A New Complexity Measure to Classify Ambulatory Patients in Rehabilitation Facilities for Financing Purposes

Elizabeth Reis[1], Abdul Suleman[1], José G. Dias[1], Dália Nogueira[1,2,*] and Claudia M. Borges[3]

[1]Instituto Universitário de Lisboa (ISCTE-IUL), BRU, Lisbon, PORTUGAL
[2] ESSA – Escola Superior de de Saúde de Alcoitão, Lisbon, PORTUGAL
[3] ACSS - DPS, Lisbon, PORTUGAL
[*]Corresponding author: Dália Nogueira, e-mail: dmsnogueira@gmail.com

Abstract

The improvement of living conditions associated with medical advances in the treatment and prevention of many diseases has led to changes characterized by increased longevity and different epidemiological public health profiles. These changes are the result of the evolution in the treatment of diseases that have become less mortal and more chronic. Many of these conditions lead to the inability to perform the Activities of Daily Living (ADL) which, in turn, cause disability and dependence. National Health Systems (NHS) face additional challenges and must respond by reorganizing institutions, resources and financing, in order to provide care in agreement with a new pattern of population's needs. The adjustment of care of outpatients in rehabilitation facilities and the dimension of episodes complexity are not only related to the etiology of the problem, but also with the level of dependency and the patient individual characteristics. The main objective of this research is to create a new indicator based on a case mixed group function, to determine the complexity level of rehabilitation facilities ambulatory patients. This variable will be used as a proxy of care needs to find homogeneous patients groups according to their complexity. Classification variables from a sample of episodes were collected in a retrospective manner. The question that arises is how to tackle such a heterogeneous population. In this study we used a fuzzy clustering approach to data analysis, namely a grade of membership (GoM) representation of data. Having noted a hierarchical structure of fuzzy clusters, individuals can be ranked by complexity, and adjustment of payment can be made accordingly. The achieved results allow determining the complexity of new inpatients by means of a linear regression of the classification variables.

Keywords: classification, fuzzy methods, grade of membership, health services.

1. Introduction

The Portuguese population is living longer than ever before, with life expectancy at age 65 and age 85 representing the largest increase. Longevity is straightly related to the increase of chronic diseases, with stroke and rheumatic conditions leading long-term illnesses, conducting to high levels disability and dependence (Kirsch 2010). To improve care of those with chronic illness, it will be required to develop plans that include the management of multiple types of care across health care settings. Many of these conditions lead to the inability to perform Activities of Daily Living (ADLs) which cause, in turn, disability and dependence. Dependence is the inability to execute, at least, one of the ADLs. Disability is any difficulty with ADLs performance and represents the interaction between individuals' health condition and the environmental barriers.

Chronic health conditions and disability negatively affect individual's quality of life, contributing to the decline in functioning and the inability to take care of themselves. Patients with chronic disease and disability states are amongst the largest consumers of health care and have become a burgeoning concern to health care systems (Passerino 1999).

Rehabilitation involves the intensive management of disorders that intends to better alter the patient's performance in ADLs. Emphasis is always placed on the optimization of function. Any ambulatory rehabilitation service should provide care to people who do not need to be in hospital but require an intensive and coordinated therapy program. These services aim to help patients achieve their highest level of function by identifying specific goals to work on within a time frame. As patients are admitted to rehabilitation programs to improve their level of functional performance this indicator is looked upon as an important predictor of the use of resources in rehabilitation facilities (Carter et al. 2003). Thus, the adjustment of care in an ambulatory regimen must consider the complexity of the episode, meaning that it must not only be related with the etiology of the disease but also with the patient level of dependency.

Health Systems must guarantee the patient's health improvement, promote the evidenced based practice and allocate resources with more equity, by real levels of needs, using better information.

Portuguese outpatient rehabilitation facilities are financed with no differentiation according to the complexity of the patient's disability. It has been long acknowledged that a financing system detached from the patient case-mix and burden of disease may lead to inequities among providers since there are potential incentives to select individuals with less complex disabilities (Gagnon et al. 2005). Since patients are treated not only by the etiological diagnosis but also by functional dependency, a financing system that does not contemplate the amount of health care needed will lead to an inadequate allocation of resources. An evidence based system for allocating financial resources would be fairer and enable the implementation of a prospective financing system according to the patient's clinical characteristics and functional status. Given the assumption that patients with higher dependency require additional time in rehabilitation and more allocation of resources, the Portuguese Ministry of Health has been studying the development of a financing system for ambulatory rehabilitation according to complexity levels.

Both for clinical and management purposes, rehabilitation professionals have been exposed to the need of using objective and valid measures to assess dependency. At a time of increasing need for accountability in the delivery of health care services from both payers and consumers, it has become imperative for health care providers to submit evidence of their results. Nevertheless, although the rehabilitation field has long struggled between the need for comprehensive and clinically sensitive outcome tools and the request for them to be feasible and easy to use in busy clinical settings, there is no overall agreement on what and how it should be measured (Jette et al. 2005).

Rehabilitation professionals in Portugal do not systematically use outcome measure instruments as part of a total outcome evaluation plan to improve clinical or health services purchasers decision making. Therefore, there is little data on the profiles and outcomes of patients undergoing rehabilitation. But the need to develop and implement a suitable process for measuring outcomes in rehabilitation has long been acknowledged. The question is what outcome measure instrument to use. Numerous outcomes are relevant in rehabilitation and are related to body structure and function, activity and participation. The measurement instrument will depend not only on the targets of the rehabilitation program, but also on the amount and type of clinical information needed to plan that program. In addition, when a measurement outcome tool becomes part of a payment formulation according to dependency levels, it also has to consider the aims of the prospective payment system it will serve.

The main objective of this research is to create a new indicator based on a case mixed group function that determines the complexity level of AN ambulatory rehabilitation system. This variable will be used as a proxy of care needs in order to find homogeneous patients groups.
The aim of the present study is to create a patient classification system for ambulatory rehabilitation identifying statistical homogeneous profiles of patients from a set of quantitative and categorical variables that includes sociodemographic profile, morbid conditions and the level of disability measured by the International Classification of Function (ICF).
The universe of rehabilitation facilities in Portugal is numerous as well as the number of patients treated by year. In order to collect data more efficiently and faster, a convenience sample was defined. Two institutions from the social sector, one specialized facility from the public sector, two from the private sector as well as two hospital settings were selected. It was our intention to guarantee that public and private sectors were represented just as they are distributed in Portugal.
This paper is structured as follows: the next section describes the ICF scale to measure dependency and presents the data collection methodology. Section 3 describes the estimation of the complexity measure and section 4 presents the main conclusions.

2. Methodology

2.1. The International Classification of Functioning, Disability and Health (ICF)

The ICF is WHO's framework for measuring health and disability at both individual and population levels. It is a classification of health and health-related domains that are classified from body functions and structures and also societal perspectives. The ICF was endorsed for use by WHO Member States as the international standard to describe and measure health and disability (Stucki et al. 2008). The ICF provides a common framework for describing functional status information in health records to make this information comparable and of value (Kostanjsek et al. 2011, Lohmann et al. 2011, Stucki et al. 2013). The authors stated that the ICF conceptualizes functioning from holistic points of view thereby allowing for planning interventions. They hold that routine collection of functional status information across settings in the health care delivery system can facilitate more effective evaluation of outcomes, comparison of the effectiveness and cost effectiveness of treatment modalities, and prediction and management of costs. The ICF would better enable social policymakers to devise interventions, strategies, and health policies.
The ICF Research Branch developed the ICF Core Sets to facilitate a systematic and comprehensive description of functioning and the use of the ICF in clinical practice and research. ICF Core Sets are lists of ICF categories relevant for specific diseases. They can serve in clinical studies and health statistics or guide multidisciplinary assessments to rate the patient’s level of functioning. An ICF Core Set for a specific condition includes as few categories as possible to be practical, but as many as necessary to be sufficiently comprehensive in describing the typical spectrum of problems in functioning of patients with a specific condition. The list needs to be as short as possible (WHO 2013).

2.2. Study design

Classification variables from a sample of episodes were collected in a retrospective manner for a sample of 1850 episodes. All instruments were applied retrospectively, according to information present in clinical records. Data was gathered in five outpatient private and public rehabilitation facilities in different Portuguese regions; it includes socio demographic information (age and gender) and medical information (disability group, main diagnoses and ICD10 chapter). From the total of 1850 episodes 344 (18.6%) were classified in Core set 1 – neurological, 1404 (75.9%) in Core set 2 – musculoskeletal, and 102 (5.5%) in Core set 3 – cardiopulmonary conditions.

3. Estimation of a complexity measure

The heterogeneity commonly associated with clinical data motivated the use of the grade of membership (GoM) model (Woodbury and Clive, 1974) to describe the uncertainty of different core sets. This model is based on a fuzzy K-partition meaning that each patient, say the patient (indexed by the letter) i, is described by the unit sum vector $g_i = (g_{i1}, g_{i2}, \dots, g_{iK}), i = 1,2, \dots, N$, where the generic coordinate $g_{ik} \geq 0$ is the membership degree of individual i in fuzzy cluster or typology k.
Let $X_i = (X_{i1}, X_{i2}, \dots, X_{iJ})$ be the vector of outcomes of individual i in J items of any core set. Each coordinates of X_i reflects the complexity of the disease associated with the core set in patient i. The level of complexity is measured by a Likert type scale and can assume an integer value from 1 to L_j. For example, for the Core set 1, *Neurologic Diseases*, there are $J = 40$ items and, as with other core sets, $L_j = 5$, for all $j = 1,2, \dots, J$ items. The levels are increasingly ordered, i.e., the value 5 corresponds to the highest complexity.
Given a fixed value of the number of typologies, K, and a random sample of size N, the likelihood function of GoM model can be written as follows (Wachter, 1999):

$$L_K = \prod_{i=1}^{N} \prod_{j=1}^{J} \prod_{l=1}^{L_j} \left(\sum_{k=1}^{K} g_{ik} \times \lambda_{kjl} \right)^{y_{ijl}}$$

where N is the sample size, and it is equal to 344, 1404 or 102, for the Core sets 1, 2 or 3, respectively. The quantity y_{ijl} is an indicator variable which is equal to 1 if individual i has outcome l in variable j and it is equal to 0 otherwise. The parameters λ_{kjl} denote the probability of observing the complexity l of item j in typology k.
The model's goodness-of-fit is assessed through a likelihood ratio test comparing L_K to the model based on a single typology. In practice several models are tested and the optimal value of K is determined by the one that fits the data better than any of the alternatives. We found out that the model based on $K = 5$ typologies best fits the Core set 1 data, while the datasets of Core set 2 and 3 are reasonably modelled by $K = 3$ fuzzy clusters, i.e. by a fuzzy 3-partition.
The question now being asked is how to construct an overall measure of complexity for each patient, based on the output of the GoM model. A scalar measure of individual complexity of the disease would be beneficial for many purposes including to provide appropriate funding. Nascimento (2005) points out the extremal natural of GoM model potentially leading it to identify a hierarchical fuzzy K-partition of the observations. Given such a possibility, we propose a measure of individual complexity as follows. Suppose we know the complexity associated with the typology k, and denote it by S_k. Then, the complexity associated with the patient i, $S(i)$, can be calculated as a convex combination of all S_k,

$$S(i) = \sum_{k=1}^{K} g_{ik} \times S_k$$

As noted earlier, the quantities λ_{kjl} accounts, in probabilistic terms, for how likely is a certain level of complexity l in the typology k. So, we can get a notion of the average complexity in typology k, and use this average as an estimate of S_k, by the formula

$$S_k = \sum_{j=1}^{J} \sum_{l=1}^{L_j} \lambda_{kjl} \times l$$

Higher values of S_k are associated with higher complexity and, in consequence, critically ill patients should present greater values of $S(i)$. So we can use $S(i)$ as a measure or proxy of hospital cost. Tables 1, 2, and 3 present the estimated average of complexity in each typology in three different core sets. They also display the estimates of different dimensions of analysis which aggregate several items. The typologies were ordered *a posteriori* by complexity. We can see at a glance that the body functions or activity limitations are correlated with the complexity of Neurologic diseases, but not necessarily with musculoskeletal or cardio respiratory disorders.

Table 1 – Estimates for Core Set 1

Typology (k)	Body functions	Activity limitations	Complexity (S_k)
1	0.20	0.21	0.21
2	0.83	1.21	0.99
3	1.11	1.54	1.29
4	1.17	2.57	1.76
5	2.43	3.73	2.98

Table 2 – Estimates for Core Set 2

Typology (k)	Body functions	Activity limitations	Complexity (S_k)
1	0.21	0.08	0.13
2	0.98	0.86	0.90
3	2.01	1.94	1.97

Table 3 – Estimates for Core Set 3

Typology (k)	Body functions	Activity limitations	Complexity (S_k)
1	1.36	1.01	1.22
2	2.54	1.52	2.09
3	2.34	2.85	2.54

At the next stage of our data analysis we are concerned with the prediction of the complexity of a new patient. For this purpose, we used the individual complexity $S(i)$ as a dependent variable and estimated a linear regression model on the basis of observed items. Prior to apply such a model, we assumed that the J items in each core set are vectors of $\mathbb{R}^J$. Although this is convenient for practical purposes, we have no formal statistical justification for such an assumption. In order to avoid the regression models to be affected by multicollinearity, we performed a principal component analysis to identify orthogonal dimensions. The general form of a regression model can be written as

$$S(i) = B_0 + \sum_{j=1}^{J'} B_j X_{ij} + \mathcal{E}_i \,, \qquad i = 1,2,..,N$$

where B_0 is the constant term, B_1, B_2, …, $B_{J'}$ are coefficients of the effective J' predictors or independent variables and $\mathcal{E}_i$ are the error terms and, by assumption, are normally and independently distributed with zero mean and σ^2 variance. All estimated models are good fit to data as measured by the adjusted coefficient of determination (R^2_{adj}).

In sum, given the assessments of a new episode, we can easily calculate an estimate of the complexity associated with it, and subsequently the respective cost. The fuzzy approach to data analysis can be thought as a mean to obtain the weights of each indicator in determining an overall individual measure of complexity of patients for every core set.

4. Conclusions

This paper introduces a new indicator for the complexity of rehabilitation ambulatory patients. Classification variables from a sample of episodes were collected in a retrospective manner. Data heterogeneity was extracted by the grade of membership (GoM) model, a fuzzy clustering approach. Then, patients can be ranked by complexity, and adjustment of payment can be made accordingly. The achieved results allow determining the complexity of new ambulatory patients by means of a linear regression of the classification variables.
Further research can explore how individual severity index can be associated with profiling characteristics of the patients, allowing an easier classification of patients into homogeneous groups. It may explore the application classification techniques (e.g., discriminant analysis, logistic regression, and CART – Classification and Regression Trees) with this allocation purpose.

References

Carter, M.G., Relles, D.A., Ridgeway, G.K. and Rimes, C.M. (2003) "Measuring Function for Medicare Inpatient Rehabilitation Payment," *HealthCare Financing Review*, 24(3), 25-44.

Gagnon, D., Nadeau, S. and Tam, V. (2005) "Clinical and administrative outcomes during publicly funded impatient rehabilitation based on a case-mix group classification model," *Journal of Rehabilitation Medicine*, 37, 45-52.

Jette, A. M. and Haley, S. M. (2005) "Contemporary measurement techniques for rehabilitation outcomes assessment," *J Rehabil Med*, 37(6): 339-345.

Kirsch, B. (2010) "Stroke rehabilitation requires knowing what level a patient needs," *Managed Care*, 19(11): 32-34.

Kostanjsek, N., Rubinelli, S. , Escorpizo, R., Cieza, A., Kennedy, C., Selb, M., Stucki, G. and Ustun, T. B. (2011) "Assessing the impact of health conditions using the ICF," *Disabil Rehabil*, 33(15-16): 1475-1482.

Lohmann, S., Decker, J., Müller, M., Strobl, R. And Grill, E. (2011) "The ICF forms a useful framework for classifying individual patient goals in post-acute rehabilitation," *J Rehabil Med*, 43, 151-155.

Nascimento, S. (2005) *Fuzzy Clustering via Proportional Membership Model*, IOS Press.

Passerino, C. (1999) "Costs of rehabilitation in the health systems," *G Ital Med Lav Ergon*, 21(1): 58-67.

Stucki, G., Kostanjsek, N., Ustun and, B. and Cieza, A. (2008) "ICF-based classification and measurement of functioning," *Eur J Phys Rehabil Med*, 44(3): 315-328.

Stucki, G. and Cieza, A. (2013) "The International Classification of Functioning, Disability and Health (ICF) Core Sets for rheumatoid arthritis: a way to specify functioning," *Ann Rheun Dis*, 63(Suppl II): ii40-ii45.

Wachter, K.W. (1999) "Grade of Membership Models in Low Dimensions," *Statistical Papers*, 40, 439-457.

WHO (2013) *ICF Core Sets*, World Health Organization.

Woodbury, M.A. and Clive, J. (1974) "Clinical Pure Types as a Fuzzy Partition," *Journal of Cybernetics*, 11, 277-298.