Artificial Intelligence in Medicine 58 (2013) 165–173

Contents lists available at SciVerse ScienceDirect

Artificial Intelligence in Medicine

journal homepage: www.elsevier.com/locate/aiim

Computational intelligence for the Balanced Scorecard: Studying performance trends of hemodialysis clinics

Isabella Cattinelli[a,*], Elena Bolzoni[a], Milena Chermisi[a], Francesco Bellocchio[a], Carlo Barbieri[a], Flavio Mari[a], Claudia Amato[b], Marcus Menzer[b], Andrea Stopper[b], Emanuele Gatti[b,c]

[a] *Healthcare and Business Advanced Modeling, Fresenius Medical Care, Else-Kröner-Strasse 1, 61352 Bad Homburg, Germany*
[b] *Fresenius Medical Care, Else-Kröner-Strasse 1, 61352 Bad Homburg, Germany*
[c] *Centre for Biomedical Technology at the Danube, University of Krems, Dr.-Karl-Dorrek-Strasse 30, 3500 Krems, Austria*

ARTICLE INFO

Article history:
Received 10 September 2012
Received in revised form 12 April 2013
Accepted 18 April 2013

Keywords:
Self-organizing maps
Clustering
Markov chains
Balanced Scorecard
Dialysis
Health care performance

ABSTRACT

Objectives: The Balanced Scorecard (BSC) is a general, widely employed instrument for enterprise performance monitoring based on the periodic assessment of strategic Key Performance Indicators that are scored against preset targets. The BSC is currently employed as an effective management support tool within Fresenius Medical Care (FME) and is routinely analyzed via standard statistical methods. More recently, the application of computational intelligence techniques (namely, self-organizing maps) to BSC data has been proposed as a way to enhance the quantity and quality of information that can be extracted from it. In this work, additional methods are presented to analyze the evolution of clinic performance over time.
Methods: Performance evolution is studied at the single-clinic level by computing two complementary indexes that measure the proportion of time spent within performance clusters and improving/worsening trends. Self-organizing maps are used in conjunction with these indexes to identify the specific drivers of the observed performance. The performance evolution for groups of clinics is modeled under a probabilistic framework by resorting to Markov chain properties. These allow a study of the probability of transitioning between performance clusters as time progresses for the identification of the performance level that is expected to become dominant over time.
Results: We show the potential of the proposed methods through illustrative results derived from the analysis of BSC data of 109 FME clinics in three countries. We were able to identify the performance drivers for specific groups of clinics and to distinguish between countries whose performances are likely to improve from those where a decline in performance might be expected. According to the stationary distribution of the Markov chain, the expected trend is best in Turkey (where the highest performance cluster has the highest probability, $P = 0.46$), followed by Portugal (where the second best performance cluster dominates, with $P = 0.50$), and finally Italy (where the second best performance cluster has $P = 0.34$).
Conclusion: These results highlight the ability of the proposed methods to extract insights about performance trends that cannot be easily extrapolated using standard analyses and that are valuable in directing management strategies within a continuous quality improvement policy.

1. Introduction

Health care units are complex entities whose good functioning can depend on many diverse factors. The technical preparation of the medical staff must be paired with adequate structures and equipment, which in turn can be supplied only if sufficient economical resources are available. The combination of these and other aspects should concur to create an optimal functioning condition that is the prerequisite for the clinic to achieve its ultimate aim: ensuring the health and welfare of its patients. It follows that performance monitoring, an important activity in any type of enterprise, can be even more crucial in health care companies. A robust system for performance assessment that can take into account many factors of a different nature (e.g., financial, medical) can provide the starting point for devising appropriate management strategies that ensure that the highest operation standards are achieved and maintained over time.

This can be particularly important for those health care units devoted to the treatment of chronic diseases, where patients visit the clinic on a regular basis for several years and therefore spend

* Corresponding author: Tel.: +39 0373974380; fax: +39 03739747380.
E-mail address: isabella.cattinelli@fmc-ag.com (I. Cattinelli).

0933-3657/$ – see front matter
http://dx.doi.org/10.1016/j.artmed.2013.04.005

a considerable amount of their lives within the clinic itself. This is the typical scenario in the management of end stage renal disease, a condition that requires frequent treatments in hemodialysis clinics (although it should be noted that in recent years peritoneal dialysis, which is performed autonomously by patients at home, has become increasingly widespread as an alternative to hemodialysis). A major provider of therapies and services related to dialysis is Fresenius Medical Care (FME), a health care company comprising more than 600 centers in 19 countries, considering only the European district.

FME has chosen, since 2007, to adopt the Balanced Scorecard (BSC) [1] as the main instrument to formulate a global corporate strategy and to support the company's performance administration, including clinic management, within its European clinics [2]. Introduced in 1992, the BSC is designed to provide a view of the overall efficiency of a company by assessing the extent to which specific sub-goals are reached. To this aim, the company operations are broken down into a number of main areas of business, usually named *perspectives*, which in turn include a set of carefully selected parameters of interest. These parameters are called *Key Performance Indicators* (KPIs) and represent the management sub-goals as defined by the BSC implementers based on the unique characteristics of the company. Each KPI is periodically evaluated to check its adherence to the predefined goal for that KPI: if the goal is reached, the KPI is assigned a high score. By combining the scores for each defined KPI across all perspectives, one can gain a clear picture of the overall performance of the company.

The BSC is a rather general conceptual framework that can be, and in fact has been, tailored to application in different types of enterprise [3,4]. Indeed the BSC is one of the most popular methodologies for business performance measurement [3] and naturally accommodates within its framework concepts from disparate areas of clinic management, such as the patient and employee perspectives, allowing, in principle, the study of potential efficiency trade-offs emerging between different perspectives. For these reasons, FME decided to build its own implementation of the BSC, the NephroCare (NC) BSC, which is organized according to four perspectives (patients, employees, shareholders, and community) and is largely based on the clinical, operational and financial data stored in the company warehouse [5,6]. In the NC BSC, each clinic is evaluated on a monthly basis by computing for each of the defined KPIs a score that reflects how close the clinic is to excellence with respect to that parameter.

The NC BSC is one of the primary inputs to FME management decision making: monthly reports at different abstraction levels (e.g. clinic groups or single clinics) note those situations that may need to be corrected and those that are entitled to receive incentives. However, these reports are limited in that they do not allow for an immediate grasp of the relations existing between different KPIs, whether for a given clinic or for all of them together, and therefore lack the ability to identify clinics that share a similar behavior in terms of KPI correlations. This motivated the introduction of computational intelligence techniques to complement the standard analysis already performed on the BSC data. Self-organizing maps (SOMs) [7] have recently been introduced as a tool for extracting relevant correlation patterns on groups of KPIs and for visualizing them in an intuitive manner [8].

The power of SOMs to automatically discover unanticipated relations among KPIs and to convey complex information in a compact visual representation makes them a valuable enhancement of the decision making process. Other advanced computational techniques have been proposed in the literature to be applied within the BSC framework. For instance, in [9], fuzzy weights are assigned to a set of banking performance indexes and are used to rank perspectives and indexes to identify the most important BSC factors for the bank's growth strategy; in [10], a similar approach is applied to the performance assessment of the IT departments in industry. Notice, however, that whereas these works aim to define weights for performance indicators expressing their relative importance *a priori*, the approach in [8] looks for *emergent* relations among them, highlighting potential links between groups of KPIs that were not apparent at the time of BSC design, but are disclosed by the actual KPI scores. In a similar spirit, [11] uses clustering and partial least square path modeling to infer cause-effect relations between performance areas of a BSC for the health care sector; however, relations are established with respect to all considered clinics, rather than differentiated for groups of clinics exhibiting distinct characteristics. The main shortcoming of all these approaches lies in that they fail to fully exploit the temporal information contained in the BSC data. In our application context, because each clinic is assessed once per month, KPI scores naturally lend themselves to be studied as they change in time to detect improving and worsening trends.

In this paper, we present a method for analyzing the temporal dimension of the BSC data that builds on top of the SOM-based analysis previously introduced. We analyze performance evolution both at a single clinic level and on groups of clinics: the temporal evolution of single clinics is quantified by means of a pair of indexes, whereas the long-term trend for a set of clinics (for example, all clinics within a given country) is described in probabilistic terms by resorting to Markov chain properties. In describing the proposed methodology, we will also present relevant results obtained from real BSC data and discuss the importance of such temporal analysis in informing management strategies in a timely manner.

2. Material and methods

2.1. The NC BSC data

The BSC from which our data were extracted [2] defines 30 KPIs organized in four perspectives (see Table 1). For any clinic, relevant raw data are collected every month and assembled to produce a value for each KPI. KPI values, in turn, are compared to their corresponding target to produce KPI scores. KPI scores range from 0 to 100, with 100 representing the full accomplishment of the target for that KPI. It follows that a single clinic in a given month can be effectively described by the vector of its KPI scores: $\boldsymbol{x}_i = (v_{i1}, \ldots, v_{iN}) \in \mathbb{R}^N$ for $i = 1, \ldots, m$ (where i indexes every clinic-month, and N is the number of considered KPIs – in the most general case, $N = 30$). Each $\boldsymbol{x}_i$ is called a KPI record. Additional details on the NC BSC can be found in the Supplemental Materials.

Within the European sector of FME, the NC BSC is adopted in 18 countries, each hosting a variable number of clinics. In this work we present our results on a subset of this massive amount of data, namely, the data from Italy, Portugal, and Turkey for the period January 2008 – April 2010 (see Table 2).

2.2. Self-organizing maps

Self-organizing maps [7] are a class of artificial neural networks widely applied to clustering and visualization tasks. Neurons in a SOM, commonly organized in a 2-dimensional grid, act as prototypes for input vectors, effectively remapping the high-dimensional input dataset to a 2D representation. The SOM is trained in an unsupervised fashion by iteratively assigning each input vector to its closest neuron and then updating the map so that this unit and its neighbors are made even closer to the input pattern. This training procedure ultimately builds a map where the neighborhood relationships in the data are mostly preserved, that is, similar input vectors are placed in close positions in the resulting map.

We trained a different SOM for each country using its KPI records as training data. Prior to training, the KPI records were preprocessed to remove KPIs with missing values or having a low standard