissatisfied with the health care received, and around 10% were dissatisfied with it. It is of concern that only 36% of the centres (35 centres out of the 97) were treating highly satisfied patients. Furthermore, within each of the clusters, the levels of satisfaction with the efficiency aspects of the system have been normally lower than the ones shown by the other variables being studied. These results have highlighted the need of an improvement in the provision of health care so that, a high quality health care service can be delivered to patients, a better quality service that, can increase patients' satisfaction with the system.

The results have also emphasised the idea that a different distribution of the health care across the region may be needed so that a more effective health care can be delivered. With that, policies could be implemented based on patient satisfaction with

the service received rather than on an administrative distribution of health units, which is not likely to benefit every user of the system.

These implications are especially relevant in a large and sparsely populated region like the one analysed here, as in such context, an optimum organisation of health care is essential so that needs are correctly identified and resources are allocated where they are really required. Considering this, the development and implementation of policies should take into consideration what patients have stated to need an improvement so that deficiencies are corrected. Additionally, managers should continue to offer a service with the same standards in relation to the aspects with which patients are satisfied. All of this should be elaborated bearing in mind where these strategies are more needed and in which direction.

REFERENCES

Adil, D., Al Ghaswyneh, O. F. M., and Albkour, A. M. (2013). SERVQUAL and SERVPERF: A review of measures in services marketing research. *Global Journal of Management And Business Research*, 13 (6).

Babakus, E. and Mangold, W. G. (1991). Adapting the SERVQUAL scale to Hospital services: An Empirical Investigation. *Health Service Research*, 26 (6), pp. 767-786.

Badri, M. A., Abdulla, M. and Al-Madani, A. (2005). Information technology center service quality: assessment and application of SERVQUAL. *International Journal of Quality & Reliability Management*, 22 (8/9), pp. 819-848.

Berry, L. L., Zeithaml, V. A. and Parasuraman, A. (1985). Quality Counts in Services too. *Business Horizons*, 28 (3), pp. 44-52.

Bitner, M. J., Booms, B. H. and Tetreault, M. S. (1990). The service encounter: diagnosis favourable and unfavourable incidents. *Journal of Marketing*, 54, pp. 71-84.

Brown, T. J., Churchil, G. A. and Peter, J. P. (1993). Improving the measurement of service quality. *Journal of Retailing*, 69 (81), pp. 127-139.

Churchill, G. A. and Suprenant C. (1982). An investigation into the determinants of customer satisfaction. *Journal of Marketing Research*, 19, pp. 491-504.

Cronin, J. and Taylor, S. (1992). Measuring service quality: a reexamination and extension. *Journal of Marketing*, 56, pp. 55-68.

Crow, R., Gage; H., Hampson, S., Hart, J., Kimber, A., Storey, L. and Thomas, H. (2002). The measurement of satisfaction with healthcare: implications for practice from a systematic review of the literature. *Health Technology Assessment*, 6 (32), pp. 1-244.

Dawson, J., Fitzpatrick, R., Murray, D. and Carr, A. (1998). Questionnaire on the perceptions of patients about total knee replacement. *Journal of Bone & Joint Surgery-British Volume*, 80 (1), pp. 63-69.

Donabedian, A. (1984). La Calidad de la Atención Médica. *La Prensa Mexicana*, México.

Fishbein, M. and Ajzen, I. (1975). Belief, attitude, intention and behavior. An introduction to theory and research. *Reading Mass: Addison-Wesley*.

Greenly, G. E., Hooley, G. J. and Rudd, J. M. (2005). Market orientation in a multiple stakeholder orientation context: implications for marketing capabilities and assets. Journal of Business Research, 58 (11), pp. 1483-1494.

Hair, J., Yerson, R., Tatham, R. and Black, W. (1999). Análisis Multivariante. Pearson Educacion, Madrid.

Jain, S. K. and Gupta, G. (2004). Measuring Service Quality: SERVQUAL vs. SERVPERF scales. *Vikalpa*, 29 (2), pp. 25-37.

Kilbourne, W. E., Duffy, J. A., Duffy, M. and Giarchi, G. (2004). The applicability of SERVQUAL in cross-national measurements of health-care quality. *Journal of Services Marketing*, 18 (6/7), pp. 524-533.

Ladhari, R. (2009). A review of twenty years of SERVQUAL research. *International Journal of Quality and Service Sciences*, 1 (2), pp. 172-198.

LeBlanc, G. and Nguyen, N. (1988). Customers satisfaction with service quality in financial institutions. *International Journal of Bank Marketing*, 6 (4), pp. 7-18.

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

López-Sánchez, J. A. and Santos-Vijande, M. L. (2015). How value creation and relationship quality coalignment affects a firms's performance: An empirical analysis. *Journal of Marketing Channels*, 22 (3), pp. 214-230.

Luque, T. (2000). "Técnicas de Análisis de Datos en Investigación de Mercados". Madrid: Ed. Pirámide.

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, pp. 281–297.

Martilla, J. A. and James, J. C. (1977). Importance-Performance analysis. *Journal of Marketing*, 41, pp. 77-79.

Mazis, M. B., Antola, O. T. and Klippel, T. E. (1975). A comparison of four multi-attribute models in the prediction of consumer attitudes. *Journal of Consumer Research*, 2 (1), pp. 35-52.

Miranda, F. J., Chamorro, A., Murillo, L. R. and Vega, J. (2010). Assessing primary health care services quality in Spain: managers vs. Patients perceptions". *The Service Industries Journal*, 30 (13), pp. 2137-2149.

Murillo-Zamorano, L. R., Vega-Cervera, J. A., Chamorro, A., Cordero, J. M., Crespo, E., de Miguel, F. J., Miranda, F. J., Nogales, L., and Sánchez, M. (2012). Eficiencia y Calidad de la atención primaria en Extremadura. Murillo and Vega eds. ISBN:978-84-614-7542-1.

Parasuraman, A., Zeithaml V., and Berry L. (1985). A Conceptual Model of Service Quality and its Implications for Future Research. *Journal of Marketing*, 49, pp. 35-48.

Parasuraman, A., Zeithaml V., and Berry L. (1988). "SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality". *Journal of Retailing*, 64 (1), pp. 12-40.

Paul, D. P. (2003). An exploratory examination of SERVQUAL versus SERVPERF for Prosthectic Dental Specialists. *Clinical Research and Regulatory Affairs*, 20 (1), pp. 89-100.

Peter, J. P., Churchil, G. A. and Brown, T. L. (1993). Caution in the use of difference scores in Consumer Research. *Journal of Consumer Research*, 19 (4), pp. 655-662.

Schäfer, W. L., Boerma, W. G. W., Murante, A. M., Sixma, J. J. M., Schellevis, F. G., Groenewegen, P. P. (2015). Assessing the potential for improvement in primary care in 34

countries: a cross-sectional survey. *Bulletin of the World Health Organization*, 93 (3), pp. 161-168.

Ward, J. H. Jr. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58 (301), pp. 236-244.

Wong, M. A. (1982). A hybrid clustering method for identifying high-density clusters. Journal of the American Statistical Association, 77 (380), pp. 841-847.

Woodruff, R. B., Cadotte, E. R. and Jenkins, R. L. (1983). Modelling consumer satisfaction processes using experience-based norms. *Journal of Marketing Research*, 20, pp. 296-304.

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

Yang, Z; Jun, M, and Peterson, R. T. (2004). Measuring customer perceived online service quality: scale development and managerial implications. *International Journal of Operations & Production Management*, 24 (11), pp. 1149-1174.

APPENDICES

Appendix 1. Satisfaction survey delivered to primary health care users in Extremadura

JUNTA DE EXTREMADURA
Consejería de Sanidad y Dependencia

UNIVERSIDAD DE EXTREMADURA

QUALITY ANALYSIS OF PRIMARY HEALTH CARE IN EXTREMADURA

Dear user of primary health care service in Extremadura,

Thanks for devoting the time to answer this questionnaire. The results obtained from it will allow us to have a better understanding of the way primary health care in the region is being delivered, so that, a better quality service can be provided.

All the information provided here is strictly confidential and it will be analysed together with the rest of questionnaires received from other primary health care users.

Instructions: Circle the option which best describes your experience when attending your primary health care centre. Hand this survey in to the reception of your corresponding centre. If you wish, you can fill the survey online following this link: http://mercado.unex.es/salud.

Thanks in advance for taking part in the survey

GENERAL INFORMATION
1.- Primary Health Centre:
2.- Gender: ☐ Male ☐ Female
3.- Age: ☐ < 30 years old ☐ 30-45 years old ☐ 45-64 years old ☐ ≥ 65 years old
5.- Family income: ☐ < 900€/m ☐ 900-1800 €/m ☐ 1800-2700 €/m ☐ ≥ 2700 €/m
6.- Education: ☐ no studies ☐ primary school ☐ secondary school ☐ university
7.- Do you live in the same town where the primary health care is located? ☐ Yes ☐ No
8.- How long does it take you to travel to the centre?:

HEALTH CENTRE FACILITIES							
Please, indicate your level of agreement with the following statements regarding your primary health care centre (where 1 means completely disagree and 7 means completely agree)							
In your opinion:	Completely disagree						Completely agree
1.- Health centre facilities are clean	1	2	3	4	5	6	7
2.- Health centre equipment is modern	1	2	3	4	5	6	7
3.- According to your place of residence, health centre location is adequate and easy to access	1	2	3	4	5	6	7

HEALTH STAFF (GENERAL PRACTITIONERS AND NURSES)							
In your opinion:	Completely disagree						Completely agree
1.- Health staff are well dressed and clean	1	2	3	4	5	6	7
2.- Health staff know well what they do (they are professionals)	1	2	3	4	5	6	7
3.- Health staff are kind and polite	1	2	3	4	5	6	7
4.- Health staff inspire trust	1	2	3	4	5	6	7
5.- I have received personalised care	1	2	3	4	5	6	7
6.- I have been well informed about the nature and objectives of the recommended treatment	1	2	3	4	5	6	7
7.- Health staff listen carefully to my problems	1	2	3	4	5	6	7
8.- Health staff have shown an interest in solving my problem	1	2	3	4	5	6	7
9.- Health staff have dealt with my problems/needs very quickly	1	2	3	4	5	6	7
10. Health staff prestige is high	1	2	3	4	5	6	7

NON-HEALTH STAFF (ADMIN STAFF, CENTRE ATTENDANTS, ETC.)							
In your opinion:	Completely disagree						Completely agree
1.- Non-health staff are well dressed and clean	1	2	3	4	5	6	7
2.- Non-health staff know well what they do (they are professionals)	1	2	3	4	5	6	7
3.- Non-health staff are kind and polite	1	2	3	4	5	6	7
4.- Non-health staff listen carefully to my problems	1	2	3	4	5	6	7
5.- Non-health staff have shown an interest in solving my problems	1	2	3	4	5	6	7

EFFICIENCY MEASURES							
In your opinion:	Completely disagree						Completely agree
1.- It is easy to contact the centre to make an appointment	1	2	3	4	5	6	7
2.- The level of bureaucracy is reasonable	1	2	3	4	5	6	7
3.- Once I arrived at the centre, waiting time to be seen was reasonable	1	2	3	4	5	6	7
4.- Complementary tests (x-ray, tests, etc.) are performed promptly	1	2	3	4	5	6	7
5.- When there has been a complaint, health centre response has been adequate	1	2	3	4	5	6	7
6.- Time I have been given has been adequate	1	2	3	4	5	6	7
7.- Opening times have been appropriate	1	2	3	4	5	6	7

OVERALL SATISFACTION							
	Completely disagree						Completely agree
1.- In general, I am very satisfied with health centre services	1	2	3	4	5	6	7
2.- Service received fulfils what I expected from it	1	2	3	4	5	6	7
3.- Health centre reputation is really good	1	2	3	4	5	6	7

OBJECTIVE MEASURES
1.- Time waited to be seen: ☐ < 10 min ☐ 10-19 min ☐ 20-29 min ☐ ≥ 30 min
2.- Years attending this particular health centre: ☐0-1 yr ☐ 2-5 yrs ☐6-10 years ☐ ≥ 10 yrs
3.- Times I have visited the centre in the last 12 months: ☐ 1-2 ☐ 3-4 ☐ 5-6 ☐ > 6 times

IMPORTANCE LEVEL FOR EACH OF THE ATTRIBUTES							
Please, state how important the following aspects of service quality are for you							
In your opinion:	Not very important						Very important
1.- Cleanliness of facilities	1	2	3	4	5	6	7
2.- Equipment of the health centre	1	2	3	4	5	6	7
3.- Location of the health centre	1	2	3	4	5	6	7
4.- Health staff cleanliness	1	2	3	4	5	6	7
5.- Health staff professionalism	1	2	3	4	5	6	7
6.- Health staff kindness and politeness	1	2	3	4	5	6	7
7.- Trust in health staff	1	2	3	4	5	6	7
8.- Personalised services	1	2	3	4	5	6	7
9.- Communication with health staff	1	2	3	4	5	6	7
10.- Health staff attention to patients' problems	1	2	3	4	5	6	7
11.- Health staff interest in solving patients' problems	1	2	3	4	5	6	7
12.- Health staff understanding of patients' problems	1	2	3	4	5	6	7
13.- Health staff prestige	1	2	3	4	5	6	7
14.- Non-health staff cleanliness	1	2	3	4	5	6	7
15.- Non-health staff professionalism	1	2	3	4	5	6	7
16.- Non-health staff kindness and politeness	1	2	3	4	5	6	7
17.- Non-health staff attention to patients' problems	1	2	3	4	5	6	7
18.- Non-health staff interest in solving patients' problems	1	2	3	4	5	6	7
19.- Ease of making an appointment	1	2	3	4	5	6	7
20.- Bureaucracy intensity	1	2	3	4	5	6	7
21.- Waiting time in the centre before entering the consultation room	1	2	3	4	5	6	7
22.- Speed of complementary tests	1	2	3	4	5	6	7
23.- Resolution of complaints	1	2	3	4	5	6	7
24.- Time spent with each patient	1	2	3	4	5	6	7
25.- Health centre opening hours	1	2	3	4	5	6	7

Appendix 2. Map of the distribution of Health Areas across the region of Extremadura and distribution of patients in each of them

Health Area	Number of Health Zones	Population, N (%)
I	21	270,317 (24.62)
II	13	165,750 (15.10)
III	14	142,448 (12.98)
IV	9	106,731 (9.72)
V	23	196,411 (17.89)
VI	7	49,530 (4.51)
VII	14	111,927 (10.20)
VIII	8	54,630 (4.98)
Total	**109**	**1,097,744 (100)**

Source: Prepared by authors

CHAPTER 3: EXPLAINING PATIENTS' SATISFACTION WITH PRIMARY HEALTH CARE: A MULTILEVEL ANALYSIS

1. INTRODUCTION

Primary health care is the first contact point with the health care system for patients. Therefore, it is essential to offer a high quality service, a quality that needs to be recognised as high quality by both, clinicians and patients. Service quality has been recognised as one of the values of central importance in organisations, regardless whether we study manufacturing or service sectors (Berry et al., 1985; Bitner et al., 1990; LeBlanc and Nguyen, 1988).

Although quality in the health care field has been considered as an objective measure, some authors indicate that we cannot fully describe quality if the user's point of view is not taken into consideration (Donabedian, 1984), as patients' provide reliable data about their health and about how they feel after seeing their doctor (Dawson et al., 1998).

Additionally, "service quality is an antecedent of consumer satisfaction" (Cronin and Taylor, 1992), hence, attention needs to be focused on offering a high quality service as perceived by the patient, which will ultimately be translated into a greater satisfaction. When measuring patients' satisfaction with a particular service, researchers normally ask users to estate, in a given scale, how satisfied they are with that service. However, previous literature has indicated that the importance that they give to the elements of the service being considered, will influence their levels of satisfaction (Martilla and James, 1997; Jain and Gupta, 2004), consequently, a combination of patients' experiences together with the importance of all the aspects for which it is being measured, will be an appropriate way to build a measure of satisfaction.

When studying satisfaction with the health care, it is also important to take into consideration that the levels of satisfaction may differ from one patient to another and across a particular health system or region. Differences in patients' satisfaction with the health care may be driven by individual characteristics as well as features of the centre patients attend or the General Practitioner (GP) they see. Not only patients' sociodemographic characteristics, such as gender or age, may be influencing their views

about a particular service, but also the context in which the care occurs has a fundamental impact on understanding differences in satisfaction and on the policies that need to be implemented (Duncan et al., 1993). In consequence, if we want to offer a high quality service with which patients are highly satisfied, we need to determine why satisfaction may be unevenly distributed, i.e. what most influences satisfaction and where actions are more needed.

Considering these ideas, the aim of the present study is to identify to what extent differences in patients' level of satisfaction, obtained as a combination of experiences and importance, are related to individual characteristics as well as to characteristics at the providers level, not only including GPs but also other aspects of the health care centres.

In doing so, we run a series of multilevel analyses in which we include characteristics at various levels, the individual patient level and the centre level. This type of model is normally used in the analysis of data that are naturally organised in hierarchical structures (Rice and Jones, 1997), for example patients nested within GPs.

There is a considerable amount of research in different fields applying multilevel analyses and using a wide range of variables. In the health care field, a lot of this research has been focused on the analysis of hospitals rather than primary health care centres and has, for example, explained variations in costs (Laudicella et al., 2010; Street et al., 2012; Dormont and Milcent, 2004) or in length of stay (Street et al., 2010). Some multilevel models have also been applied to primary health care although this has been more reduced than in the case of hospital care. For instance, we can mention Schäfer et al. (2015) analysing the potential for improvement in a series of countries or Pullicino et al. (2016) focussing on the influence of patients characteristics on health care-seeking behaviour.

In regard to the study of patients reported satisfaction with primary health care, many pieces of research look at what patients' characteristics most influence this variable. However, few of them use a multilevel approach in which patients' characteristics together with aggregate level data are combined (Salisbury et al., 2010; Sixma et al., 1998).

Among the patients' characteristics that influence satisfaction, previous research has looked at variables such as age, gender, education or morbidity status, among others. Higher levels of satisfaction are often reported by older, compared to younger, patients (Campbell et al., 2001; Sixma et al., 1998; Millar, 2001; Hall and Dorman, 1990), although some exceptions have also been found (Baker, 1996). With reference to gender, some authors have found a relationship between gender and satisfaction, with women reporting lower levels than men (Millar, 2001), while others, have not found statistically significant differences when comparing satisfaction levels by gender (Campbell et al., 2001). When looking at levels of education, those with a lower level of qualification often report a higher satisfaction than those who have further studies (Millar, 2001; Hall and Dorman, 1990; Lewis, 1994). Other research has also looked at how satisfaction varies in relation to health status. Sixma et al. (1998) found that those considering themselves as being in good health are more satisfied than those in poor health.

Regarding the aggregate level variables, previous literature has considered aspects such as the characteristics of the GP, the practice size or the experience of doctors. Sixma et al., (1998) run a multilevel analysis to identify which characteristics may be influencing satisfaction with the accessibility and availability of care, the patients' perception of the humaneness of the GP and the patients' perception of the amount of information given by the GP. For their second level, they include characteristics of the GP and the practices and out of the sixteen variables included, they find that only getting an appointment for the same day is significantly associated with their dependent variables. Salisbury et al. (2010) also explore whether patients' experience and satisfaction are driven by individual and aggregate features. Among the aggregate variables, they include practice level and GP characteristics, finding that, for overall satisfaction the training practices treat more satisfied patients and that if the GP is male or qualified in the UK, patients' satisfaction is also higher.

Our research extends the existing literature in several ways. First of all, we study patients' satisfaction with primary health care not only focussing on the overall level of satisfaction or the satisfaction with the GP, but also considering specific elements that

build the primary health care system. Secondly, our measure of satisfaction gathers more information than a general question asking patients to report their general satisfaction, as it is obtained as a combination of experiences with a series of elements of the primary care and the importance patients attain to each them. Finally, it extends the existing multilevel literature that studies the source of variation in patients' satisfaction with primary health care, which, as mentioned above, has been very scarce.

To the best of our knowledge, there is not any piece of research looking at how satisfaction with primary health care and with the elements of it, obtained as a combination of patients' experiences and importance judgements varies with patients' and centre characteristics. This extension of the literature contributes to a better understanding of the health service received that, in the last instance, will lead to the definition of adequate health policies that not only secure the delivery of high quality services but also its recognition as such by patients, with the subsequently reflection in their levels of satisfaction.

In order to achieve all the previously mentioned, we structure the study as follows. First, we describe the main characteristics of the multilevel models as well as the advantages of this type of methodology. Next, we present the dataset used in the analysis. Specifically, we use two datasets, the first of them containing information at the patient level, and the second one comprising information about the primary health care centres where those patients receive care. Following this section the results of the multilevel analysis are displayed, taking an additive approach, firstly, estimating the null models, secondly, adding variables at the patient level and thirdly, incorporating the centre level variables. Finally, we present the discussion and conclusions of the research.

2. METHODS

In the research presented in this chapter, we analyse whether satisfaction with primary health care is driven by characteristics at the patient level or at the level of the centre they go to. In order to analyse so, we run a series of multilevel analyses. Multilevel analyses, also known as two-level multilevel models, are commonly used when data are

naturally organised in hierarchies (Rice and Jones, 1997). The goal of this type of analysis is to predict the value of some dependent variables in terms of both, individual and environmental or aggregate variables (Luke, 2004; Von Korff et al., 1992). As indicated by Diez-Roux (2008) the incorporation of group level variables will provide essential information that it is not taken into account by individual level data. Some of the advantages of this kind of models is that it presents changes in the pattern of a particular variable even if the data has been collected at different points in time (Luke, 2004); the coefficients and standard errors estimated from the application of a multilevel model are more accurate than the obtained from a linear model, and the explained variance can be divided into the different levels being considered (Sixma et al., 1998).

The procedure commonly used when applying a multilevel model consists of an additive approach. First of all, a model without any exogenous variable is estimated, so that the explained variance can be decomposed in the different levels included in the model. The general specification of the null model is shown in equation 1.

$$Y_{ij} = \beta_0 + \mu_j + \varepsilon_{ij} \quad [1]$$

$$\textit{where } \mu_j \sim N(0, \sigma^2);\ \varepsilon_{ij} \sim N(0, \tau^2)$$

After that, explanatory variables are incorporated into the model, first, at the individual level, in our case at the patient level (equation 2), and then at the aggregate level, in our analysis including characteristics of the centres (equation 3).

$$Y_{ij} = \beta_0 + \beta_1 X_{ij} + \mu_j + \varepsilon_{ij} \quad [2]$$

$$\textit{where } \mu_j \sim N(0, \sigma^2);\ \varepsilon_{ij} \sim N(0, \tau^2)$$

$$Y_{ij} = \beta_0 + \beta_1 X_{ij} + \beta_2 Z_j + \mu_j + \varepsilon_{ij} \quad [3]$$

$$\textit{where } \mu_j \sim N(0, \sigma^2);\ \varepsilon_{ij} \sim N(0, \tau^2)$$

In the previous equations the subscript *i* refers to the individual patients and *j* to the centre; *Y* is the dependent variable; β_0 is the intercept and represents the average satisfaction score that we would expect from a patient randomly selected from the pool

of centres participating in the study; X_{ij} includes the variables at the patient level; Z_j refers to the variables at the centre level; $\mu_{0j} + \varepsilon_{ij}$ represent the random components for the patient level and the centre level, where ε_{ij} is the variance due to between-patient differences and μ_{0j} is the variance due to between-centre differences.

3. DATA

We use two different datasets, one with information about patients and the other one containing data at the level of the health centre that patients attend when seeking health care. All this information refers to primary health care in the region of Extremadura, an Autonomous Community located in the South West of Spain. The region is characterised for being very large in extension and sparsely populated[15]. Because of these particular characteristics, the primary health care system in the region is organised in two levels of aggregation: Health Areas and Health Zones, with each of the Health Areas containing several Health Zones, and the Health Zones organised around a primary health care centre as the main provider.

For the patient level variables, we use a satisfaction questionnaire (presented in Appendix 1) specifically designed to explore patients' views with primary health care in the region. The questionnaire was delivered (between September and October 2008) to patients that had attended primary health care centres in Extremadura in May 2008. It is organised in five blocks of questions: (1) background information about the patients, such as the health centre they go to, gender or age; (2) questions evaluating patients perceptions of the performance of the service in relation to health care facilities, health staff, non-health staff and efficiency elements; (3) questions asking patients about their level of overall satisfaction with the health service; (4) objectives measures such as time waited to be seen or number of years registered in the practice, and (5) questions asking patients to give an importance score to the same aspects of the service for which they have been asked to rate their experience, i.e. importance scores for the health care

[15] Specifically the extension of the region is 41,634.43 km^2, and in 2008 it had a population of 1,097,744 inhabitants, having a population density of 26.36 pop. per km^2.

facilities, health staff, non-health staff and efficiency of the centre. We sent the questionnaire to a total of 20,271 primary health care users in the region, obtaining 2,402 completed versions from 97 primary health care centres.

For the centre level variables, we use data from a series of reports, called APEX[16], that gather information about primary care in Extremadura and that have been published from the year 2006 to 2009. In the analysis here presented, we use the information from one of those reports, specifically the one related to 2008 – APEX08 (Murillo-Zamorano et al., 2011), as it contains the records for the 109 health centres existing in that year, for which the data from the patients perspective was also collected. This dataset presents the particular characteristic that it allows for the identification of pattern behaviours related to a high or low quality and efficiency in the delivery of primary health care services. Apart from that, APEX08 contains data related to the population attending the services, the accessibility of the different centres; the health and non-health staff working in the different centres; the activity of each of them, which refers, for example, to number of consultations or frequency of attendance in each speciality; the costs of different aspects of primary health care, such as staff or prescriptions; quality indicators, and environmental variables related to economic characteristics or of the population.

As mentioned before, for the analysis here presented we include variables at both levels the individual and the centre. The selection of exogenous variables is based on previous literature (Hall and Dorman, 1998; Lewis, 1994; Baker 1996; Sixma et al., 1998; Millar, 2001; Campbell et al., 2001; Salisbury et al., 2010) as well as on the availability of variables in the different datasets previously described. The variables finally selected at the patient level refer to sociodemographic information about service users and to some objective measures, such as age, gender, waiting times, among others. At the centre level, the variables used in the analysis refer to daily caseload in different specialities, the amount available staff and its experience, among others. All these variables and its main descriptive statistics are specified in Table 1.

[16] APEX is the acronym used in Spanish to specify that the data refers to Primary Health Care in Extremadura.

Table 1. Exogenous variables used in the analysis

Variables	Description		Characteristics	
Exogenous variables at the patient level			**N**	**%**
SEX	Patients' gender	0 = male	916	39.28
		1 = female	1,416	60.72
AGE	Patients' age	0 = <30 years old	226	9.62
		1 = 30-45 years old	660	28.10
		2 = 45-64 years old	573	24.39
		3 = ≥ 65 years old	890	37.89
MARITAL_STATUS	Patients' marital status	0 = single	387	16.64
		1 = married	1,399	60.15
		2 = separated	232	9.97
		3 = divorced	58	2.49
		4 = widowed	250	10.75
INCOME	Patients' monthly income level	0 = < 900 €/month	906	40.49
		1 = 900-1,800 €/month	852	38.36
		2 = 1,800-2,700 €/month	354	15.94
		3 = > 2,700 €/month	109	4.91
EDUCATION	Patients' highest level of education	0 = no studies	579	25.08
		1 = primary school	761	32.96
		2 = secondary school	516	22.35
		3 = university studies	453	19.62
WAITING_TIME	Waiting time when attending the practice	0 = < 10 minutes	90	3.89
		1 = 10-19 minutes	617	26.66
		2 = 20-29 minutes	591	25.54
		3 = ≥ 30 minutes	1,016	43.91
FREQUENCY	Frequency with which the patients has visited the centre in the previous 12 months	0 = 1-2 times	185	7.90
		1 = 3-4 times	312	13.33
		2 = 5-6 times	367	15.68
		3 = > 6 times	1,477	63.09
Exogenous variables at the centre level			**Mean**	**SD**
DAYVISTGP	Daily caseload in general practice		41.34	10.05
DAYVISTP	Daily caseload in paediatrics		23.48	8.36
DAYVISTN	Daily caseload in nursery		27.71	11.09
EXPERIENCE	Experience, in days, of the health staff		5900.66	442.45
HOSTESTS	Tests requested to hospitals to patient ratio		0.50	0.20
REFERRALS	Referrals from primary to secondary care to patient ratio		0.34	0.08
HLAB	Health staff to patient ratio		0.0025	0.00
NHLAB	Non-health staff to patient ratio		0.0009	0.00
PRESCRIPTIONS	Number of prescriptions to patient ratio		21.51	5.12
AREA	Area of the primary care centre to patient ratio		0.10	0.05

Source: Prepared by authors

As dependent variables we use the satisfaction of patients with (1) the facilities of the centre (FAC), (2) the health staff (HS), (3) the non-health Staff (NHS), and (4) the efficiency (EFF) with which patients perceive service is being delivered. An overall satisfaction (SAT) is also used as dependent variable by combining the previous four.

In order to build those dependent variables we use sections two and five of the satisfaction questionnaire (Appendix 1) designed for this. These two sections of the questionnaire refer to a series of importance and performance scores that service users assign to a total of twenty-five items about the health care and that can be grouped in different categories (Table 2).

Table 2. Attributes used to build the dependent variables

Category	Definition
Facilities	1. - Cleanliness of facilities
	2.- Equipment of the health centre
	3.- Location of the health centre
Health Staff	4.- Health staff cleanliness
	5.- Health staff professionalism
	6.- Health staff kindness and politeness
	7.- Trust in health staff
	8.- Personalised service
	9.- Communication with health staff
	10.- Health staff attention to patients' problems
	11.- Health staff interest in solving patients' problems
	12.- Health staff understanding of patients' problems
	13.- Health staff prestige
Non-health staff	14.- Non-health staff cleanliness
	15.- Non-health staff professionalism
	16.- Non-health staff kindness and politeness
	17.- Non-health staff attention to patients' problems
	18.- Non-health staff interest in solving patients' problems
Efficiency	19.- Ease of making an appointment
	20.- Bureaucracy intensity
	21.- Waiting time in the health centre before entering the consulting room
	22.- Speed of complementary tests
	23.- Resolution of complaints
	24.- Time devoted to each patient
	25.- Health centre opening hours

Source: Prepared by authors

First of all, we obtain a measure of satisfaction for each of the twenty-five individual items by multiplying the importance and the performance scores of each of them (Fishbein and Ajzen, 1975; Paul, 2003; Jain and Gupta, 2004). The combination of these two indicators adequately measures patients' satisfaction with the health care as it combines users' experiences with a series of elements of the systems and the importance they give to those elements.

Secondly, we build the dependent variables FAC, HS, NHS and EFF by calculating the average score of the items within every category of attributes, i.e. an average value of

the three facility variables, a score for the ten variables describing characteristics of the health staff, the average value of the five items characterising aspects related to the non-health staff, and a mean score for the attributes showing patients' satisfaction with the efficiency of the system.

Finally, we obtain the general satisfaction variable (SAT) that includes all the aspects of the service considered in the questionnaire by computing an average score of all the items. The main descriptive statistics of the dependent variables are specified in Table 3.

Table 3. Dependent variables used in the analysis

Variables	Description	Characteristics	
Dependent variables		**Mean**	**SD**
SAT	Overall patients' satisfaction with primary health care [1-49]	29.73	10.69
FAC	Patients' satisfaction with the facilities of the centre [1-49]	33.41	11.01
HS	Patients' satisfaction with the health staff [1-49]	32.77	12.09
NHS	Patients' satisfaction with the non-health staff [1-49]	29.44	12.53
EFF	Patients' satisfaction with the efficiency of the centre [1-49]	24.03	11.82

Source: Prepared by authors

4. RESULTS

The final sample size for our analysis consists of 96 centres and 2,384 patients. As previously mentioned, our patient level dataset contains information of patients attending 97 health centres while the dataset with the information at the aggregate level refers to 109 centres. However, we work with 96 centres because they are present in both datasets. Additionally, the literature has recommended that, when applying a multilevel model, we should have at least 20 groups of size at least 5 (De Leeuw and Kreft, 1995). In our analysis, we have investigated this issue in the 96 centres, each of which has a minimum of 5 patients.

Using this final data, we follow the additive approach, as previously described, in the presentation of our results. First of all we present the null models, where none of the explanatory variables are included. Secondly, we run a series of regressions including variables only at the patient level. Finally, we add to the previous models the variables at centre level.

4.1 Null models

In this section we present the results of the null models, i.e. the models run without the inclusion of any of the explanatory variables. The specification of these models is presented in equation 4.

$$Y_{ij} = \beta_0 + \mu_j + \varepsilon_{ij} \quad [4]$$

As mentioned before, the null model allows us to decompose the estimated variance in its patient and centre level components. Table 4 shows the proportion of variance explained by the model at these two levels. Furthermore, it also exhibits the Intraclass Correlation Coefficient (ICC), which indicates the proportion of the total variance that it is explained by the differences in the groups, i.e. by the differences in the centres. The ICC takes values in a range from zero to one. If the ICC is close to one it would be indicating that most of the variance of the dependent variables is due to aspects related to the groups, i.e. within each of the groups the individual units are very similar. If the ICC is near zero, it is indicating that within each of the groups, the individual units are very different. A small value of the ICC would be suggesting that we could ignore the multilevel structure, given that the proportion of variance to be explained at the group level is very low.

Table 4. Decomposition of the variance at the patient and centre levels

	Dependent variable				
	SAT	**FAC**	**HS**	**NHS**	**EFF**
Patient level	103.18	102.06	133.57	145.27	126.27
Centre level	13.09	20.73	12.91	13.74	17.82
ICC	11.26%	16.88%	8.81%	8.64%	12.37%

Source: Prepared by authors

As seen from the figures presented in Table 4, most of the variance of our satisfaction variables is explained at the individual level, although the centre level variables also have an influence in explaining patients' satisfaction. The highest percentages of the ICC are found for the models analysing satisfaction with the facilities and the efficiency of the centres.

4.2 Models with variables at the patient level only

The second step in our multilevel analysis is to include explanatory variables only at the first level, i.e. patient level variables. The general specification of these models is shown in equation 5.

$$Y_{ij} = \beta_0 + \beta_1 SEX_{ij} + \beta_2 AGE_{ij} + \beta_3 MARITAL\_STATUS_{ij} + \beta_4 INCOME_{ij} + \beta_5 EDUCATION_{ij} + \beta_6 WAITING\_TIME_{ij} + \beta_7 FREQUENCY_{ij} + \mu_j + \varepsilon_{ij} \quad [5]$$

Table 5 shows the estimation coefficients and the statistical significance of each of the variables included in the analyses and for the five models that are estimated.

Table 5. Regression coefficients for the regressions including patient level data only

Dependent variables / Exogenous variables	SAT	FAC	HS	NHS	EFF
INTERCEPT	33.20***	35.95***	34.62***	31.20***	31.41***
SEX	0.39	0.51	0.28	0.79	0.23
AGE					
30-45 years old	-0.12	-0.17	-0.76	1.47	-0.33
45-64 years old	1.03	1.40	0.55	2.53**	0.53
≥ 65 years old	3.56***	1.94*	2.77**	5.93***	3.79***
MARITAL_STATUS					
married	-0.32	-0.19	0.04	-0.86	-0.54
separated	-1.32	-1.19	-0.52	-1.78	-2.21**
divorced	-1.30	-0.68	-2.08	-1.53	-0.23
widowed	-0.78	-0.99	-0.39	-1.45	-0.80
INCOME					
900-1,800 €/month	0.67	0.13	1.22**	0.88	-0.05
1,800-2,700 €/month	0.43	-0.09	1.00	0.57	-0.26
> 2,700 €/month	0.78	-1.00	1.73	0.86	0.17
EDUCATION					
primary school	0.45	-0.18	0.40	0.39	0.84
secondary school	1.60**	0.50	1.78**	1.69*	1.78**
university studies	-0.51	-2.22**	-0.18	-0.98	0.21
WAITING_TIME					
10-19 minutes	-2.74**	-1.70	-1.93	-2.84*	-4.31***
20-29 minutes	-6.17***	-3.62***	-5.32***	-6.00***	-8.63***
≥ 30 minutes	-8.84***	-4.61***	-7.77***	-8.05***	-12.78***
FREQUENCY					
3-4 times	-1.62*	-2.25**	-1.39	-1.63	-1.72*
5-6 times	-1.07	-2.32**	-0.73	-1.60	-0.68
> 6 times	0.77	0.10	1.37	0.20	0.55
ICC	12.88%	20.10%	10.44%	8.96%	12.40%
***p<0.1, **p<0.05; ***p<0.001**					

Source: Prepared by authors

For the patients' general satisfaction with health care, our regression shows that the variables having a significantly association with it are patients' age, education, waiting times and frequency of attendance in the previous twelve months. In relation to age, out of the four age groups, those being 65 or older are statistically more satisfied with primary health care than the youngest group of patients. Focussing on education, the statistically significant relationship refers to patients having a maximum of secondary school. Service users with such qualification show a better satisfaction than those who have no education. Waiting time to be seen is statistically significant related to patients' overall satisfaction with health care. In general, we find that the longer patients wait to enter the consultation room, the lower the satisfaction with primary health care. Those that wait for 30 minutes or more, have a satisfaction score of almost nine points less, compared to those that have to wait for less than 10 minutes. Finally, considering the frequency with which patients have attended primary care service in the previous twelve months, our results show that only those that have gone to the centre between three or four times in the previous year are less satisfied than those visiting their centre once or twice.

After looking at which patients' characteristics most influence satisfaction with primary health care, we study how these characteristics affect satisfaction with the facilities, health and non-health staff and the efficiency of the centres in the region.

In relation to the facilities of the centres, we find that those aged 65 or older report a better satisfaction than the youngest groups, although this relationship is only significant at the 10% level of significance. Patients with university studies report significant lower levels of satisfaction compared to those patients without any kind of studies. In relation to waiting times, we find a similar relationship with satisfaction with the facilities of the centre than we found when studying general satisfaction, i.e. the longer the time the patients wait to be seen, the lower the satisfaction with the facilities of the centre, although, this relationship is only statistically significant for a waiting time of 20 minutes or longer. Finally, the frequency with which patients have required primary health care services in the previous twelve months is significantly associated with

patients' satisfaction with the facilities of the primary health care centre. Specifically, visiting the centre between three and six times, reduces the satisfaction with primary care in more than two points compared to service users attending the centre only once or twice during the year.

Regarding the health staff, our multilevel regression shows that patients aged 65 or older are significantly more satisfied with this aspect of primary health care than the youngest patients. This difference in satisfaction is of almost three points. Unlike, the rest of regressions that have not shown any statistically significant relationship between satisfaction and income level, when looking at satisfaction with the health staff we find that those having a monthly income of between 900 and 1800 €/month are more satisfied than those earning less than 900 € per month. In relation to patients' education, out of the four education groups, we find that only having a maximum of secondary education is associated with a higher satisfaction with the health staff compared to patients that have no studies. As in the previous regressions, waiting times before entering the consultation room is highly related to our dependent variable. Results report that the longer the patients wait for, the worse their satisfaction with the health staff, with differences of more than eight points for those that wait the most compared to patients waiting for less than 10 minutes.

We also study satisfaction with the non-health staff. Out of the seven variables used in the regression, three of them show a statistically significant relationship with our dependent variable: age, education and waiting time. For age, we find that the older the patient the better his/her satisfaction with the non-health staff, although this relationship is only significant for the two oldest groups. For patients' education, our regression states that service users with a maximum of secondary school report better levels of satisfaction than those without any kind of qualifications. Nevertheless, this relationship is only significant at the 10%. As in the previous regressions, waiting time is the most significant variable, showing that the longer the time waited to be seen the lower the satisfaction with the non-health staff reported by the patient.

Finally and in relation to satisfaction with the efficiency of primary health care, our model reveals statistically significant relationships with age, marital status, education level, waiting times and frequency. In relation to age, patients being 65 and older report a significant higher satisfaction with the efficiency of the centre than the youngest group. Patients indicating being separated are less satisfied than single patients. For the variable education, and as in previous cases, we find that service users having a maximum qualification of secondary school report better satisfaction than patients without any kind of education. Time patients have to wait to be seen have a very high impact in patients' satisfaction with the efficiency. As in previous regressions, the longer the waiting time, the lower the satisfaction, however, in this case, the variable is significant even at the 1% level of significance, and those having to wait for more than 30 minutes to be seen, have a satisfaction score of thirteen points lower than those who wait for less than 10 minutes. Finally, and in relation to the frequency with which patients attend the centres, we find that visiting the centre three or four times compared to once or twice during the year, significantly reduces patients' satisfaction in almost 2 points.

4.3 Models with variables at the patient and at the centre level

Finally, we add variables at the centre level. It is important to note that we include all the available variables in the model with the general satisfaction as dependent variable. However, for the rest of models, the inclusion of exogenous variables at the centre level varies based on the dependent variable being considered. For example, we do not include the number of health staff in the regression run using the satisfaction with the non-health staff as dependent variable. The final specification of each of these models is presented in equations 6 to 10.

$$SAT_{ij} = \beta_0 + \beta_1 SEX_{ij} + \beta_2 AGE_{ij} + \beta_3 MARITAL\_STATUS_{ij} + \beta_4 INCOME_{ij} + \beta_5 EDUCATION_{ij} + \beta_6 WAITING\_TIME_{ij} + \beta_7 FREQUENCY_{ij} + \beta_8 DAYVISITGP_j + \beta_9 DAYVISITP_j + \beta_{10} DAYVISITN_j + \beta_{11} EXPERIENCE_j + \beta_{12} HOSTESTS_j + \beta_{13} REFERRALS_j + \beta_{14} HLAB_j + \beta_{15} NHLAB_j + \beta_{16} PRESCRIPTIONS_j + \beta_{17} AREA_j + \mu_j + \varepsilon_{ij} \quad [6]$$

$$FAC_{ij} = \beta_0 + \beta_1 SEX_{ij} + \beta_2 AGE_{ij} + \beta_3 MARITAL\_STATUS_{ij} + \beta_4 INCOME_{ij} + \beta_5 EDUCATION_{ij} + \beta_6 WAITING\_TIME_{ij} + \beta_7 FREQUENCY_{ij} + \beta_8 DAYVISITGP_j + \beta_9 DAYVISITP_j + \beta_{10} DAYVISITN_j + \beta_{11} AREA_j + \mu_j + \varepsilon_{ij} \quad [7]$$

$$HS_{ij} = \beta_0 + \beta_1 SEX_{ij} + \beta_2 AGE_{ij} + \beta_3 MARITAL_{STATUS\,ij} + \beta_4 INCOME_{ij} + \beta_5 EDUCATION_{ij} + \beta_6 WAITING_{TIME\,ij} + \beta_7 FREQUENCY_{ij} + \beta_8 DAYVISITGP_j + \beta_9 DAYVISITP_j + \beta_{10} DAYVISITN_j + \beta_{11} EXPERIENCE_j + \beta_{12} HOSTESTS_j + \beta_{13} REFERRALS_j + \beta_{14} HLAB_j + \beta_{15} PRESCRIPTIONS_j + \beta_{16} AREA_j + \mu_j + \varepsilon_{ij} \quad [8]$$

$$NHS_{ij} = \beta_0 + \beta_1 SEX_{ij} + \beta_2 AGE_{ij} + \beta_3 MARITAL\_STATUS_{ij} + \beta_4 INCOME_{ij} + \beta_5 EDUCATION_{ij} + \beta_6 WAITING\_TIME_{ij} + \beta_7 FREQUENCY_{ij} + \beta_8 DAYVISITGP_j + \beta_9 DAYVISITP_j + \beta_{10} DAYVISITN_j + +\beta_{11} NHLAB_j + \beta_{12} PRESCRIPTIONS_j + \beta_{13} AREA_j + \mu_j + \varepsilon_{ij} \quad [9]$$

$$EFF_{ij} = \beta_0 + \beta_1 SEX_{ij} + \beta_2 AGE_{ij} + \beta_3 MARITAL\_STATUS_{ij} + \beta_4 INCOME_{ij} + \beta_5 EDUCATION_{ij} + \beta_6 WAITING\_TIME_{ij} + \beta_7 FREQUENCY_{ij} + \beta_8 DAYVISITGP_j + \beta_9 DAYVISITP_j + \beta_{10} DAYVISITN_j + \beta_{11} NHLAB_j + \beta_{12} AREA_j + \mu_j + \varepsilon_{ij} \quad [10]$$

The results of the aforementioned regressions are presented in Table 6. In relation to general satisfaction with primary health care, results show that for the patient level variables, only age, education, waiting time and frequency have a significant impact on satisfaction. These are the same significant variables than in the previous section where we only considered explanatory variables at the patient level. For the variables at the centre level, the daily caseload in paediatrics and the number of prescriptions are the ones having a statistically significant impact on patients' general satisfaction with primary health care. For the first of them, our results show that the higher the daily caseload in paediatrics, the lower the satisfaction with primary healthcare. In relation to the number of prescriptions, we also find an inverse relationship, although this is only significant at the 10% level.

Table 6. Regression coefficients for the regressions including patient and centre level data

Dependent variables / Exogenous variables	SAT	FAC	HS	NHS	EFF
INTERCEPT	36.26***	33.96***	37.12***	34.78***	31.67***
SEX	0.41	0.52	0.30	0.82	0.26
AGE					
30-45 years old	-0.08	-0.13	-0.71	1.50	-0.24
45-64 years old	1.08	1.44	0.58	2.60**	0.69
≥ 65 years old	3.63***	2.00**	2.83**	6.01***	3.96***
MARITAL_STATUS					
married	-0.35	-0.23	0.02	-0.88	-0.59
separated	-1.38	-1.22	-0.56	-1.79	-2.24**
divorced	-1.22	-0.64	-1.98	-1.43	-0.18
widowed	-0.82	-1.01	-0.40	-1.50	-0.86
INCOME					
900-1,800 €/month	0.69	0.17	1.19*	0.94	0.03
1,800-2,700 €/month	0.47	-0.04	0.99	0.64	-0.16
> 2,700 €/month	0.78	-0.95	1.68	0.87	0.23
EDUCATION					
primary school	0.54	-0.11	0.50	0.50	1.01
secondary school	1.72**	0.61	1.85**	1.88**	2.05**
university studies	-0.31	-2.07**	-0.06	-0.73	0.57
WAITING_TIME					
10-19 minutes	-2.74**	-1.75	-1.97	-2.77*	-4.35***
20-29 minutes	-6.17***	-3.67***	-5.38***	-5.90***	-8.61***
≥ 30 minutes	-8.79***	-4.62***	-7.79***	-7.88***	-12.72***
FREQUENCY					
3-4 times	-1.63*	-2.25	-1.36	-1.67	-1.75*
5-6 times	-1.08	-2.32	-0.69	-1.66	-0.73
> 6 times	0.80	0.11	1.45	0.19	0.53
DAYVISTGP	0.01	0.05	0.01	0.01	0.02
DAYVISTP	-0.14**	-0.13**	-0.13**	-0.12**	-0.16***
DAYVISTN	0.00	0.01	0.01	-0.02	-0.03
EXPERIENCE	0.00	-	0.00	-	-
HOSTESTS	1.59	-	0.66	-	-
REFERRALS	-2.55	-	-3.04	-	-
HLAB	-248.93	-	74.26	-	-
NHALB	1535.12	-	-	2221.25**	1509.71*
PRESCRIPTIONS	-0.19*	-	-0.22*	-0.19	-
AREA	15.78	20.51**	20.06***	6.32	11.84
ICC	10.27%	18.34%	8.79%	7.18%	9.24%

*p<0.1, **p<0.05; ***p<0.001

Source: Prepared by authors

As in the previous section, we also run a series of regressions using as dependent variable satisfaction with different aspects of primary health care: facilities, health staff, non-health staff and efficiency.

In the case of the facilities, the variables at the patient level that were statistically significant in the previous section are still significant in this one except for the case of the

frequency with which patients had attended primary health care centres in the previous twelve months. For the second level variables, the daily caseload in paediatrics is significantly associated with the satisfaction with the facilities of the centre; specifically, this relationship is negative, indicating that the higher the daily caseload in this speciality, the lower the satisfaction with the efficiency with which primary care centres deliver their services. Furthermore, the area of the centre also has significant impact in the satisfaction with the facilities of primary health care, indicating that the bigger the area of the centre the better the level of satisfaction with the facilities reported by patients.

Focussing on satisfaction with the health staff, being 65 and older, having a monthly income between 900 and 1800 €, a maximum of secondary school and waiting for 20 minutes or more to be seen by the doctor continues to be statistically associated with patients' satisfaction with the health staff. Furthermore, the daily caseload in paediatrics, the number of prescriptions and the area of the centre are variables at the centre level having also an impact on satisfaction with the this dimension of primary health care. Results show that the highest the daily caseload in this speciality, the lower the satisfaction with the health staff reported by the patients. In the case of the prescriptions, our regression indicates that patients do not want to be prescribed more drugs, as this variable is negatively associated with satisfaction. In the case of the area of the centre, as in the previous model, patients report that the bigger the area of the centre the better they feel.

In relation to the non-health staff, besides the variables at the patient level that we already found to be statistically significant related to satisfaction with this aspect of the health care, we also identify several variables at the centre level that are influencing patients' satisfaction with the non-health staff: the daily caseload in paediatrics and the number of non-health staff. Results reveal that the greater the daily caseload in paediatrics, the lower the satisfaction with the non-health staff. In the case of the number of staff in this speciality, we find that an increase in the number of this type of workers would be translated in a better service users' satisfaction.

Finally, and attending to the efficiency of the service, we found that, besides the variables at the patient level that we already pointed to be statistically associated with efficiency, at the centre level, the daily caseload with the paediatrician and the number of non-health staff affect patients' satisfaction with this aspect of health care. Specifically, regressions show that the greater the daily caseload with the paediatrician, the lower the satisfaction with the efficiency of the centre and that having more non-health staff employed in the centre is likely to increase patients' satisfaction with the efficiency of it.

5. DISCUSION AND CONCLUSION

Our study has analysed what characteristics, at the patient and at the centre levels most influence patients' general satisfaction with primary health care and patients' satisfaction with a series of elements of the system such as the facilities, the health staff, the non-health staff and the efficiency of the centre. The analysis has been performed employing data provided by patients about their satisfaction with the service received and data about the main characteristics of the primary health care system in the Spanish Autonomous Community of Extremadura.

We have applied several multilevel models to determine which variables, at the patients and at the centre level, most influence patients' satisfaction with the different elements of the primary health: general satisfaction and satisfaction with the facilities, the health and non-health staff and the efficiency of the centre. The use of this type of methodology introduces a great potential to the research, as it includes data not only at the individual level but also at the group level, providing with that essential information (Diez-Roux, 2008). Furthermore, this type of analysis facilitates the decision making process because besides identifying the characteristics, at any level, that influence satisfaction; it provides with valuable information about which of the levels included have a higher impact on satisfaction and allowing policy makers to focus more in one level or another. In our analysis, the patient level explains more of the variation in satisfaction than does the centre level; therefore, without disparaging centres, the focussed should be more on trying to improve patients' satisfaction.

Unlike previous literature that has analysed patients' satisfaction with primary health care (Sixma et al., 1998; Salisbury et al., 2010, among others), we have not directly asked service users to state how satisfied they are with the service or with a particular aspect of it. Instead, we have calculated several satisfaction variables based on patients' reported experiences with the care received and the importance they have assigned to different aspect of the service, following recommendations of previous literature (Fishbein and Ajzen, 1975; Paul, 2003; Jain and Gupta, 2004) and obtaining, with that, a variable that apart from appropriately measure satisfaction, gathers more information about patients than a general satisfaction question.

Comparison of results with previous literature may be difficult as there exist inconsistencies between studies in for example, how they define the measures of patients' satisfaction, the range of attributes of primary care that are explored, the definition of the exogenous variables, etc. Despite that, some common themes may be identified.

In relation to the variables at the patient level, similarly to previous research, we have not found any statistical association between satisfaction and the gender of the patient (Campbell et al., 2001). According to previous literature, older patients are more satisfied with the health care received than younger ones (Campbell et al., 2001; Sixma et al., 1998; Millar, 2001; Hall and Dorman, 1990), supporting our study this relationship and being especially relevant for those aged 65 and older.

Focussing on education, previous literature has pointed to an inverse relation between qualifications and satisfaction (Millar, 2001; Hall and Dorman, 1990; Lewis, 1994). In our study this is only confirmed when the satisfaction with the facilities is considered. For the rest of satisfaction variables having an education level up to secondary school was statistically significant associated with a higher satisfaction than those without any qualifications. Therefore, it cannot be generally established that satisfaction with primary health care decreases when the patients' level of education of the increases.

For the variable waiting time, it is of high relevance the fact that regardless the aspects of the primary health care for which we study satisfaction, the time patients need to wait before seeing their doctor has a very strong impact on their satisfaction. Waiting times are an important aspect of health care and much research has focused on the topic, not only looking at its relationship with satisfaction, for which previous results have been similar to ours (Dansky and Mile, 1997 or Anderson et al., 2007 as some examples of it), but also analysing its association with socioeconomic status, equity, among other variables (Laudicella et al., 2012; Cooper et al., 2009). Given the high impact that this variable has in satisfaction, policy makers should focus on trying to eliminate or reduce those waiting times to a minimum, so that patients are more satisfied with the care received.

For the aggregate level variables, similarly to what Sixma et al. (1998) found, our results suggested that no association can be established between satisfaction and referrals to secondary care and between satisfaction and number of drugs prescribed (although with some minor exceptions for the latter variable). In line with findings from Salisbury et al. (2010) the number of health staff did not influence general satisfaction with primary health care in our analysis, despite the fact that the definition of the variables in both pieces of research was different. Furthermore, the experience of the GP was not statistically associated with satisfaction similar to what Salisbury et al. (2010) found.

To the best of our knowledge, previous literature has not considered how the daily caseload influences patients' satisfaction. Our research has analysed this aspect and results have indicated that the daily caseloads in paediatrics is statistically associated with satisfaction. Specifically, this relationship has been negative, indicating that the higher the caseload, the lower the satisfaction. Therefore, efforts need to be focused on trying to reduce those caseloads so that a high quality health care that increases patients' satisfaction can be delivered. If doctors can face a lower caseload that means they could spend more time in each visit and offer a more personalised health care than when they have a long list of patients to see every day. This latter relationship has been studied by

previous research, showing the existence of a correlation between the length of the visits and patients satisfaction (Cape, 2002). Additionally, existing literature has also indicated that patients highly value the time that doctors devote to them (Sirdifield et al., 2016; Miranda et al., 2010a, 2010b). To finish with the variables at the centre level, our analysis has also revealed that in order to improve patients' levels of satisfaction with the non-health staff, additional personnel may be necessary.

In view of all the above, to increase patients satisfaction with the primary health care services, the focus should be primarily on some main aspects. First, efforts need to be done in reducing the waiting times as this variable has been the most significant one, regardless which aspect of the satisfaction with primary health care we have considered. Previous research has adapted some principles from industrial engineering to the health field concluding that delays can be satisfactorily reduced without requiring additional resources (Murray and Berwick, 2003), so similar applications could be done in our health system. Second, the daily caseloads, which have had a negative impact on satisfaction, could be reduced by either increasing the number of staff, reducing the number of patients that each doctor sees on a day or increasing the length of the visits, although this may need an investment of additional resources. Finally, older people has stated to be more satisfied with primary health care, therefore, it is important that we try to maintain their levels of satisfaction as they are normally the group of people that require the use of health service more frequently.

REFERENCES

Anderson, R.T., Camacho, F.T. and Balkrishnan, R. (2007). Willing to wait? the influence of patient wait time on satisfaction with primary care. *BMC health services research*, 7 (1), p. 1.

Baker, R. (1996). Characteristics of practices, general practitioners and patients related to levels of patients' satisfaction with consultations. *British Journal of General Practice*, 46 (411), pp. 601-605.

Berry, L. L., Zeithaml, V. A. and Parasuraman, A. (1985). Quality Counts in Services too. *Business Horizons*, 28 (3), pp. 44-52.

Bitner, M. J., Booms, B. H. and Tetreault, M. S. (1990). The service encounter: diagnosis favourable and unfavourable incidents. *Journal of Marketing*, 54, pp. 71-84.

Campbell, J.L., Ramsay, J. and Green, J. (2001). Age, gender, socioeconomic, and ethnic differences in patients' assessments of primary health care. *Quality in health care*, 10 (2), pp.90-95.

Cape, J., (2002). Consultation length, patient-estimated consultation length, and satisfaction with the consultation. *British Journal of General Practice*, 52 (485), pp. 1004-1006.

Cooper Z. N., McGuire, A., Jones, S., and Le Grand, J. (2009). Equity, waiting times and NHS reforms: restrospective study, *BMJ*, 339.

Cronin, J. and Taylor, S. (1992). Measuring service quality: a reexamination and extension. *Journal of Marketing*, 56, pp. 55-68.

Dansky, K.H. and Miles, J. (1997). Patient satisfaction with ambulatory healthcare services: waiting time and filling time. *Journal of Healthcare Management*, 42 (2), pp. 165.

Dawson, J., Fitzpatrick, R., Murray, D. and Carr, A. (1998). Questionnaire on the perceptions of patients about total knee replacement. *Journal of Bone and Joint Surgery-British Volume*, 80 (1), pp. 63-69.

De Leeuw, J. and Kreft, I.G. (1995). Questioning multilevel models. *Journal of Educational & Behavioral Statistics*, 20 (2), pp. 171-189.

Diez-Roux, A.V. (1998). Bringing context back into epidemiology: variables and fallacies in multilevel analysis. *American Journal of Public Health*, 88 (2), pp. 216-222.

Donabedian, A. (1984). La Calidad de la Atención Médica. *La Prensa Mexicana*, México.

Dormont, B. and Milcent, C. (2004). The sources of hospital cost variability. *Health Economics*, 13 (10), pp. 927-939.

Duncan, G., Jones, K. and Moon, G. (1993). Do places matter? A multi-level analysis of regional variation in health-related behaviour in Britain. *Social Science & medicine*, 37 (6), pp. 725-733.

Fishbein, M. and Ajzen, I. (1975). Belief, attitude, intention and behavior. An introduction to theory and research. *Reading Mass: Addison-Wesley*.

Hall, J.A. and Dornan, M.C. (1990). Patient sociodemographic characteristics as predictors of satisfaction with medical care: a meta-analysis. *Social Science & Medicine*, 30 (7), pp. 811-818.

Jain, S. K. and Gupta, G. (2004). Measuring Service Quality: SERVQUAL vs. SERVPERF scales. *Vikalpa*, 29 (2), pp. 25-37.

Laudicella, M., Olsen, K. R. and Street, A. (2010). Examining cost variation across hospital departments–a two-stage multi-level approach using patient-level data. *Social Science & Medicine,* 71 (10), pp. 1872-1881.

Laudicella, M., Siciliani, L. and Cookson, R. (2012). Waiting times and socioeconomic status: evidence from England. *Social Science & Medicine*, 74 (9), pp. 1331-1341

LeBlanc, G. and Nguyen, N. (1988). Customers satisfaction with service quality in financial institutions. *International Journal of Bank Marketing*, 6 (4), pp. 7-18.

Lewis, J.R. (1994). Patient views on quality care in general practice: literature review. *Social Science & Medicine*, 39 (5), pp. 655-670.

Luke, D.A. (2004). Multilevel modelling. *Sage*, 143.

Martilla, J. A. and James, J. C. (1977). Importance-Performance analysis. *Journal of Marketing*, 41, pp. 77-79.

Millar, M. (2001). Patient satisfaction with general practice in Ireland. *Irish medical journal*, 94 (4), pp. 106-108.

Miranda, F. J., Chamorro, A., Murillo, L. R. and Vega, J. (2010a). Assessing primary health care services quality in Spain: managers vs. Patietns perceptions". *The Service Industries Journal*, 30 (13), pp. 2137-2149.

Miranda, F. J., Chamorro, A., Murillo, L. R. and Vega, J. (2010b). An Importance-Performance Analysis of primary health care services: managers vs. patients perceptions. *Journal of Service Science and Management*, 3 (02), pp. 227-234.

Murillo-Zamorano, L.R., Vega, J., Corrales, N., Fernández, Y., de Miguel, F.J., 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.

Murray, M. and Berwick, D.M. (2003). Advanced access: reducing waiting and delays in primary care. *Jama*, 289 (8), pp. 1035-1040.

Paul, D. P. (2003). An exploratory examination of SERVQUAL versus SERVPERF for Prosthectic Dental Specialists. *Clinical Research and Regulatory Affairs*, 20 (1), pp. 89-100.

Pullicino, G., Sciortino, P., Camilleri, L., Schäfer, W. and Boerma, W. (2016). The Influence of Patient Characteristics on Healthcare-Seeking Behaviour: A Multilevel Analysis of 70

Primary Care Practices in Urban-Suburban Regions in Malta. *Quality in Primary Care*, 24 (3), pp.106-110.

Rice, N. and Jones, A. (1997). Multilevel models and health economics. *Health Economics,* 6 (6), pp. 561-575.

Salisbury, C., Wallace, M. and Montgomery, A.A. (2010). Patients' experience and satisfaction in primary care: secondary analysis using multilevel modelling. *BMJ*, *341*, p. 5004.

Schäfer, W.L., Boerma, W.G., Murante, A.M., Sixma, H.J., Schellevis, F.G. and Groenewegen, P.P. (2015). Assessing the potential for improvement of primary care in 34 countries: a cross-sectional survey. *Bulletin of the World Health Organization*, 93 (3), pp. 161-168.

Sirdifield, C., Godoy-Caballero, A., Windle, K., Jackson, C., McKay, S., Schäfer, W. and Siriwardena, A.N. (2016). Comparing importance and performance from a patient perspective in English general practice: a cross-sectional survey. *Family practice*, 33 (2), pp. 179-185.

Sixma, H.J., Spreeuwenberg, P.M. and van der Pasch, M.A., 1998. Patient satisfaction with the general practitioner: a two-level analysis. *Medical care*, 36 (2), pp. 212-229.

Street, A., Kobel, C., Renaud, T. and Thuilliez, J. (2012). How well do diagnosis-related groups explain variations in costs of length of stay among patients and across hospitals? Methods for analysing routine patient data. *Health Economics,* 21 (S2), pp. 6-18.

Von Korff, M., Koepsell, T., Curry, S. and Diehr, P. (1992). Multi-level analysis in epidemiologic research on health behaviors and outcomes. *American Journal of Epidemiology*, 135 (10), pp.1077-1082.

APPENDICES

Appendix 1. Satisfaction questionnaire delivered to primary health care users in Extremadura

JUNTA DE EXTREMADURA
Consejería de Sanidad y Dependencia

UNIVERSIDAD DE EXTREMADURA

QUALITY ANALYSIS OF PRIMARY HEALTH CARE IN EXTREMADURA

Dear user of primary health care service in Extremadura,

Thanks for devoting the time to answer this questionnaire. The results obtained from it will allow us to have a better understanding of the way primary health care in the region is being delivered, so that, a better quality service can be provided.

All the information provided here is strictly confidential and it will be analysed together with the rest of questionnaires received from other primary health care users.

Instructions: Circle the option which best describes your experience when attending your primary health care centre. Hand this survey in to the reception of your corresponding centre. If you wish, you can fill the survey online following this link: http://mercado.unex.es/salud.

Thanks in advance for taking part in the survey

GENERAL INFORMATION
1.- Primary Health Centre:
2.- Gender: ☐ Male ☐ Female
3.- Age: ☐ < 30 years old ☐ 30-45 years old ☐ 45-64 years old ☐ ≥ 65 years old
5.- Family income: ☐ < 900€/m ☐ 900-1800 €/m ☐ 1800-2700 €/m ☐ ≥ 2700 €/m
6.- Education: ☐ no studies ☐ primary school ☐ secondary school ☐ university
7.- Do you live in the same town where the primary health care is located? ☐ Yes ☐ No
8.- How long does it take you to travel to the centre?:

HEALTH CENTRE FACILITIES							
Please, indicate your level of agreement with the following statements regarding your primary health care centre (where 1 means completely disagree and 7 means completely agree)							
In your opinion:	Completely disagree						Completely agree
1.- Health centre facilities are clean	1	2	3	4	5	6	7
2.- Health centre equipment is modern	1	2	3	4	5	6	7
3.- According to your place of residence, health centre location is adequate and easy to access	1	2	3	4	5	6	7

HEALTH STAFF (GENERAL PRACTITIONERS AND NURSES)							
In your opinion:	Completely disagree						Completely agree
1.- Health staff are well dressed and clean	1	2	3	4	5	6	7
2.- Health staff know well what they do (they are professionals)	1	2	3	4	5	6	7
3.- Health staff are kind and polite	1	2	3	4	5	6	7
4.- Health staff inspire trust	1	2	3	4	5	6	7
5.- I have received personalised care	1	2	3	4	5	6	7
6.- I have been well informed about the nature and objectives of the recommended treatment	1	2	3	4	5	6	7
7.- Health staff listen carefully to my problems	1	2	3	4	5	6	7
8.- Health staff have shown an interest in solving my problem	1	2	3	4	5	6	7
9.- Health staff have dealt with my problems/needs very quickly	1	2	3	4	5	6	7
10. Health staff prestige is high	1	2	3	4	5	6	7

NON-HEALTH STAFF (ADMIN STAFF, CENTRE ATTENDANTS, ETC.)							
In your opinion:	Completely disagree						Completely agree
1.- Non-health staff are well dressed and clean	1	2	3	4	5	6	7
2.- Non-health staff know well what they do (they are professionals)	1	2	3	4	5	6	7
3.- Non-health staff are kind and polite	1	2	3	4	5	6	7
4.- Non-health staff listen carefully to my problems	1	2	3	4	5	6	7
5.- Non-health staff have shown an interest in solving my problems	1	2	3	4	5	6	7

EFFICIENCY MEASURES							
In your opinion:	Completely disagree						Completely agree
1.- It is easy to contact the centre to make an appointment	1	2	3	4	5	6	7
2.- The level of bureaucracy is reasonable	1	2	3	4	5	6	7
3.- Once I arrived at the centre, waiting time to be seen was reasonable	1	2	3	4	5	6	7
4.- Complementary tests (x-ray, tests, etc.) are performed promptly	1	2	3	4	5	6	7
5.- When there has been a complaint, health centre response has been adequate	1	2	3	4	5	6	7
6.- Time I have been given has been adequate	1	2	3	4	5	6	7
7.- Opening times have been appropriate	1	2	3	4	5	6	7

OVERALL SATISFACTION							
	Completely disagree						Completely agree
1.- In general, I am very satisfied with health centre services	1	2	3	4	5	6	7
2.- Service received fulfils what I expected from it	1	2	3	4	5	6	7
3.- Health centre reputation is really good	1	2	3	4	5	6	7

OBJECTIVE MEASURES
1.- Time waited to be seen: ☐ < 10 min ☐ 10-19 min ☐ 20-29 min ☐ ≥ 30 min
2.- Years attending this particular health centre: ☐ 0-1 yr ☐ 2-5 yrs ☐ 6-10 years ☐ ≥ 10 yrs
3.- Times I have visited the centre in the last 12 months: ☐ 1-2 ☐ 3-4 ☐ 5-6 ☐ > 6 times

IMPORTANCE LEVEL FOR EACH OF THE ATTRIBUTES							
Please, state how important the following aspects of service quality are for you							
In your opinion:	Not very important						Very important
1.- Cleanliness of facilities	1	2	3	4	5	6	7
2.- Equipment of the health centre	1	2	3	4	5	6	7
3.- Location of the health centre	1	2	3	4	5	6	7
4.- Health staff cleanliness	1	2	3	4	5	6	7
5.- Health staff professionalism	1	2	3	4	5	6	7
6.- Health staff kindness and politeness	1	2	3	4	5	6	7
7.- Trust in health staff	1	2	3	4	5	6	7
8.- Personalised services	1	2	3	4	5	6	7
9.- Communication with health staff	1	2	3	4	5	6	7
10.- Health staff attention to patients' problems	1	2	3	4	5	6	7
11.- Health staff interest in solving patients' problems	1	2	3	4	5	6	7
12.- Health staff understanding of patients' problems	1	2	3	4	5	6	7
13.- Health staff prestige	1	2	3	4	5	6	7
14.- Non-health staff cleanliness	1	2	3	4	5	6	7
15.- Non-health staff professionalism	1	2	3	4	5	6	7
16.- Non-health staff kindness and politeness	1	2	3	4	5	6	7
17.- Non-health staff attention to patients' problems	1	2	3	4	5	6	7
18.- Non-health staff interest in solving patients' problems	1	2	3	4	5	6	7
19.- Ease of making an appointment	1	2	3	4	5	6	7
20.- Bureaucracy intensity	1	2	3	4	5	6	7
21.- Waiting time in the centre before entering the consultation room	1	2	3	4	5	6	7
22.- Speed of complementary tests	1	2	3	4	5	6	7
23.- Resolution of complaints	1	2	3	4	5	6	7
24.- Time spent with each patient	1	2	3	4	5	6	7
25.- Health centre opening hours	1	2	3	4	5	6	7

CHAPTER 4: USING PATIENT REPORTED OUTCOME MEASURES (PROMs) IN THE ANALYSIS OF HOSPITAL QUALITY PERFORMANCE

1. INTRODUCTION

Providers of secondary care in the English National Health Service (NHS) have been required to collect patient reported outcome measures (PROMs) since April 2009. The English Department of Health define PROMs as those questionnaires filled by patients to evaluate their own health status before and after certain interventions (Department of Health, 2008). PROMs are collected before and three or six months after surgery for four elective procedures: hip and knee replacement, varicose vein surgery and hernia repair.

Questionnaires delivered to patients can be both generic as well as specific for a particular disease. Generic questionnaires are common for all the interventions without including any reference to a particular condition, allowing, therefore, for the comparison among them. These questionnaires are the European Quality of life-5 Dimensions (EQ-5D) compound of 5 questions related to different aspects of patients' quality of life, and the European Quality of life-Visual Analogue Scale (EQ-VAS) consisting of a scale measuring patients' health status where the minimum value corresponds to the worst health status that patient can imagine and the maximum to the best possible health status.

Specific questionnaires for each of the diseases are: Oxford Hip Score (OHS) for the hip replacement, Oxford Knee Score (OKS) for the knee replacement and Aberdeen Varicose Vein Questionnaire (AVVQ) for the varicose vein surgery. Each of them presents a series of questions specifically related to the disease to which they refer, without containing questions related to patients' general health status.

The inclusion of PROMs (both generic and specific) in the analysis of quality in the health care allows for the consideration of patients' perspective about their health and about their health related quality of life (Devlin and Appleby, 2010). Patients' point of view is important given that, following to Dawson et al. (1998), they "provide reliable and valid judgements of health status and of the benefits of treatment"

The collection of PROMs marks a change in the way performance of secondary providers of care is assessed. Many analyses of provider performance have focused on

the activity or output of different health centres. More recently, the analysis of hospital performance has also included measures of patient outcomes rather than just hospital outputs, with a change in emphasis from the production of health care to the production of health itself (Devlin and Appleby, 2010). Outcome analysis had tended to be limited to measures of mortality or emergency readmissions (Thomas et al., 1994; Dimick et al., 2012; Selim et al., 2002; Chua et al., 2010), but the use of PROMs potentially allows for greater insight into the changes of the Health Related Quality of Life (HRQoL) that a patient may enjoy as a result of hospital activity (Gutacker et al., 2013).

Despite the existence of both generic and specific measurement instruments, the English NHS has preferred the use the generic instrument EQ-5D when hospital health care has been evaluated (NICE, 2008). Nevertheless, and in order to assess that choice, in this research we consider both generic as well as specific PROMs.

Considering all the aforementioned, the aim of this research is to analyse quality in the provision of health service in the NHS, using different PROMs, as well as to study whether hospital performance varies depending on the utilisation of one instrument or another. In doing so, we contribute to the literature in three ways.

First of all, we consider the subjective quality in the delivery of health services. Previous literature has analysed quality in health care, but in most of the cases, this has referred to indicators related to quality of the activity delivered (Salinas-Jiménez and Smith, 1996; García et al., 1999; Murillo-Zamorano and Petraglia, 2011; Godoy-Caballero and Murillo-Zamorano, 2012; Cordero et al., 2014). The quality instruments used in the present study are directly provided by the user of the health service and, in consequence, informing about the quality perceived by the patients, based on their health status, and not on the way the system is operating measured with the use of activity indicators.

Secondly, we perform a comparative analysis of the different PROMs. Most of the existing literature has focused on the use of only one of the questionnaires, without comparing them and, a reduced number of papers have performed comparative analyses, but never in order to assess the performance of the different health units. To the best of our knowledge, this is the first research that compares PROMs with the aim of evaluating

the quality with which health services are being delivered, being this, the idea of present research.

Finally, we obtain an estimation of the hospital performance with which we can recognise whether differences between centres exist. This approach identifies those units that require to be studied in the detail because they have an "unusual performance". Consequently, it allows us to answer the question of whether the approach taken to judge hospitals depends on the indicators used to make that judgement, and more specifically, if their quality differs based on whether a generic or a specific measurement instrument is used.

The structure of the chapter is as followed. Firstly, we describe the main characteristics, advantages and disadvantages of each of the PROMs, both generic and specific. Secondly, we present the empirical approach followed, which refers to a multilevel methodology, and that has been applied for one of the interventions for which the NHS collects PROMs, hip replacement. In relation to the multilevel model, we have used a novel methodology that allows us to obtain an estimation of hospital effect that can be used as an approximation of hospital performance (Jacobs et al., 2006). Thirdly, we show the characteristics that describe our sample as well as its main descriptive statistics. The information used involves a total of 20,509 patients undergoing hip replacement and treated in 153 hospitals during 2009. Fourthly, we display the results of the application of the multilevel methodology for each of the analyses considered. Finally, we present discussion and main conclusions of the study.

2. CONCEPTUAL FRAMEWORK

This section is devoted to the description of the main characteristics of the Patient Reported Outcome Measures (PROMs), as well as the advantages and disadvantages of each of them both generic and disease-specific instruments.

As mentioned in the introduction, the English Department of Health defines PROMs as "self-completed questionnaires administered to patients to assess their self-

reported health status before and after certain elective healthcare interventions funded by the NHS" (Department of Health, 2008). The consideration of these measures is based on the idea that the best source of information of how a patient feels is the patient himself (Devlin et al., 2010).

Different authors have also defined PROMS as measurements of any aspect of a patients' health status, obtained directly from the patients, i.e. without the help of physicians or other observers (Ousey and Cook, 2001; Valderas et al., 2008; Valderas and Alonso, 2008; Wylde et al., 2009). Given that PROMs consider the patient's view, they assist to an increase in their participation in health care (Marshall et al., 2006).

Across PROMs we can distinguish between generic and disease-specific measures. The generic measures of health-related quality of life are being collected for all procedures in the PROMs survey. They offer the advantage of allowing for comparisons of hospitals for individual procedures as well as across interventions given that they do not contain any question referring to a particular disease. Despite this positive aspect, the generic measures present some disadvantages. For example, the items they include are broader and not directly related to the condition. Because of that, patients' answers may include health aspects not related to the surgery for which the questionnaire has been completed.

On the other side, the disease-specific measures are particular for each procedure. They are hypothesised to be more sensitive to changes in health status within a given procedure, as they only consider information for the particular disease they analyse. Therefore, they can complement the information collected by the generic PROMs, helping to examine that any relevant aspect of patient health related to the medical condition for which patient has received treatment has not been missed (Devlin and Appleby, 2010). This implies that we can only use them to make comparisons within a particular procedure and not across patients diagnosed with different conditions (Devlin and Appleby, 2010).

Currently in the NHS both, generic as well as disease-specific measures are being collected. The generic measures are the European Quality of life-5 Dimensions (EQ-5D)

and the European Quality of life-Visual Analogue Scale (EQ-VAS), while the specific measures for each of the interventions are: Oxford Hip Score (OHS), Oxford Knee Score (OKS) and Aberdeen Varicose Vein Questionnaire (AVVQ), respectively for hip and knee replacement and varicose vein surgery, not having any specific measure for the case of hernia repair. Here we present the main characteristics for each of them.

The EQ-5D descriptive system, presented in Appendix 1, measures patients' self-reported health-related quality of life in terms of five health domains: mobility, self-care, usual activity, pain/discomfort and anxiety/depression. For each of these domains, patients indicate the degree of problems they experience using a three-point scale, where 1 indicates that the patient does not present problems in relation to that domain; 2, refers to some problems, and 3, to extreme problems. As a result, the EQ-5D can describe 243 different health states, where the health state defined as 11111 (which would correspond to a person with no problems in any of the dimensions) reflects full health, and 33333 is the worst possible health state, given that it would be indicating that the patient presents extreme problems in all the five dimensions considered in the questionnaire. This health profile can be translated into a weighted utility score using the UK population weights (Dolan, 1997). The resulting EQ-5D index is a cardinal measure that ranges from one, representing perfect health, to -0.594, where zero represents a state equivalent to being dead and utility scores lower than zero represent health states worse than being dead.

The use of EQ-5D to measure patient health and health-related quality of life has the advantage of being simple in use, responsive to change and reliable (Hurst et al., 1997). However, the disadvantages are that it can lead to losses of information when obtaining the EQ-5D index from the EQ-5D profile (Devlin and Appleby, 2010; Gutacker et al., 2013). For example, the differences between scores can be related to a particular dimension and it may be interesting knowing in which dimension the differences in health arise. This aspect cannot be determine when we work with the EQ-5D index given that all the information relative to a patient is gathered in a value between -0.594 and 1.

The EQ-5D also contains a visual analogue scale (EQ-VAS). The EQ-VAS, presented in Appendix 2, can be defined as a measure of the patient's valuation of their own global health status. This scale ranges from zero to one hundred where zero is the worst health state that a patient can imagine and 100 is the best imaginable health state. Patients are asked to report their health-related quality of life by indicating the point on the scale that reflects their current health state. Despite the simplicity of the questionnaire, the EQ-VAS is not based on a utility theory and, therefore, it is not much used in the scope of economic analysis.

Following the terminology used by the English Department of Health (DH), we will refer to the EQ-5D descriptive system simply as the EQ-5D and will treat the EQ-5D descriptive system and the EQ-VAS as two independent measures (NHS, 2011)[17].

Moving on to the specific PROMs, and given the similarities between the OHS and the OKS, we jointly present the characteristics of these two indicators, highlighting their main differences. These two questionnaires, presented in Appendix 3 and 4, are designed to evaluate disability in patients undergoing total hip (OHS) and knee (OKS) replacement respectively (Dawson et al., 1996; Dawson et al., 1998; Wylde et al., 2009). Each questionnaire contains a total of twelve questions about pain and physical limitations which have existed during the past four weeks due to the hip or knee. Ten of the questions are identical in the OHS and OKS, while the remaining two are specific to the condition. Each of the questions has five categories of response, from least to most difficulty or severity, resulting in more than 244 million possible health states (Oppe et al., 2011). Each of the answers results in an item score between zero and four and the total score is obtained by adding the individual item scores. Thus, the total score ranges from zero to 48 were lower scores indicate higher disability. Both measures have been found to be "practical, reliable, valid and sensitive to clinically important changes over the time" (Dawson et al., 1998).

[17] Note that this is at odds with the terminology used by the EuroQol group who developed and maintain the EQ-5D.

Finally, the last specific questionnaire (presented in Appendix 5) focuses on the analysis of the aspects related to the varicose vein surgery, and consists of 13 questions. The responses can be aggregated into an index taking values from zero to 100 (Garratt et al., 1993) using weights provided by the developers. A score of zero is defined to be the best health state and higher scores reflect worse health states.

3. METHODS

The methods section refers, firstly, to a multilevel model that allows us to analyse the effect of a series of hospitals using the information provided by patients in the different PROMs. Secondly, it describes the process followed in order to identify those hospitals with a different behaviour than the average hospital has, and that have been catalogued as "outliers" (Gutacker et al., 2013).

In relation to the empirical model, we have information about patients grouped in a series of hospitals, leading to a hierarchical data structure. Multilevel models are commonly used when the data to be analysed fall into a "hierarchical structure consisting of multiple macro units and multiple micro units within each macro units" (Rice and Jones, 1997). Our data present this multilevel structure, as we have many hospitals and many individuals within each hospital. Differently to traditional multilevel modelling, here we apply a fixed effect multilevel modelling that allow us to estimate the hospital effect and that can be used as an indicator of organisational performance (Jacobs et al., 2006).

We estimate multilevel models with individual patient characteristics and hospital fixed effects, which will allow us to investigate whether some hospitals have a differential performance than others based on the outcome data provided by patients within different hospitals. Furthermore, this methodology allows us to account for the observed heterogeneity in patient characteristics rather than rely on aggregate data about patient severity at hospital level (Laudicella et al., 2010).

Previous literature has applied multilevel models following this methodology. Among them, we can mention Street et al. (2012) and Laudicella et al. (2010). Street et

al., (2012) use fixed effect multilevel models to explain why the resource use (costs or length of stay) differs among patients and hospitals. In order to do it, they make use of a two-stage model. In the first stage they analyse the influence of a set of patient level explanatory variables on individual resource use and extract hospital fixed effects. In the second stage, they analyse these fixed effects to identify hospital level factors that are associated with performance variation.

Laudicella et al. (2010) use the same approach to examine to what extent costs of English obstetrics departments are explained by the characteristics of patients admitted to their respective diagnosis related groups (DRGs). After controlling for those characteristics, they analyse why some departments have higher costs than others. As in the previous case, estimates of departmental fixed effects are expected to reflect the relative performance of each department with values above the national average indicating worse performance, as they would represent higher average costs.

Following these two approaches we estimate three multilevel models with fixed effects. To our concern, this is the first time that this type of analysis is applied to the scope of patient reported outcome measures. This issue presents the utility of allowing for the evaluation of the hospitals performance based on how patients' health status has been altered as a consequence of a particular intervention.

As dependent variables we use the change in patient health or health-related quality of life as measured by the different PROM instruments, specifically, for EQ-5D (equation 1), EQ-VAS (equation 2) and OHS (equation 3). The change experienced by each of the indicators is calculated as the difference between the post and pre-surgery questionnaires. As previously mentioned, we are interested in the unexplained variation at hospital captured by the hospital fixed effect. Because we control for patient characteristics, any remaining variation at provider level (the "hospital effect") can be interpreted as systematic variation in performance.

As regressors in the different models we use a set of explanatory variables that reflect a series of patients' sociodemographic characteristics and of their health status, an error term to represent the fixed effect of each hospital and the random error term.

Among variables describing patients' sociodemographic characteristics we include age (age), which refers to patients older than 15; the square term of age (age^2) in order to allow for a non-linear relationship between age and change in health status; sex, which is a dummy variable taking value 1 if the patient is male and 0 if female; the variable revision, indicating if the intervention, for which the PROM is being collected, refers to a revision from a previous surgery; a series of variables reflecting other diseases or comorbidities that the patient may present (shown in Table 1), and finally, we include the initial health status reported in the pre-operative questionnaire.

The rationale for this is that not all patients are likely to improve to the same extent. Potentially, those patients who are in worse health before the surgery are more likely to experience a greater change as surgery can have higher impacts on those individuals. Similarly, those patients who report to be healthier at the baseline cannot improve to the same extent given what has been termed a "ceiling effect" and the "inability of the scores to detect top-end differences" (Baker et al., 2012).

The final specification of each of the three models with their corresponding explanatory variables is as follows:

$$changeEQ5D_{ij} = \beta_1 age_{ij} + \beta_2 age^2{}_{ij} + \beta_3 sex_{ij} + \beta_4 preEQ5D_{ij} + \beta_5 revision_{ij} + \beta_6 AMI_{ij} + \beta_7 CHF_{ij} + \beta_8 PVD_{ij} + \beta_9 CD_{ij} + \beta_{10} Dem_{ij} + \beta_{11} COPD_{ij} + \beta_{12} Reuma_{ij} + \beta_{13} PED_{ij} + \beta_{14} MLD_{ij} + \beta_{15} Dia_{ij} + \beta_{16} DiaCom_{ij} + \beta_{17} HP_{ij} + \beta_{18} RD_{ij} \ldots + \beta_{19} Cancer_{ij} + \beta_{20} MSLD_{ij} + \beta_{21} MC_{ij} + \beta_{22} AIDS_{ij} + u_j + e_{ij} \quad [1]$$

$$changeEQVAS_{ij} = \beta_1 age_{ij} + \beta_2 age^2{}_{ij} + \beta_3 sex_{ij} + \beta_4 preEQ5D_{ij} + \beta_5 revision_{ij} + \beta_6 AMI_{ij} + \beta_7 CHF_{ij} + \beta_8 PVD_{ij} + \beta_9 CD_{ij} + \beta_{10} Dem_{ij} + \beta_{11} COPD_{ij} + \beta_{12} Reuma_{ij} + \beta_{13} PED_{ij} + \beta_{14} MLD_{ij} + \beta_{15} Dia_{ij} + \beta_{16} DiaCom_{ij} + \beta_{17} HP_{ij} + \beta_{18} RD_{ij} \ldots + \beta_{19} Cancer_{ij} + \beta_{20} MSLD_{ij} + \beta_{21} MC_{ij} + \beta_{22} AIDS_{ij} + u_j + e_{ij} \quad [2]$$

$$changeOHS_{ij} = \beta_1 age_{ij} + \beta_2 age^2{}_{ij} + \beta_3 sex_{ij} + \beta_4 preEQ5D_{ij} +$$
$$\beta_5 revision_{ij} + \beta_6 AMI_{ij} + \beta_7 CHF_{ij} + \beta_8 PVD_{ij} + \beta_9 CD_{ij} + \beta_{10} Dem_{ij} +$$
$$\beta_{11} COPD_{ij} + \beta_{12} Reuma_{ij} + \beta_{13} PED_{ij} + \beta_{14} MLD_{ij} + \beta_{15} Dia_{ij} +$$
$$\beta_{16} DiaCom_{ij} + \beta_{17} HP_{ij} + \beta_{18} RD_{ij} \ldots + \beta_{19} Cancer_{ij} + \beta_{20} MSLD_{ij} +$$
$$\beta_{21} MC_{ij} + \beta_{22} AIDS_{ij} + u_j + e_{ij} \quad [3]$$

The subscript i refers to the individual patient while the subscript j is used to identify the hospitals. The term u_j is the fixed effect for the j^{th} hospital. This is the term we are interested in, as it reflects the performance of each individual hospital j. In our case, higher values of this term will indicate that such a particular hospital is performing better than another one with a smaller value of u_j, given that patients express a higher scores in the different questionnaires. The term e_{ij} denotes the random error. This error term is assumed to fulfil the standard properties of the disturbances (Wooldridge, 2009): random disturbances, homoscedasticity, no serial correlation and normal distribution.

Table 1. Comorbidities used in the estimation of the models

Variable	Description
AMI	Acute Myocardial Infarction
CHF	Congestive Heart Failure
PVD	Peripheral Vascular Disease
CD	Cerebrovascular Disease
Dem	Dementia
COPD	Chronic Obstructive Pulmonary Disease
RD	Rheumatoid Disease
PED	Peptic Ulcer Disease
MLD	Mild Liver Disease
Dia	Diabetes
Dia + Com	Diabetes + Complications
H/P	Hemiplegia or Paraplegia
RD	Renal Disease
Cancer	Cancer
M/SLD	Moderate/Severe Liver Disease
MC	Metastatic Cancer
AIDS	AIDS

Source: Prepared by authors

After the estimation of the different hospital effects we identify those hospitals with an "unusual" performance and that have been termed as "outliers". These outliers are hospitals that perform differently from the average with respect to at least one of the

PROMs considered. It is important to note that some hospitals may only be considered outlier on one PROM, whereas others may be outliers on both PROMs which are being compared.

The procedure we have used to identify these outliers is as follows (Gutacker, 2013): we first obtain an estimate of the hospital effect together with its 95% confidence intervals. After that, we defined a variable that only takes values for each observations for which the confidence intervals contains the value zero and that, therefore, are not statistically different from the average, i.e. they are not outliers. Hospital which confidence intervals do not include the value of zero, are classified as positive or negative outliers, depending on whether their values are all greater or lower than zero, respectively. Hence, positive outliers are represented by those hospitals which performance is above the average, while for negative outliers, the performance is below the average.

4. DATA

In this section we present the main characteristics of the data used in the analysis, as well as its descriptive statistics. As previously mentioned, data refer to patients that have undergone hip replacement. Hence, the questionnaires used in the analysis are the generic EQ-5D and EQ-VAS and the specific OHS.

The sample is extracted from the English National Health Service (NHS) database which contains the PROMs questionnaires completed by patient who have undergone the hip replacement surgery between April 2009 and March 2010. These patients are asked to complete the pre-operative questionnaire before being admitted for the surgery and the post-operative questionnaire is sent to them six months after the intervention so they can returned it once completed.

Our data consist of 20,509 patients grouped in 153 hospitals. Out of these 20,509 patients, 41.45% (n = 8,501) are males while the 58.55% (n = 12,008) are females. The age of patients ranges from 15 to 94, with an average age of 68 years. From all of them, 7.40%

(n = 1,516) of patients are characterised for undergoing a revision of a previous hip replacement.

We now present a summary of the main descriptive statistics for the PROM responses both, at the individual level and the hospital level. At the patient level (Table 2), the mean pre-operative score is 0.357 for EQ-5D, 66.36 for EQ-VAS and 18.38 for OHS. For the post-operative scores all the mean values experience an improvement. These values are, 0.765, 75.52 and 38.16 respectively for, EQ-5D, EQ-VAS and OHS. This improvement is particularly big in the case of OHS and EQ-5D, with an increase of 107.62% and of 114.29%, respectively. In relation to the generic EQ-VAS, there are also some improvements; however, these are substantially smaller (13.80%).

Table 2. Descriptive statistics at patient level

Patient level	EQ-5D		EQ-VAS		OHS	
	Pre-op	Post-op	Pre-op	Post-op	Pre-op	Post-op
Mean	0.357	0.765	66.36	75.52	18.38	38.16
Minimum	-0.594	-0.594	0	0	0	0
Maximum	1	1	100	100	48	48

Source: Prepared by authors

An interesting aspect to be considered in the analysis of the individual level data is related to the number of patients that stated to be in perfect health before and after the surgery under each of the measurement instruments (Table 3). In the case of the specific OHS, the proportion of people that indicated to be in perfect health, i.e. obtaining a score of 48 in that questionnaire, increases from 0.05% of surveyed users before the intervention to 11.67% after it. For the generic EQ-5D, this difference changes from 3.12% before the surgery to 35.62% of patients obtaining, after the hip replacement, a score of one in the EQ-5D index. In the case of the EQ-VAS scale, there is also an increase in this percentage, although it is not as pronounced as in the other cases: 1.77% before the surgery and 4.55% of patients pointing a value of 100 in the scale after the hip replacement.

Table 3. Proportion of patients reporting to be in perfect health

	Pre-intervention	Post-intervention
EQ-5D	3.12	35.62
EQ-VAS	1.77	4.55
OHS	0.05	11.67

Source: Prepared by authors

The differences between PROMs at patient level are also reflected in the provider level statistics (Table 4). The mean scores of each of the questionnaires at the hospital level are similar to the ones obtained at the patient level. For the case of the pre-operative questionnaires, these scores are 0.346 for EQ-5D, 65.54 for EQ-VAS and 18.02 for OHS. Focussing on post-operative answers, we can observe that all the mean values assist to an improvement, especially relevant for the case of EQ-5D (increasing to 0.753 points – 109.05%) and OHS (increasing to 37.67 points – 117.63%). This change is a bit smaller when we analyse the generic EQ-VAS, which increases in a 13.78%, having a mean hospital score of 74.57 for the post-operative questionnaire.

Table 4. Descriptive statistics at hospital level

Hospitals	EQ-5D		EQ-VAS		OHS	
	Pre-op	Post-op	Pre-op	Post-op	Pre-op	Post-op
Mean	0.346	0.753	65.54	74.57	18.02	37.67
Minimum	0.185	0.604	55.65	60.00	13.67	29.30
Maximum	0.639	0.908	72.78	81.50	29.00	44.50

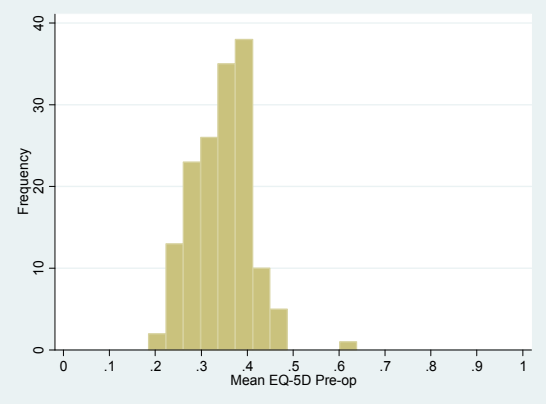
Source: Prepared by authors

Improvements commented above are also reflected in the minimum and maximum values of each of the questionnaires (Table 4). For example, for the case of OHS, the minimum pre-operative value is 13.67, while after the intervention this value increases to 29.30. The same happens for the maximum scores, changing from 29 before the surgery to 44.50 after it.

Another aspect to consider and that it is also reflected in the previous table, refers to the difference between the minimum and maximum scores in the responses before and after the medical intervention. Scores are more dispersed before surgery than after surgery. For example, the difference between the minimum and maximum values on the EQ-5D is 0.454 points, while after surgery this reduces to 0.304. This reduction in variance is also observed for the OHS, albeit less pronounced. However, in the case of EQ-VAS, the difference between the minimum and maximum value after surgery is bigger than before

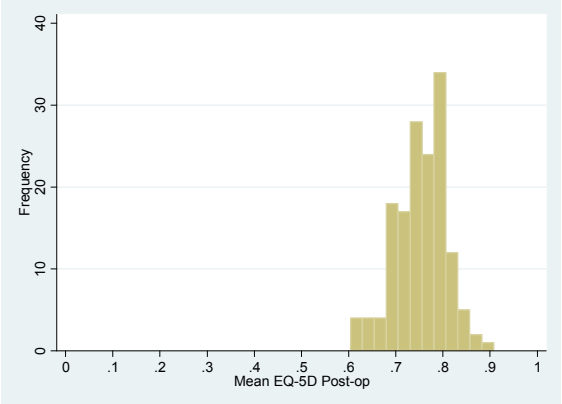
surgery. These facts can be also observed in the histograms for the pre-operative and post-operative scores presented in Figures 1 to 6. In these figures we can also observe the general improvement experienced by the mean values in the different questionnaires.

Figure 1. Histogram for EQ-5D. Pre-op



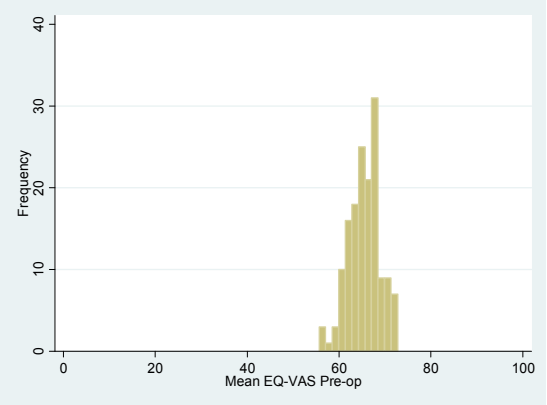
Source: Prepared by authors

Figure 2. Histogram for EQ-5D. Post-op



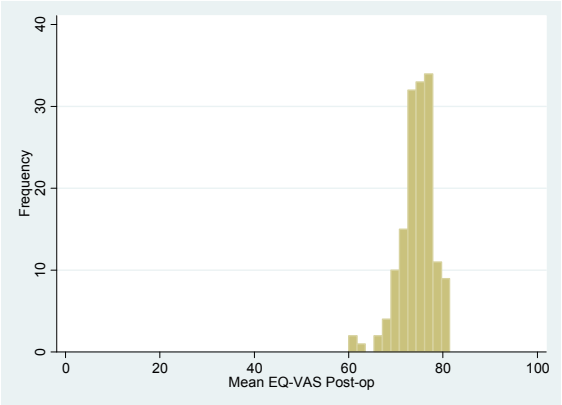
Source: Prepared by authors

Figure 3. Histogram for EQ-VAS. Pre-op



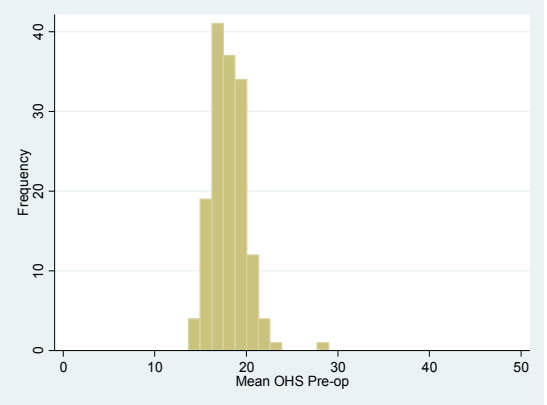
Source: Prepared by authors

Figure 4. Histogram for EQ-VAS. Post-op



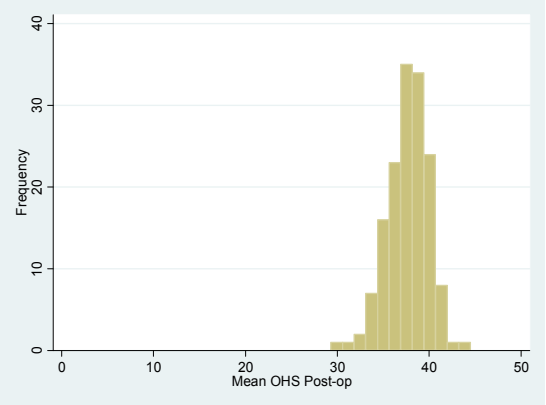
Source: Prepared by authors

Figure 5. Histogram for OHS. Pre-op



Source: Prepared by authors

Figure 6. Histogram for OHS. Post-op



Source: Prepared by authors

Once we have commented the mean, minimum and maximum values of the different indicators, we next present the correlations between PROMs, both at the individual and at the patient level (Tables 5 and 6). The correlations showed in the tables refer to the answers given to the pre-operative questionnaires and to the post-operative questionnaires.

Table 5. Correlation between PROMs at the patient level

Patients	Pre-op	Post-op
OHS and EQ-5D	0.7325	0.7617
OHS and EQ-VAS	0.3783	0.5982
EQ-5D and EQ-VAS	0.3586	0.6386

Source: Prepared by authors

Table 6. Correlation between PROMs at the hospital level

Hospitals	Pre-op	Post-op
OHS and EQ-5D	0.8904	0.9337
OHS and EQ-VAS	0.5496	0.7893
EQ-5D and EQ-VAS	0.5500	0.8282

Source: Prepared by authors

The highest correlations are observed between the indicators OHS and EQ-5D, both at the individual and at the hospital level. For the rest of indicators there also exist strong correlations being them higher when hospital data rather than individual data are considered. Previous literature has already found high correlations between generic PROMs such as the EQ-5D and specific ones such as the OHS (Oppe et al., 2011; Ostendorf et al., 2004; Dawson et al., 2001). Nevertheless, the correlations that we have obtained in this study are not only consistent with previous work; we have also reported higher scores than the existing literature.

Finally, we have also derived a series of indicators that record the percentage of patients who report about improvements or deteriorations in their health status after surgery or who did not experienced any changes. Patients improving as a result of the surgery are those ones for which the score in the post-operative PROM is greater than in the pre-operative questionnaire (indicator “better”). On the contrary, those with smaller values after the intervention are represented by the indicator “worse”. There is also a proportion of patient informing about the lack of changes as a result of the surgery and

are represented by the indicator “neutral”. These percentages, both at the patient and at the hospital level, are reported in Tables 7 and 8.

Table 7. Indicators better/neutral/worse at the patient level

Patients	Better	Neutral	Worse
OHS	96.07%	0.53%	3.40%
EQ-5D	87.45%	6.16%	6.39%
EQ-VAS	61.49%	11.49%	27.02%

Source: Prepared by authors

Table 8. Indicators better/neutral/worse at the hospital level

Hospitals	Better	Neutral	Worse
OHS	98.04%	0.00%	1.96%
EQ-5D	87.58%	7.84%	4.58%
EQ-VAS	60.79%	15.03%	24.18%

Source: Prepared by authors

More patients report improvements when outcomes are measured by the OHS instead of the generic measure. This may be because the OHS is measuring more specific aspects of health related to a particular intervention than the EQ-5D or EQ-VAS do, as the latter two refer to more general aspects. Therefore, the former measure may take into account smaller changes in patient’s health status, which are related to the surgery performed. Comparing the two generic measures, we observe that there are more people experiencing an improvement when considering the EQ-5D rather than EQ-VAS. Additionally, an interesting value is the percentage of people being worse-off when we consider EQ-VAS. This value is over 20% in both cases, at patient and at hospital level data.

5. RESULTS

In this section we present the results obtained from the estimation of the econometric models specified in equations 1, 2 and 3 specified above. First of all, we show the results of the estimation of the multilevel models, analysing the significant relations between our explanatory variables and the change in patients’ health status as measured by the different PROMs. After that, we present the results obtained from the comparison of the hospital effects between each of the generic and the specific

instruments considered in the analysis. Specifically, between EQ-5D and OHS, between EQ-VAS and OHS and, finally, between EQ-5D and EQ-VAS.

5.1 Estimation of the multilevel models

The coefficients obtained from the estimation of the multilevel models for each of the questionnaires, EQ-5D, EQ-VAS and OHS are presented in Table 9, showing that most of the variables used in the analysis are statistically significant.

The majority of non-significant variables are related to the different comorbidities that patients may have. Their significance varies depending on the questionnaire we use as dependent variable. Some of them are significantly associated with changes in health in all the estimations (chronic obstructive pulmonary disease, rheumatoid disease, mild liver disease and diabetes), others are significant in some estimations (congestive heart failure, peptic ulcer disease, diabetes and complications, hemiplegia or paraplegia, cancer and metastatic cancer) and a range of comorbidities are not significant in any of the estimations (acute myocardial infarction, peripheral vascular disease, cerebrovascular disease, dementia, renal disease or moderate/severe liver disease), indicating that the fact that the patients are diagnosed with any of these diseases does not influence their scores in the EQ-5D, EQ-VAS or OHS.

Looking at the rest of variables used in the estimation of the model, and starting with age, results show that both age effects (age and age^2) are significant, indicating the existence of a non-lineal relationship between age and the change in patients' health status. The variable sex is significant in two out of the three estimations, not showing any significant effect when the model is run using as dependent variable the change experience by the patient in the EQ-VAS ($p = 0.118$). For models that consider OHS and EQ-5D as dependent variables, we find that the coefficient for male gender is positive, indicating that men experience a greater change than women and, therefore, benefit more from surgery.

Table 9. Results obtained from the estimation of the multilevel models

	EQ-5D (eq. 1)	EQ-VAS (eq. 2)	OHS (eq. 3)
age	0,010***	0,441***	0,354***
age^2	-0,0008***	-0,004***	-0,003***
sex	0,017***	0,375	0,761***
pre-op score	-0,774***	-0,716***	-0,656***
revision	-0,111***	-4,990***	-6,106***
AMI	-0,122	-1,556	-0,805
CHF	-0,072***	-4,596**	-1,227
PVD	-0,025	-0,805	-0,188
CD	-0,029	0,043	-0,667
Dem	0,005	-2,877	0,880
COPD	-0,046***	5,022***	-1,480***
Reuma	-0,080***	-6,094***	-0,995**
PED	-0,064	-4,550	-3,050*
MLD	-0,138**	-10,014**	-3,460*
Dia	-0,049***	-4,417***	-1,939***
Dia+Com	-0,071	-6,061*	-3,241*
H/P	-0,036	-9,699**	-3,117
RD	-0,001	-1,524	0,397
Cancer	-0,027	-3,829**	-0,778
M/SLD	-0,135	-6,166	-1,848
MC	-0,064	-10,442**	-2,042
Adj. R^2 without hospital effect	0,519	0,445	0,299
Adj. R^2 with hospital effect	0,527	0,451	0,313

* p <0.05; ** p < 0.01; *** p < 0.001

Source: Prepared by authors

The pre-operative score has obtained a negative sign in each of the regressions, indicating that those patients being worse at the beginning (before the surgery) are experiencing a higher improvement that those ones being better. This is consistent with the results presented in Baker et al. (2012). However, as pointed out by these authors, the reason may be that those patients with a better pre-operative score are not able to improve to the same extent, given the existence of a ceiling effect. In any case, this information has important implications in policy terms, given that it allows us to identify who benefits more from the surgery to analyse why others do not benefit in the same way.

The negative sign of the he revision variable (informing about whether the intervention is because of a revision of a previous surgery) is indicating that those patients undergoing surgery as a revision procedure experience a smaller change than those ones who are admitted as a primary procedure.

Finally, with respect to the estimation, we computed the R^2 for two situations: with and without the hospital effects. This value represents the proportion of variation in the PRO measures explained by the variables we are using in the analysis, with and without the hospital effects. We can see from Table 9 that this value is affected by the incorporation of these effects, so there is a proportion of the change in the measures explained by the hospital effects.

The inclusion of the hospital effect not only improves the specification of the models, it also enables to perform a disaggregated analysis with which we can study the performance of each of the individual health units. This is a very relevant aspect given that the lack of its consideration could lead to the introduction of important biases in the results of the analysis, and consequently, in the policy decisions made based on them.

In the literature there exist some references that have carried out comparative analyses between the different PROMS (Baker et al., 2012; Oppe et al., 2011; Ostendorf et al., 2004; Dawson et al., 2001). However, in all of them those comparisons have been made in an aggregate way, without making any distinction based on the hospital where patient has been assisted. As we have just seen, the no consideration of each hospital as an independent unit, may introduce bias in terms of health policy given that the indicators obtained would not be representative all the population considered.

The inclusion of the "hospital effect", as considered in this research, it does not only allow us to perform a comparison of the different scores, but also to detect those units that have a performance significantly different to the average and that we catalogue as "outliers". Following this consideration, we now present a detailed analysis of the outliers found in each possible comparisons between the PROMs used in our study.

5.2 Comparison of hospital effects for OHS and EQ-5D

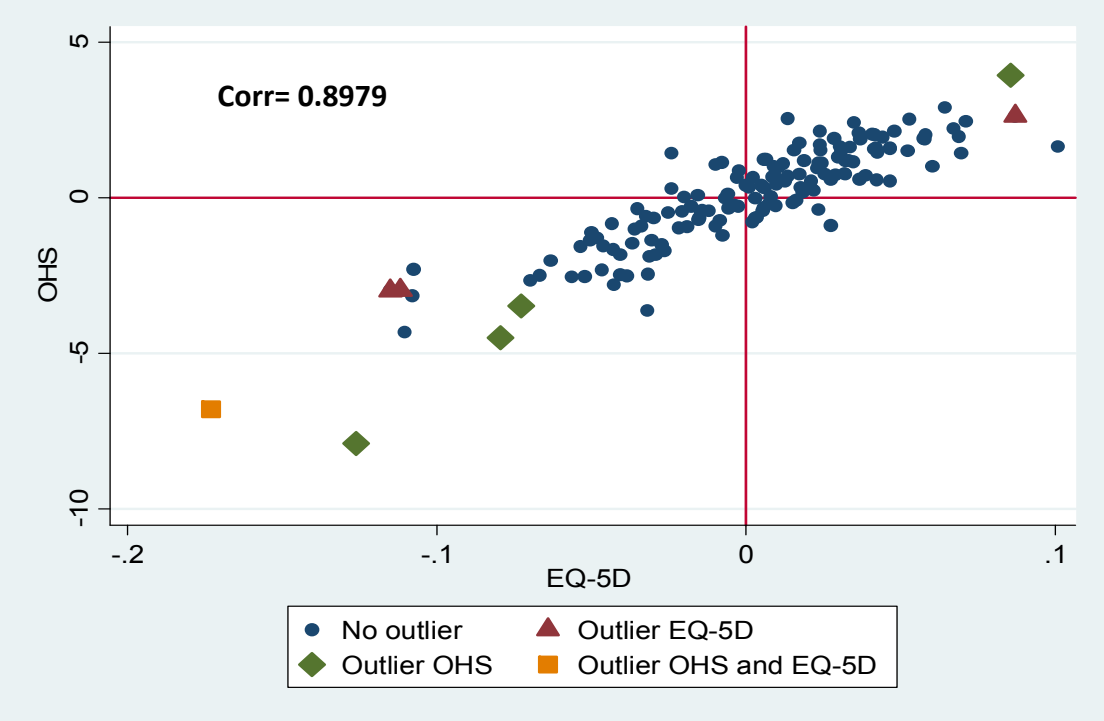
We now move on to the comparisons of the hospital effect (u_j) obtained from the estimations which dependent variables have been the specific questionnaire OHS and the generic EQ-5D. Performing this as a comparative analysis is more informative than presenting the hospital effects for each of the individual indicators. The comparison of

both indicators by its graphical representation will allow us to evaluate the performance of each centre, and with this to identify the unusual values or outliers.

The hospital effects for the case of OHS and EQ-5D are represented in Figure 7. We observe a strong positive association between both measures. This positive association indicates that a good performance on EQ-5D is associated with a good performance in OHS. The Pearson correlation coefficient is 0.8979. This strong correlation between the hospitals effects analysed with the two measures is consistent with the correlations obtained before the estimation of the models. Previous literature also found high correlations, but not with coefficients as high as the ones in our analysis. Furthermore, the values obtained are superior to previous ones. This is not only because the coefficient obtained in this analysis is higher in value than the found in the existing literature, but also because it has been calculated considering the hospital effects, which allows to make better decisions and reduce the biases of considering the whole population as an aggregate unit.

From the analysis of the hospital effects, we find eight hospitals behaving significantly different to the average hospital, and that therefore, are outliers, obtaining scores below or above the average performance. These particular hospitals are highlighted in Figure 7 by using different markers. Three of these observations are outliers only with respect to EQ-5D, performing on the average under the consideration of the OHS. One of them is above the average while the other two present values below the average of the set of hospitals. With respect to OHS we find four outliers, three of them refer to centres in which patients consider their health status as being worse than the average, while the third one present the contrary situation, its performance is better than the average hospital. Finally, there is one hospital which is catalogued as a negative outlier under the consideration of both, the EQ-5D and the OHS.

Figure 7. Hospital fixed effects for OHS and EQ-5D



Source: Prepared by authors

This analysis of the outliers according to EQ-5D and OHS can be summarised in Table 10. In this table we can see the number of outliers identified separately for each of the measures as well as the ones identified by both measures.

Table 10. Outliers in OHS and EQ-5D.

		OHS			TOTAL
		+ outlier	mean	- outlier	
EQ-5D	+ outlier	0	1	0	**1**
	mean	1	145	3	**149**
	- outlier	0	2	1	**3**
TOTAL		**1**	**148**	**4**	**153**

Source: Prepared by authors

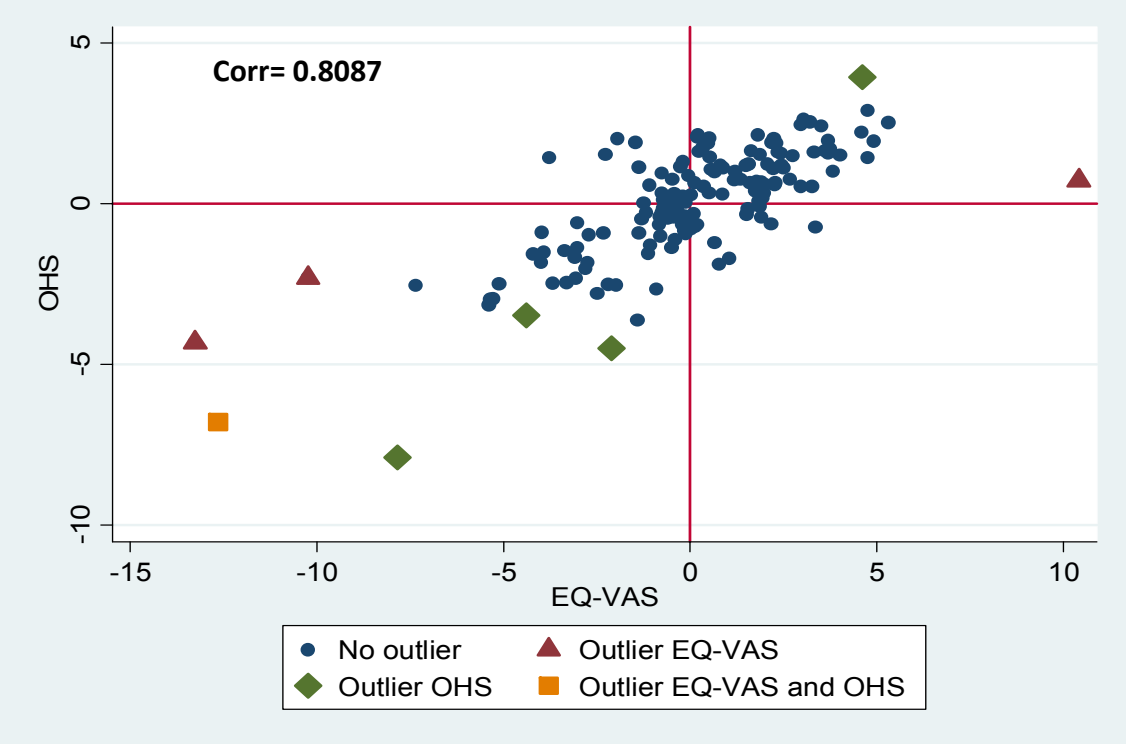
5.3 Comparison of hospital effects for OHS and EQ-VAS

As in the previous case and as a result of the comparative analysis of the indicators OHS and EQ-VAS, we obtain the existence of a high and positive association between these two questionnaires. The value of the correlation coefficient is 0.8087, indicating that in general, those patients reporting to have a good health related quality of life in the generic EQ-VAS do the same under the evaluation of the specific OHS, and vice versa.

Even if the correlation is weaker than before there is still a strong correlation, and higher than the correlations found at the individual level in the previous literature.

In spite of this general concordance we have also identified several outliers following the same procedure described before. This procedure obtained a total of eight outliers (Figure 8). We have found one hospital being a negative outlier according to the two measures. Three hospitals present an unusual performance under the consideration of the EQ-VAS being on the average according to OHS. One of those hospitals are performing above the average with respect to EQ-VAS (positive outlier) while the other two are performing below the average, indicating that, patients report that their health status as measured by the EQ-VAS is below the average hospital. Finally, four centres are valued according to the average under the consideration of the EQ-VAS, but not when the OHS is considered. One of them presents above the average scores on the specific questionnaire, while for the remaining three, those scores are below the mean, being negative outliers.

Figure 8. Hospital fixed effects for OHS and EQ-VAS



Source: Prepared by authors

Table 11 exhibits a summary of those hospitals that have been catalogued as outliers based on the comparison of EQ-VAS and OHS, as well as those that are outliers in each of the measures independently.

Table 11. Outliers in OHS and EQ-VAS

		OHS			TOTAL
		+ outlier	mean	- outlier	
EQ-VAS	+ outlier	0	1	0	**1**
	mean	1	145	3	**149**
	- outlier	0	2	1	**3**
TOTAL		**1**	**148**	**4**	**153**

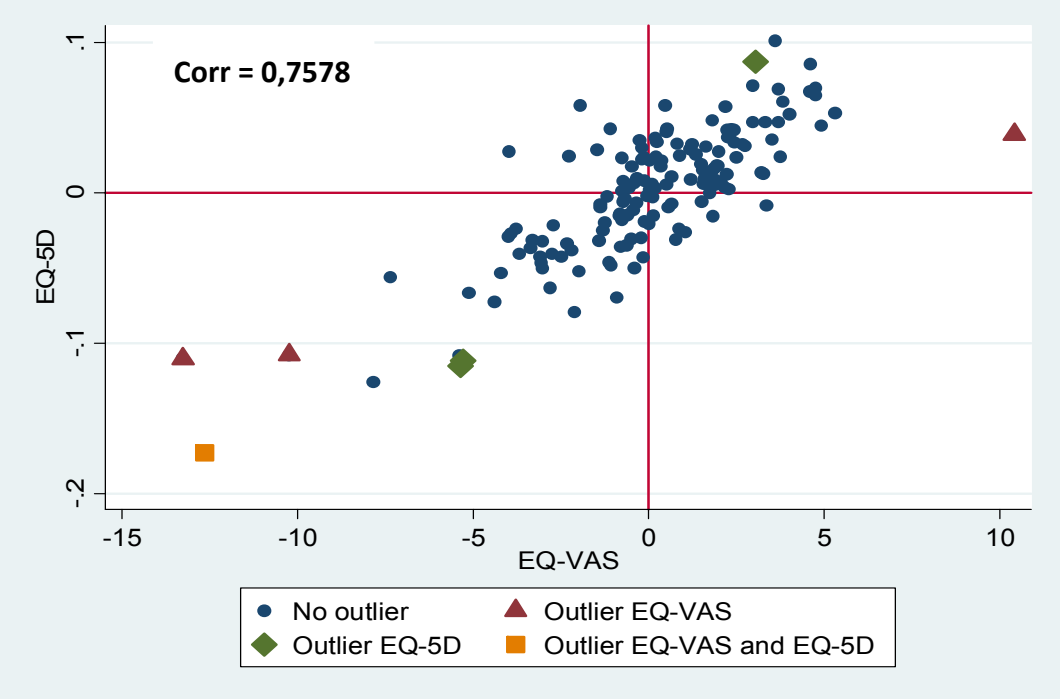
Source: Prepared by authors

5.4 Comparison of hospital effects for EQ-5D and EQ-VAS

Finally, we focus on the comparison between the two generic questionnaires, EQ-5D and EQ-VAS. The hospital effects for the two generic measures are presented in Figure 9. The association between the two generic measures, despite being a bit smaller than the association of each of the generics with the disease-specific, is still high; being the correlation between the two measures 0.7578.

Again, we have identified several outliers following the same procedure described in the methods section. In this case we have identified a total of seven outliers. As before, we have found one hospital being a negative outlier according to the two measures, EQ-5D and EQ-VAS. The rest of centres, present this unusual performance only in relation to one of the indicators. Specifically, there are three outliers with EQ-5D and three outliers with EQ-VAS. In each case, one of the three hospitals are performing above the average with respect to one of the measures (positive outlier) while with respect to the other measure they are performing on the average. The other two hospitals are negative outliers in one of the measures and are performing on the average with respect to the other questionnaire used in the comparison.

Figure 9. Hospital fixed effects for EQ-5D and EQ-VAS.



Source: Prepared by authors

The analysis of the outliers for the comparison of the generic measures is summarised in Table 12. This table shows the number of outliers that each of the indicators has identified separately as well as the outliers identified by both PROMs simultaneously.

Table 12. Outliers in EQ-5D and EQ-VAS

		EQ-VAS			TOTAL
		+ outlier	mean	- outlier	
EQ-5D	+ outlier	0	1	0	**1**
	mean	1	146	2	**149**
	- outlier	0	2	1	**3**
TOTAL		**1**	**149**	**3**	**153**

Source: Prepared by authors

Once we have identified the outliers in each of the comparisons between the generic and the disease specific measures and to finish with the analysis of these centres with an unusual performance, we present the similarities existing between those analyses.

In the first two analyses (OHS vs. EQ-5D and OHS vs. EQ-VAS) we found that there are four hospitals being outliers according to OHS but performing on the average in both

cases, when the comparison is made against EQ-5D and when the comparison is made with EQ-VAS. From an analysis in detail of these four hospitals we found that they are the same in both comparisons. Therefore, none of the outliers that the specific measure identifies are considered as such under the generic measures EQ-5D and EQ-VAS.

Contrary to this, we also identified three hospitals performing on the average with respect to OHS and being outliers according to the generic measures. These three hospitals are different when we consider EQ-5D than when we use EQ-VAS as the comparative measure for OHS. These three outliers identified in each of the generic measures are the outliers we mentioned before when we analysed the comparison between the two generic measures, three in relation to EQ-5D and 3 under the consideration of EQ-VAS.

Finally, in each of the three comparisons, we found the existence of one hospital performing below the average regardless whether the comparison was made between generic and specific measures or between the two generic EQ-5D and EQ-VAS. We considered this aspect in more detail and found that that observation refers to the same hospital in the three analyses. This particular hospital is performing below the average according to all the measures we have studied. In case we had carried out an aggregated analysis we would have ignored that particular centre, with the possibility of introducing important biases.

In this section we have presented the results obtained from the estimation of the different multilevel models. Those models have allowed us to analyse whether the different scores provided by patients in each of the questionnaires used differ depending on the indicator we consider. This empirical analysis has shown that, in general, high scores in one of the indicators are associated with high scores in the other measure used in the comparison, obtaining correlation coefficients that have been greater than 75% in all the cases.

Despite this common pattern, we have found a set of hospitals that we have catalogued as outliers, for which that general relationship did not exist. In view of the

results we could consider that the number of outliers that we found (between seven and eight depending on the comparison) is reduced and that, any decision would not have a big impact on them. Nevertheless, these outliers represent between 7% and 8% of the sample. Given that the sample is representative of the population and that the NHS assist more than 120 million cases every year, such apparently reduced proportion could be representing around 9 million consultations per year. Any decision making process that could ignore the outliers could have important implications, as it would not only be affecting patients in the sample, but also the total number of patients treated in those hospitals. The lack of consideration of those unusual observations could introduce an important bias which could have important implications in policy terms.

6. DISCUSSION AND CONCLUSION

The inclusion of PROMs in the analysis of hospital quality performance allows for the consideration of patients' perspective. Unlike others indicators that only inform about the objective quality, these questionnaires allow for the inclusion of certain aspects that reflect how patients feel as result of the treatment received. Such opinion provides with valuable information about patients' health related quality of life and allow to, in the last instance, offering a better quality in the provision of health care.

Nevertheless, the problem we have to face is that we have no gold standard measurement to which we can compare alternative measures. Rather we have disease-specific measures which are, a priori, hypothesised to be more sensitive to changes in the disease condition, and generic measures that present the advantage of allowing for comparisons between interventions, although with the drawback of not being able to reflect small changes experienced by patients in their health status.

Our objective is then to compare the three measures that the English National Health Service (NHS) currently uses and to observe whether the conclusion drawn differ depending on the instrument used, i.e. what is the general correlation between the results provided by each of the measures? Can we identify different outliers based on which indicator we are using?

In order to answer those questions, we have carried out an empirical analysis using the aforementioned indicators with the aim of analysing if the responses on one of the questionnaires are consistent with the ones provided in another one. That is, if a high score in the health status measured by one of the indicators correspond with high scores in the other indicators.

We estimated three multilevel models with fixed effect, one for the PROM EQ-5D, another one for EQ-VAS and a third one for the OHS. This methodological approach unlike a traditional multilevel analysis, allow us to obtain an estimation of hospital effect that can be used as an approximation of hospital performance. In each of these models, the dependent variable has been obtained as the change in the score of every measure, calculated as the difference between the post-operative (questionnaire completed before surgery) and pre-operative (questionnaire completed after surgery) scores. As explanatory variables we used a set of variables describing patients' sociodemographic characteristics as well as of their health status.

The multilevel approach developed in this research has allowed us to perform a disaggregated analysis, in which instead of analysing the answers given by all the patients to the questionnaires, we have been able to consider the effect of each hospital separately and, with that, to perform comparisons between the measures at the level of hospital instead of considering the data aggregated at the patient level. This aspect supposes an important contribution to the existing literature, given that the studies comparing generic and specific PROMs have considered the set of patients as a whole, without any reference to the centres where patients are treated, ignoring with that, any biases that this aspect could introduce.

After the estimations of the models, we performed a comparative analysis to evaluate the hospital effects in detail. As we have already mentioned, the comparison of PROMs has been very scarce and when it is has been done, it has not been with the purpose of analysing the performance of the different health units. We have carried out three comparisons: between the generic EQ-5D and the specific for total hip replacement OHS, between EQ-VAS and OHS and between the two generic EQ-5D and EQ-VAS. From

those comparisons we have found that, in general, when we analyse the performance assessment using generic measures and disease-specific instruments the correlation is high. Specifically, we obtained values above 70% in all the cases. These correlations were also above the ones found in previous literature which reported values of around 50%. The coefficients obtained in this study do not only improve the ones found by the existence literature in terms of value, they have the potential of having been obtained after the application of the multilevel models allowing us to consider each hospital as an independent unit.

However, we have found the existence of a set of observations that we have catalogued as outliers, for which the general pattern was not observed. Without the application of the methodology here considered we would not have been able to identify these unusual observations as we would have not been able to perform the disaggregated analysis. To this respect, the results obtained and the conclusions drawn would have not been representative of all the population and we could have introduced biases in the analysis, as we would be making decisions that would not be appropriated for everyone using the health care.

One possible explanation for these differences in performance assessment may be that disease-specific measures are exclusive for a particular disease and hence, may be more sensitive to smaller changes in the patients' health status related to the particular disease for which it was collected. Consequently, these measures will be able to reflect smaller improvements or worsening that the generic measures cannot detect given that it considers wider aspects of the patients' health. In the case that the EQ-5D or EQ-VAS was able to reflect more specific changes, maybe the differences between generic and disease-specific measures would have been smaller and we would not have identified any outlier or the number of outliers identified would have been smaller. Another possible answer is that the generic measures can reflect aspects of patients' health status that are not related to the particular surgery we are considering, but that are affecting patients' general health. Because of all that, it seems reasonable to jointly consider generic as well as disease specific questionnaires.

In general, and based on the results obtained in the present research, it can be concluded that the consideration of patients' opinions is of high importance when the performance of the different units needs to be assessed given that they are the actual users of the service and towards who it should be organised.

Furthermore, an analysis that does not consider the aggregated information, but that considers every hospital as an independent unit is of high relevance. First of all it allows us to know and fulfil different existing needs and secondly, in terms of health policy, ensures that decisions are made considering all the units participating in the system, and not only the general behaviour that may not be representative of the wider population. Finally, the use of only one type of indicators, generic or specific, would not be reflecting all the available information. Therefore, the consideration of both types of measurements will allow us to get richer information in relation to the behaviour of all health units, and with that, to offer a better and individualised health care and adapted to every user participating in it.

As future research, a further study including all the procedures would show whether the results we found in the analysis of hip replacement are similar when considering generic and disease-specific instruments for the study of knee replacement as well as varicose vein surgery.

Additionally, the application of this methodology to primary health care would allow us to analyse the effect of the different primary care centres using the same procedure as applied here and by asking patients about any improvements or worsening in their health status after the treatment prescribed by their general practitioners.

REFERENCES

Baker, P.N., Jameson, S.S., Avery, P.J., Reed, M.R. and Gregg., P.J. (2012). Comparison of patient-reported outcome measures following total and unicondylar knee replacement. *Journal of Bone & Joint Surgery - British Volume,* 94 (7), pp. 919-927.

Chua, C.L., Palangkaraya, A. and Yong, J. (2010). A two-stage estimation of hospital quality using mortality outcome measures: an application using hospital administrative data. *Health Economics,* 19 (12), pp. 1404-1424.

Cordero, J.M., Crespo, E. and Murillo-Zamorano, L.R. (2014). The effect of quality and socio-demographic variables on efficiency measures in primary health care. *The European Journal of Health Economics,* 15 (3), pp. 289-302

Dawson, J., Fitzpatrick, S.F., Gundle, O., McLardy-Smith, P. and Murray, D. (2001). Evidence for the validity of a patient-based instrument for assessment of outcome after revision hip replacement. *Journal of Bone & Joint Surgery - British Volume,* 83 (8), pp. 1125-1129.

Dawson, J., Fitzpatrick, R., Murray, D. and Carr, A. (1998). Questionnaire on the perceptions of patients about total knee replacement. *Journal of Bone & Joint Surgery - British Volume,* 80 (1), pp. 63-69.

Dawson J, Fitzpatrick R, Carr A, and Murray D. (1996). Questionnaire on the perceptions of patients about total hip replacement. *Journal of Bone & Joint Surgery – British Volume*, 78 (2), pp. 185-190.

Department of Health (2008). Guidance on the routine collection of Patient Reported Outcome Measures (PROMs), The Stationary Office, London.

Devlin, N.J. y Appleby, J. (2010). Getting the most out of PROMs. *London: Kings Fund*.

Devlin, N.J., Parkin, D. and Browne, J. (2010). Patient-reported outcome measures in the NHS: new methods for analysing and reporting EQ-5D data. *Health economics,* 19 (8), pp. 886-905.

Dimick, J.B. Staiger, D.O., Baser, O. and Birkmeyer, J.D. (2012). Composite Measures for Rating Hospital Quality with Major Surgery. *Health Services Research*, 47 (5), pp.1861-1879.

EuroQol Group (2016). www.euroqol.org. Date of last consultation: 07/08/2016.

García, F., Marcuello, C., Serrano, D. and Urbina, O. (1999). Evaluation of Efficiency in Primary Health Care Centres: An Application of Data Envelopment Analysis. *Financial Accountability and Management*, 15 (1), pp. 67-83.

Garratt, A.M. Macdonald, L.M., Ruta, D.A., Russell, I.T., Buckigham, J.K. and Krukowski, Z. H. (1993). Towards measurement of outcome for patients with varicose veins. *Quality in Health Care,* 2 (1), pp. 5-10.

Godoy-Caballero, A.L. and Murillo-Zamorano, L.R. (2012). Análisis Económico de la Asistencia Sanitaria en Atención Primaria. El caso de la Comunidad Autónoma de Extremadura. *Editorial Académica Española*. I.S.B.N.: 978-3-8473-6526-6.

Gutacker, N., Bojke, C., Daidone, S., Devlin, N. and Street, A. (2013). Hospital Variation in Patient-Reported Outcomes at the Level of EQ-5D Dimensions: Evidence from England. *Medical Decision Making*, 33 (6), pp. 804-818

Hurst, N., Kind, P., Ruta, D., Hunter, M. and Stubbings, A. (1997). Measuring health-related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of EuroQol (EQ-5D). *Rheumatology,* 36 (5), pp. 551-559.

Jacobs, R., Smith, P.C. and Street, A., (2006). Measuring efficiency in health care: analytic techniques and health policy. *Cambridge University Press*, ISBN: 978-0-521-85144-2.

Laudicella, M., Olsen, K. R. and Street, A. (2010). Examining cost variation across hospital departments–a two-stage multi-level approach using patient-level data. *Social Science & Medicine,* 71 (10), pp. 1872-1881.

Marshall, S., Haywood, K. and Fitzpatrick, R. (2006). Impact of patient-reported outcome measures on routine practice: a structured review. *Journal of evaluation in clinical practice,* 12 (5), pp. 559-568.

Murillo-Zamorano, L.R. and Petraglia, C. (2011). Technical Efficiency in Primary Health Care: Does Quality Matter? *The European Journal of Health Economics*, 12 (2), pp. 115-125.

National Institute for Health and Clinical Excellence, (2008). Guide to the methods of technological appraisal. ISBN: 1-84629-741-9.

NHS (2011). http://www.ic.nhs.uk/statistics-and-data-collections/hospital-care/patient-reported-outcome-measures-proms/finalised-patient-reported-outcome-measures-proms-in-england--april-2009-to-march-2010-pre-and-post-operative-data-experimental-statistics. Date of last consultation: 07/08/2016.

Oppe, M., Devlin, N. and Black, N. (2011). Comparison of the underlying constructs of the EQ-5D and Oxford Hip Score: implications for mapping. *Value in Health,* 14 (6), pp. 884-891.

Ostendorf, M., van Stel, H.F., Buskens, E., Schrijvers, A. J.P., Marting, L.N., Verbout, A. J. and Dhert, W.J.A. (2004). Patient-reported outcome in total hip replacement. A comparison of five instruments of health status. *Journal of Bone & Joint Surgery - British Volume,* 86 (6), pp. 801-808.

Ousey, K. and Cook, L. (2011). Understanding patient reported outcome measures (PROMs). *British Journal of Community Nursing,* 16 (2), pp. 80-82.

Rice, N. and Jones, A. (1997). Multilevel models and health economics. *Health Economics,* 6 (6), pp. 561-575.

Salinas-Jiménez, J. and Smith, P.C. (1996). Data envelopment analysis applied to quality in primary health care. *Annals of Operations Research*, 67 (1), pp. 141-161.

Selim, A.J., Berlowitzm D.R., Fincke, G., Rosen, A.K., Ren, X., Christiansen, C.L., Cong, Z., Lee, A. and Kazis, L. (2002). Risk-adjusted mortality rates as a potential outcome indicator for outpatient quality assessments. *Medical care,* 40 (3), pp. 237-245.

Street, A., Kobel, C., Renaud, T. and Thuilliez, J. (2012). How well do diagnosis-related groups explain variations in costs of length of stay among patients and across hospitals? Methods for analysing routine patient data. *Health Economics,* 21 (S2), pp. 6-18.

Thomas, N., Longford, N. T. and Rolph, J. E. (1994). Empirical Bayes methods for estimating hospital-specific mortality rates. *Statistics in medicine,* 13 (9), pp. 889-903.

Valderas, J., Kotzeva, A., Espallargues, M.., Guyatt, G., Ferrans, C.E., Halyard, M.Y., Revicki, D.A., Symonds, T., Parada, A. and Alonso, J. (2008). The impact of measuring patient-reported outcomes in clinical practice: a systematic review of the literature. *Quality of Life Research,* 17 (2), pp. 179-193.

Valderas, J.M. y Alonso, J. (2008). Patient reported outcome measures: a model-based classification system for research and clinical practice. *Quality of Life Research,* 17 (9), pp. 1125-1135.

Wooldridge, J.M. (2009). *Introductory econometrics: A modern approach*. South-Western Pub.

Wylde, V., Blom, A.W., Whitehouse, S.L., Taylor, A.H., Pattison, G.T. and Bannister, G.C. (2009). Patient-reported outcomes after total hip and knee arthroplasty: comparison of midterm results. *Journal of Arthroplasty,* 24 (2), pp. 210-216.

APPENDICES

Appendix 1. Generic PROM EQ-5D

By placing a tick in one box in each group below, please indicate which statements best describe your own health state today

Mobility

I have no problems in walking about	☐
I have some problems in walking about	☐
I am confined to bed	☐

Self-Care

I have no problems with self-care	☐
I have some problems washing or dressing myself	☐
I am unable to wash or dress myself	☐

Usual activities (e.g. work, study, housework, family or leisure activities)

I have no problems with performing my usual activites	☐
I have some problems with performing my usual activites	
I am unable to perform my usual activites	☐

Pain/Disconmfort

I have no pain or discomfort	☐
I have moderate pain or discomfort	☐
I have extreme pain or discomfort	☐

Anxiety/Depression

I am not anxious or depressed	☐
I am moderately anxious or depressed	☐
I am extremely anxious or depressed	☐

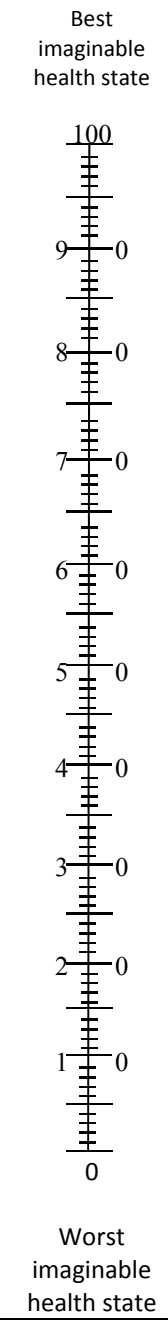
Source: EuroQol Group, 2016

Appendix 2. Generic PROM EQ-VAS

To help people say how good or bad a health state is, we have drawn a scale (rather like a thermometer) on which the best state you can imagine is marked 100 and the worst state you can imagine is marked 0.

We would like you to indicate on this scale how good or bad your own health is today, in you opinion. Please do this by drawing a line from the bow belos to whichever point on the scale indicates how good or bad your health state is today.

Your own health state today



Source: EuroQol Group, 2016

Appendix 3. Specific PROM OHS

Please answer the following 12 multiple choice questions.

During the past 4 weeks…

1. **How would you describe the pain you usually have in your hip?**

☐ None
☐ Very mild
☐ Mild
☐ Moderate
☐ Severe

2. **Have you been troubled by pain from your hip in bed at night?**

☐ No nights
☐ Only 1 or 2 nights
☐ Some nights
☐ Most nights
☐ Every night

3. **Have you had any sudden, sever pain (shooting, stabbing, or spasms) from your affected hip?**

☐ No days
☐ Only 1 or 2 days
☐ Some days
☐ Most days
☐ Every day

4. **Have you been limping when walking because of your hip?**

☐ Rarely/never
☐ Sometimes or just at first
☐ Often, not just at first
☐ Most of the time
☐ All of the time
☐ Not at all

Source: Dawson et al. (1996)

Appendix 3. Specific PROM OHS (Cont.)

5. **For how long have you been able to walk before the pain in you hip becomes severe (with or without a walking aid?**

☐ No pain for 30 minutes or more

☐ 16 to 30 minutes

☐ 5 to 15 minutes

☐ Around the house only

6. **Have you been able to climb a flight of stairs?**

☐ Yes, easily

☐ With little difficulty

☐ With moderate difficulty

☐ With extreme difficulty

☐ No, impossible

7. **Have you been able to put on a pair of socks, stockings or tights?**

☐ Yes, easily

☐ With little difficulty

☐ With moderate difficulty

☐ With extreme difficulty

☐ No, impossible

8. **After a meal (sat at a table), how painful has it been for you to stand up from a chair because of your hip?**

☐ Not at all painful

☐ Slightly painful

☐ Moderately painful

☐ Very painful

☐ Unbearable

Source: Dawson et al. (1996)

Appendix 3. Specific PROM OHS (Cont.)

9. **Have you had any trouble getting in an out of the car or using public transportation because of your hip?**

☐ No trouble at all

☐ Very little trouble

☐ Moderate trouble

☐ Extreme difficulty

☐ Impossible to do

10. **Have you had any trouble with washing and driying yourself (all over) because of your hip?**

☐ No trouble at all

☐ Very little trouble

☐ Moderate trouble

☐ Extreme difficulty

☐Impossible to do

11. **Could you do the household shopping on your own?**

☐ Yes, easily

☐ With little difficulty

☐ With moderate difficulty

☐ With extreme difficulty

☐ No, impossible

12. **How much has pain from you hip interfered with your usual work, inclusing housework?**

☐ Not at all

☐ A little bit

☐ Moderately

☐ Greatly

☐ Totally

Source: Dawson et al. (1996)

Appendix 4. Specific PROM OKS

Please answer the following 12 multiple choice questions.

During the past 4 weeks…

1. **How would you describe the pain you usually have in your knee?**

☐ None

☐ Very mild

☐ Mild

☐ Moderate

☐ Severe

2. **Have you had any trouble washing and driying yourself (all over) because of your knee?**

☐ No trouble at all

☐ Very little trouble

☐ Moderate trouble

☐ Extreme difficulty

☐Impossible to do

3. **Have you had any trouble getting in an out of the car or using public transportation because of your knee? (with or without a stick)**

☐ No trouble at all

☐ Very little trouble

☐ Moderate trouble

☐ Extreme difficulty

☐ Impossible to do

4. **For how long are you able to walk before the pain in you knee becomes severe? (with or without a stick)**

☐ No pain > 60 minutes

☐ 16 - 30 minutes

☐ 5 - 15 minutes

☐ Around the house only

☐ Not at all – severe on walking

Source: Dawson et al. (1998)

Appendix 4. Specific PROM OKS (Cont.)

5. **After a meal (sat at a table), how painful has it been for you to stand up from a chair because of your knee?**

☐ Not at all painful
☐ Slightly painful
☐ Moderately painful
☐ Very painful
☐ Unbearable

6. **Have you been limping when walking because of your knee?**

☐ Rarely/never
☐ Sometimes or just at first
☐ Often, not just at first
☐ Most of the time
☐ All of the time

7. **Could you kneel down and get up again afterwards?**

☐ Yes, easily
☐ With little difficulty
☐ With moderate difficulty
☐ With extreme difficulty
☐ No, impossible

8. **Are you troubled by pain in your knee at night in bed?**

☐ Not at all
☐ Only one or two nights
☐ Some nights
☐ Most nights
☐ Every night

9. **How much has pain from you knee interfered with your usual work? (inclusing housework)**

☐ Not at all
☐ A little bit
☐ Moderately
☐ Greatly
☐ Totally

Source: Dawson et al. (1998)

Appendix 4. Specific PROM OKS (Cont.)

10. **Have you felt that your knee night suddenly ½ give away or ½ let you down?**

☐ Rarely/never

☐ Sometimes or just at first

☐ Often, not at first

☐ Most of the time

☐ All of the time

11. **Could you do the household shopping on your own?**

☐ Yes, easily

☐ With little difficulty

☐ With moderate difficulty

☐ With extreme difficulty

☐ No, impossible

12. **Could you walk down a flight of stairs?**

☐ Yes, easily

☐ With little difficulty

☐ With moderate difficulty

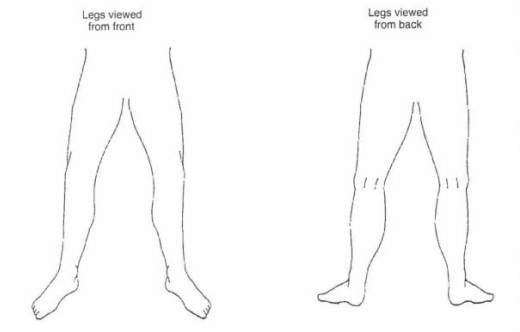
☐ With extreme difficulty

☐ No, impossible

Source: Dawson et al. (1998)

Appendix 5. Specific PROM AVVQ

1. Please, draw in your varicose veins



2. In the last 2 weeks for how many days did your veins cause you pain or ache?

	Right Leg	Left Leg
Not at all	☐	☐
Between 1 and 5 days	☐	☐
Between 5 and 10 days	☐	☐
For more than 10 days	☐	☐

3. During the last 2 weeks, on how many days did you take painkilling tablets for you varicose veins?

	Right Leg	Left Leg
Not at all	☐	☐
Between 1 and 5 days	☐	☐
Between 5 and 10 days	☐	☐
For more than 10 days	☐	☐

4. In the last 2 weeks, how much ankle swelling have you had?

Not at all	☐
Between 1 and 5 days	☐
Between 5 and 10 days	☐
For more than 10 days	☐

Source: Garrat et al. (1993)

Appendix 5. Specific PROM AVVQ (Cont.)

5. In the last 2 weeks, have you worn support stocking or tights?

	Right Leg	Left Leg
No	☐	☐
Yes, those I bouhgt myself without prescription	☐	☐
Yes, those prescribed by my doctor which I wear ocassionaly	☐	☐
Yes, those prescribed by my doctor which I wear every day	☐	☐

6. In the last 2 weeks, have you had any itching in association with your varicose veins?

	Right Leg	Left Leg
No	☐	☐
Yes, above the knee only	☐	☐
Yes, below the knee only	☐	☐
Yes, above and below the knee	☐	☐

7. Do you have purple discolouration caused by tiny blood vessels in the skin, in association with your varicose veins?

	Right Leg	Left Leg
No	☐	☐
Yes	☐	☐

8. Do you have a rash or eczema in the area of you ankle?

	Right Leg	Left Leg
No	☐	☐
Yes, but it does not require treatment from a doctor or district nurse	☐	☐
Yes, and it requires treatment from a doctor or district nurse	☐	☐

9. Do you have a skin ulcer associated with your varicose vein?

	Right Leg	Left Leg
No	☐	☐
Yes	☐	☐

10. Does the appearance of your varicose veins cause you concern?

Not	☐
Yes, their appearance causes my slight concern	☐
Yes, their appearance causes my moderate concern	☐
Yes, their appearance causes my a great deal of concern	☐

Source: Garrat et al. (1993)

Appendix 5. Specific PROM AVVQ (Cont.)

11. Does the appearance of your varicose veins influence your choice of clothing, including tights?

- No ☐
- Occasionally ☐
- Often ☐
- Always ☐

12. During the last 2 weeks , have your varicose veins interfered with your work/housework or other activities?

- No ☐
- I have been able to work byt my work has suffered to a slight extent ☐
- I have been able to work byt my work has suffered to a moderate extent ☐
- My veins have prevented me working one day or more ☐

13. During the last 2 weeks, have your varicose veins interfered with your leisure activities? (including sport, hobbies and social life)

- No ☐
- Yes, my enjoyment has suffered to a slight extent ☐
- Yes, my enjoyment has suffered to a moderate extent ☐
- Yes, my veins have prevented me taking part in any leisure activities ☐

Source: Garrat et al. (1993)