Benchmarking in healthcare:  
management tool or mechanism of governance?

Thesis presented by

Silvia Zett

to

The Class of Social Sciences

for the degree of Doctor of Philosophy

in the subject of

International PhD in Management: Innovation, Sustainability and Healthcare

Tutor: Prof. Sabina Nuti

Advisor: Dott. Milena Vainieri

Scuola Superiore Sant’Anna

Y.A. 2013-2014
# TABLE OF CONTENTS

CHAPTER 1 – Introduction .......................................................................................................................... 6

1. Background ............................................................................................................................................... 6
2. Research questions ............................................................................................................................... 9
3. Final outputs of the PhD research ...................................................................................................... 14
References .................................................................................................................................................. 17

CHAPTER 2 – Which are positive and negative effects of voluntary and compulsory benchmarking models? A review of literature in the healthcare sector .......................................................... 20

1. Introduction ............................................................................................................................................... 21
2. Aim of the research ............................................................................................................................... 22
3. Compulsory and voluntary benchmarking .......................................................................................... 22
4. Methodology ........................................................................................................................................... 25
5. Results ..................................................................................................................................................... 27
6. Discussing and concluding comments ............................................................................................... 37
References .................................................................................................................................................. 39

CHAPTER 3 – Assessment and improvement of the Italian Healthcare system: first evidences from the national performance evaluation system .............................................................................. 42

1. Introduction ............................................................................................................................................... 44
2. The Italian Performance Evaluation system: the conceptual framework ................................................. 46
3. The PES reporting system ..................................................................................................................... 53
4. Policy implications ............................................................................................................................... 56
5. Conclusions ............................................................................................................................................. 58
References .................................................................................................................................................. 61
CHAPTER 4 - The performance evaluation system of prevention: the experience of the Experimental Institutes for Animal Disease Prevention (II.ZZ_SS)……………………………..67

1. Introduzione: il settore della prevenzione nel contesto Italiano…………………………...68
2. La storia degli II.ZZ_SS…………………………………………………………………………………72
3. Il modello “toscano” verso gli II.ZZ_SS……………………………………………………………..76
   - I soggetti coinvolti……………………………………………………………………………………77
   - La metodologia………………………………………………………………………………………77
4. I risultati ……………………………………………………………………………………………………80
5. Conclusioni………………………………………………………………………………………………89

References ………………………………………………………………………………………………………92

Final Conclusions ……………………………………………………………………………………………95

Acknowledgements ………………………………………………………………………………………….98
CHAPTER 1 – Introduction

1. Background
The publication of new management on “how to do benchmarking” appears to have peaked by the late 1990s (Watson 1993; Zairi 1998; and for the non-profit sector, Letts et al. 1999). In the literature, various classifications of benchmarking are used. The following section provides an overview of selected benchmarking categories.

Many writers on benchmarking attempt to produce a classification framework that adequately describes some definitions of benchmarking.

The most commonly cited typology of benchmarking is Camp’s (1995) distinction between internal, competitive, functional and generic benchmarking. These four categories have subsequently been expanded upon by other writers: for example, ‘best in class benchmarking’ is used by Spendolini (1992) to emphasize the organization-independent nature of generic benchmarking; McGonagle and Fleming (1993) identified also shadow benchmarking that defined against a competitor, industrial benchmarking that compares similar organizations, but not exactly the same functions within the same industry/sector, often against the industry leaders (it is similar to functional benchmarking) and transnational benchmarking that analyzes processes, comparison with any industry and with world leading organizations (similar to world class or generic benchmarking). Also Shetty (1993) distinguish operational benchmarking that focusing on costs and product differentiation, management benchmarking related to supporting planning functions and strategic benchmarking analysis of successful business strategies.

Still Bhutta and Huq (1999) proposed the following matrix (see Table 1), which includes the key characteristics and relevance (high, medium, and low) of three types of benchmarking: performance, process and strategic.
Others, Coopers & Lybrand (1994) and Mayle et al. (2001), have reported that organizations often start by benchmarking internally, probably at least in part because of the complexities of establishing partnerships, particularly with competitors. An alternative to grappling with some of the more problematic aspects of competitive benchmarking is to adopt generic benchmarking with unlike partners; indeed, Camp (1995) and others point out that truly innovative ideas are more likely to be found by looking at key processes outside one’s own industry (Francis et al., 2007).

The Next Table shows the main typologies which have been developed to describe or classify the characteristics of benchmarking.

<table>
<thead>
<tr>
<th>Study</th>
<th>Typology</th>
<th>Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Camp (1995)</td>
<td>Internal, competitive, functional and</td>
<td>Internal: comparison among similar operations within one’s own organization;</td>
</tr>
<tr>
<td></td>
<td>generic benchmarking</td>
<td>Competitive: comparison to the best of the direct competitors; Functional:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>comparison of methods to companies with similar processes in the same function</td>
</tr>
<tr>
<td></td>
<td></td>
<td>outside one’s industry; Generic process: comparison of work processes to</td>
</tr>
<tr>
<td></td>
<td></td>
<td>others who have innovative, exemplar work processes. (Adapted from Camp</td>
</tr>
<tr>
<td></td>
<td></td>
<td>examples of generic benchmarking.</td>
</tr>
<tr>
<td>McGonagle and Fleming</td>
<td>Internal, competitive, functional,</td>
<td>Shadow benchmarking defined against a competitor; industrial benchmarking</td>
</tr>
<tr>
<td>(1993)</td>
<td>generic, shadow, industrial,</td>
<td>compared similar organizations, but not exactly the same functions within the</td>
</tr>
<tr>
<td></td>
<td>transnational benchmarking</td>
<td>same industry/sector; transnational benchmarking analyzes processes.</td>
</tr>
<tr>
<td>Trosa and Williams (1996)</td>
<td>Results benchmarking and process</td>
<td>‘Results benchmarking’ is concerned (merely) with the comparative data</td>
</tr>
<tr>
<td></td>
<td>benchmarking</td>
<td>generated by benchmarking. ‘Process benchmarking’ considers how results</td>
</tr>
<tr>
<td></td>
<td></td>
<td>were achieved, so that the performance gaps identified in the ‘results’ can</td>
</tr>
<tr>
<td></td>
<td></td>
<td>be closed by investigating and learning from others’ practices.</td>
</tr>
<tr>
<td>CIPFA (1996)</td>
<td>International or global benchmarking</td>
<td>The comparison with exemplary performance by organizations overseas; can be</td>
</tr>
<tr>
<td></td>
<td></td>
<td>competitive, functional or generic.</td>
</tr>
<tr>
<td>El Nathan et al. (1996)</td>
<td>Unilateral and co-operative</td>
<td>Distinguish in particular between ‘unilateral’ (often covert and independent)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>and ‘co-operative’ approaches (where information is voluntarily shared</td>
</tr>
<tr>
<td></td>
<td></td>
<td>between participants). The latter includes ‘group’, ‘indirect/third party’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>and ‘database’ forms.</td>
</tr>
<tr>
<td>Murdoch (1997)</td>
<td>Lateral benchmarking</td>
<td>Another term used to describe generic benchmarking.</td>
</tr>
<tr>
<td></td>
<td>benchmarking</td>
<td>of information-gathering exercises such as a survey undertaken by a national</td>
</tr>
<tr>
<td></td>
<td></td>
<td>agency) and ‘explicit benchmarking’ (a deliberate and structured process</td>
</tr>
<tr>
<td></td>
<td></td>
<td>to facilitate comparison and identify directions for change that will lead</td>
</tr>
<tr>
<td></td>
<td></td>
<td>to improvement.</td>
</tr>
<tr>
<td>Bhutta and Huq (1999)</td>
<td>Performance, process and strategic</td>
<td>The study includes the key characteristics and relevance (high, medium, and</td>
</tr>
<tr>
<td></td>
<td>benchmarking</td>
<td>low) of three types of benchmarking</td>
</tr>
<tr>
<td>Jackson and Lund (2000)</td>
<td>Implicit benchmarking and explicit</td>
<td>Jackson and Lund (2000) also use this typology in their discussion of the</td>
</tr>
<tr>
<td></td>
<td>benchmarking</td>
<td>emergence of benchmarking in higher education. Focusing on the</td>
</tr>
<tr>
<td></td>
<td></td>
<td>information-gathering roles of participants.</td>
</tr>
<tr>
<td>Bowerman et al. (2002)</td>
<td>Voluntary or compulsory</td>
<td>Consider the differences between public and private sector benchmarking and</td>
</tr>
<tr>
<td></td>
<td></td>
<td>classifies public sector benchmarking in terms of whether it was voluntary</td>
</tr>
<tr>
<td></td>
<td></td>
<td>or compulsory.</td>
</tr>
</tbody>
</table>


Table 2: Some definitions of benchmarking

According to recent studies in international management, other classifications about benchmarking process have been done considering the degree of relations within the structures. Dossi and Patelli (2008, 2010) conducted an interesting research basis on the relationship between headquarters and subsidiaries, and the managerial mechanisms employed in order to effectively govern these relationships.
With reference to overview, my research focused on the study of Bowerman (2002) that classified public sector in 2 different types of benchmarking, voluntary and compulsory.

2. Research questions
The first research question analyzed is RQ1 “Which are positive and negative effects of voluntary and mandatory benchmarking models in the healthcare sector?”.

Benchmarking in the healthcare sector has undergone several modifications (Ellis, 2006). Initially, benchmarking was essentially the comparison of performance outcomes to identify disparities. Then its utilization in the mid-1990s, as a structured method, emerged in the United States and the United Kingdom with the imperative of comparing hospital outcomes to rationalize their funding (Camp, 1998).

Dewan et al. (2000) presented the application in a model collaborative outcome management project in the United States to improving the quality of behavioral health care through data and best practices sharing; other authors (Ettorchi-Tardy A. et al., 2012) reported the significance of benchmarking as competition or compulsory in the healthcare, seeking to predict the nature of the relationships between benchmarking parties.

Moreover seems that compulsory benchmarking system utilized when the holding decides strategies about subsidiaries, otherwise when exits the network organization a right model could be voluntary system. In the first case Dossi and Patelli (2010) sustained that traditional hierarchical control mechanisms are not appropriate to govern interdependent relationships between headquarters and subsidiaries within modern international organizations.

Nevertheless the study of my work has been to demonstrate that compulsory type of benchmarking is also required as Italian regions highlighted. The Tuscany Region has sought to introduce, with the performance evaluation system (PES), an instrument of government of the regional health system where the strategic long-term were control together the short-term
objectives, and where government spending was integrated with measures result, in order to highlight the value produced for the citizen (Nuti, 2008).

Instead, if the system is not structured but the method of benchmarking has shared within the network organization, the voluntary model will be immediately introduced in the current environment.

In line with the context before explicated I reported second and third research questions that I have been analyzed:

RQ2 “Which tools to compare performance at national level through compulsory benchmarking?” and RQ3 “How to build a voluntary Performance Evaluation System for II.ZZ.SS.”

In order to answer to the three research questions three papers are presented:

1. Which are positive and negative effects of voluntary and compulsory benchmarking models? A review of literature in the healthcare sector
2. Assessment and improvement of the Italian Healthcare system: first evidences from a pilot national performance evaluation system
3. The performance evaluation system in the prevention sector: the results of the Experimental Institutes for Animal Disease Prevention (II.ZZ.SS.)

They respectively answer to the first, the second and the third research question. For each paper the respective abstract is reported.

*RQ1 “Which are positive and negative effects of voluntary and mandatory benchmarking in the healthcare sector?”*

The paper 1 “Which are positive and negative effects of voluntary and compulsory benchmarking models? A review of literature in the healthcare sector” aims to understand the
differences of two approaches and to consider when to use that or the other model. Benchmarking is recognized as an essential tool for continuous improvement of quality. A large number of publications by various authors reflect the interest in this technique (Dattakumar, 2003).

The purpose of article is to analyse the recent evolution of benchmarking both compulsory and voluntary models (see Bowerman et al., 2002), and to understand which form could be used in the healthcare sector.

The method followed is a systematic search of the English-language literature. In addition I identified articles that have cited the articles found in the search and backward snowballing (from the reference list). This paper provides a review of the 15 important articles published until December 2012 using “benchmarking & healthcare” as key words in the EBSCO database.

In order to compare the two models, I look for how many articles were voluntary and compulsory in healthcare, and where and when the studies were done.

I found that several authors attempted to classify benchmarking process in the healthcare sector as either voluntary or compulsory, however they conclude that the compulsory approach is said to be flawed (Bowerman et al., 2002).

Some authors said that the voluntary model probably has the highest potential for generating benefits to the healthcare sector if the context is composed by organizations that sharing idea, processes and strategies; others support that the compulsory type of benchmarking can be more useful to help policy makers to plan the reform of local government.
RQ2 “Which tools to compare performance at national level through compulsory benchmarking?”

RQ2 was analyzed throughout the paper “Assessment and improvement of the Italian Healthcare system: first evidences from a pilot national performance evaluation system”. It is based on an empirical study carried out in the Tuscan Health System in Italy.

The Italian National Health System (NHS), established in 1978, follows a model similar to the Beveridge model developed by the British NHS (Beveridge 1942; Musgrove 2000). Like the British NHS, healthcare coverage for the Italian population is provided and financed by the government through taxes. Universal coverage provides uniform healthcare access to citizens and is the characteristic usually considered the added value of a welfare system financed by tax revenues.

Nevertheless, in Italy the strong policy of decentralization, which has been taking place since the early 1990s, has gradually shifted powers from the state to the 21 Italian regions. Consequently, the state now retains limited supervisory control and continues to have overall responsibility for the NHS in order to ensure uniform and essential levels of health services across the country. In this context, it has become essential, both for the ministry and for regions, to adopt a common performance evaluation system (PES).

The subsequent Regional Outcome Evaluation Program, called P.Re.Val.E., was conducted in the Lazio region of Italy in order to compare the evaluation of regional healthcare outcomes. After National Outcome Program has launched at national level by National Agency of Regional Health Services (AGENAS) relate to link outcomes in hospitals and local health units in Italy (Fusco et al., 2012).

This article reports the definition, implementation, and first evidences of a pilot PES at a national level. It shows how this PES can be viewed as a strategic tool supporting the Ministry
of Health (MoH) in ensuring uniform levels of care for the population and assisting regional managers to evaluate performance in benchmarking.

Finally, lessons for other health systems, based on the Italian experience, are provided.

**RQ3 “How to build a voluntary Performance Evaluation System for II.ZZ.SS.”**

This research question has been analyzed throughout the paper “The performance evaluation system in the prevention sector: the results of the Experimental Institutes for Animal Disease Prevention (II.ZZ.SS.)”

The study conducted by the Laboratory of Management and Health, with the collaboration of the Ministry of Health in 2011, has involved 9 out of 10 institutes present locally in a well-structured and shared benchmarking process, using the same framework of the Performance Evaluation System of Tuscan Health Authorities.

The choice to continue on this path, which began with the testing of the two II.ZZ.SS. Lazio-Toscana and Umbria-Marche in 2009, was born from the needs of the institutes to use a performance evaluation system, based on objective data to support decision-making processes.

Thanks to the work done by the practitioners belonging to different occupational groups through various meetings useful for correct data retrieval, we have identified 80 indicators, of which 21 concern performance evaluation and 59 related to observational.

The results are represented thorough the “target” diagram, instrument well-known since 2004 by the Region of Tuscany, has allowed to clearly identify the strengths and weaknesses of the Institutes.

The first results obtained and the critical situations emerged have determined a continuous evolution towards effective responses to the needs of human and animal health. It outlined a scenario of important perspectives on which will not be less than the commitment to create
conditions useful to measure the appropriateness and quality of services provided by II.ZZ.SS. benefit of the population (Nuti et al., 2012).

3. Final outputs of the PhD research

In the following table details of the three papers are reported in a snapshot title, date/cases, methodology, difficulties, results and contributions, in particular:

- One or two research questions about my work
- Some papers/books which best represent the background/theory of my work
- Which data/cases I used and how and when I collected them
- Which methodology I used
- Which difficulties I think I had to face in my work
- How my work updates something which has been done in my group or represents a new piece of research
- What is original/nice/interesting about my research
- Some works which have done similar things outside my research group.

<table>
<thead>
<tr>
<th>Paper Title</th>
<th>Research Question</th>
<th>Literature Background</th>
<th>Data/cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>3. The performance evaluation system in the prevention sector: the results of the Experimental Institutes for Animal Disease Prevention (II.ZZ.SS.)</td>
<td>9 “visits” to everyone Institute. Group: 5-6 people (IZS Umbria-Marche; IZS Lazio-Toscana; 2 researches of MeS; more 2 II.ZZ.SS.; one representative of Ministry of Health). What were we doing? To verify the informatization level about records and data mining. Moreover: Development of a glossary of terms used (continuously updated) Using a web shared folder where to find the updated material Web platform where load data for their reporting</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
1. Author: **Silvia Zett** – working paper

2. Authors: Sabina Nuti, Chiara Seghieri, Milena Vainieri, **Silvia Zett**. This paper is published on the Journal of Healthcare Management

Finally, the additional deliverables were carried out during the period of PhD: Zett S., 2011, *Pay for performance in healthcare: a review of literature*, presented at the Italian Health Economics Association (AIES), 16th Annual Conference hosted by Second University of Naples, “National rights, regional differences and federalism. The Italian Health Care System 150 Years after the Unification”, Naples, 29-30 September 2011, and other contributions included in the annual reports of the performance evaluation systems carried out by Laboratorio Management e Sanità.
References


Nuti S., 2008, La valutazione per la governance del PSR, Salute e Territorio, Vol. 166, pp. 53-58


CHAPTER 2 – Which are positive and negative effects of voluntary and compulsory benchmarking models? A review of literature in the healthcare sector

Background: Benchmarking is one of the private sector-grown managerialist tools whose application and significance is rapidly increasing in the public sector, in particular health care (Wolfram Cox et al., 1997; Davis, 1998; Ball et al., 2000). The comparative can be analyzed through the nature (competitive or comparative), the process (based on indicators or ideas) and the outcomes (standards or “best practice”) in public healthcare.

Purpose: The aim of article is to analyse the recent evolution of benchmarking both compulsory and voluntary models (see Bowerman, 2002), and to understand which form could be used in the healthcare sector.

Methodology: The method followed is a systematic search of the English-language literature. In addiction from the articles identified initially, I expanded my search using a snowball sampling method. This paper provides a review of the 15 significant articles published until December 2012. In order to compare the two models, I look for how many articles were voluntary and compulsory in healthcare, and where and when the studies were done.

Results: I found that several authors attempt at classifying benchmarking in the healthcare sector as either voluntary or compulsory, however they conclude that the compulsory approach is said to be flawed (Bowerman et al., 2002).

Conclusions: Although it is agreed that the voluntary, improvement-oriented type of benchmarking probably has the highest potential for generating benefits to the healthcare sector, some authors who support that the compulsory type of benchmarking can be useful to help policy makers to plan the reform of local government. An appropriate balance between the use of benchmarking for control and improvement purposes is yet to be achieved.

Keywords: benchmarking, healthcare sector, compulsory, voluntary
1. Introduction

The management tool of benchmarking has successfully migrated from the private to the public sector, and is today being adapted to health care organizations. Hospitals and other health care providers are developing ways to compare their practices, processes, and resulting outcomes with other organizations in order to discover “best practices.”

The purpose of using benchmarking in this context is often a direct replication of the motivation for using benchmarking in private sector companies, i.e. learning from others, as a tool in various improvement schemes.

This paper examines the process of benchmarking and the ways in which it can be used to secure improvement and best practice. Classic definitions are presented and their appropriateness to present-day conditions is considered.

The following widely-accepted definition invokes its use as a driver of organizational performance and learning:

“Benchmarking is a positive, proactive process to change operations in a structured fashion to achieve superior performance. The benefits of using benchmarking are that functions are forced to investigate external industry best practices and incorporate those practices into their operations. This leads to profitable, high-asset utilization businesses that meet customer needs and have a competitive advantage (Camp, 1989a, p. 62).”

Camp argues that performance measurement is not an end in itself but a means of comparing practice and the impact of changes in practice intended to secure improvement.

Other reviews by Cox and Thompson (1998), Dorsch and Yasin (1998) and Zairi and Youssef (1995, 1996) traced this growth in benchmarking literature after its first decade and a half ’90. Indeed, new literature reviews have continued to emerge (e.g. Anderson and McAdam 2004; Dattakumar et al. 2003; Yasin 2002) along with new journals and books.
Previous studies show that there are several classifications of benchmarking. For example, taking Camp’s (1995) typology, he developed the concepts of \textit{external}, \textit{internal}, \textit{functional}, \textit{generic benchmarking} (see also Benson, 1994); others authors make a similar distinction between \textit{results and process benchmarking} (Trosa and Williams, 1996); otherwise categorization presented in form of \textit{voluntary and compulsory benchmarking} (Bowerman et al., 2002); another taxonomy provided in the name of \textit{unilateral and cooperative benchmarking} (Elnathan, 1996); and the last categorization made is \textit{implicit and explicit benchmarking} for the differences experienced by organizations, particularly where (inter- or intra-organization) structural variables are involved (Shofield, 1998; Jackson and Lund, 2000).

With regard to what has already been said, the most interesting classification is to analyze compulsory and voluntary benchmarking because one of the successful elements of benchmarking mechanisms is to analyze the important behavior of individuals and organizations. It utility permits to look at the key role of professionals in the healthcare sector.

2. \textbf{Aim of the research}

The goal of this paper is to review the important themes of benchmarking in the health sector focusing on the application of compulsory and voluntary benchmarking.

The following research presented a brief overview of the literature that aims to explain the nature of this tool which are the conditions utilized effectively in healthcare of the effective use of compulsory or voluntary benchmarking, also reporting some examples.

3. \textbf{Compulsory and voluntary benchmarking}

The next table reported some definitions of compulsory and voluntary benchmarking in healthcare.
<table>
<thead>
<tr>
<th>Compulsory benchmarking</th>
<th>Voluntary benchmarking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is imposed by an authority with power to prescribe such an initiative (Ball et al., 2000)</td>
<td>...for implementing collaborative benchmarking efforts and the benefits such efforts bring to the organizations involved...This model allows all participants to learn and improve…helps to go beyond data comparison to the analysis and understanding of key processes that contribute to effective health services delivery (Gift et al., 1994)</td>
</tr>
<tr>
<td>Is characterized by the engagement of public sector bodies in collecting and comparing performance data in the instruction of an external agency (Bowermann et al., 2002)</td>
<td>...provide proof of concept for the feasibility of a novel, voluntary, nationwide system for benchmarking the quality of Breast Cancer (BC) care based on collaboration between a network of specialist breast centres and an independent external provider of scientific data collection and data analysis services (Brucker et al., 2008)</td>
</tr>
<tr>
<td>…requirements, imposed by government policy, are articulated in terms of comparative ideas - benchmarking with the stated objective of “sharing best practices”, but are operationalized and disseminated in the form of indicator league tables with standardized benchmarks for performance (Northcott et al, 2005)</td>
<td>…all provinces and territories voluntarily agreed to participate and all Canadian acute care hospitals (over 600) are participating (Baranek et al., 2011)</td>
</tr>
</tbody>
</table>

Table 1: Two different typologies of benchmarking in healthcare

a) The previous table showed that the definitions of compulsory benchmarking are various. As Bowerman said, this type of comparison has led to focus on the measurable results (Bowerman et al., 2000). The accountability and compulsory are strictly linked because it has developed to support accountability by providing proof of efficiency. In fact the emphasis is placed on the quantitative, and therefore the measurable, aspects of performance.
b) Voluntary benchmarking is distinct from the compulsory model by its development primarily as a management tool. Mosel and Gift (1994) recommend also a collaborative approach to benchmarking in healthcare. Their model uses common work processes as the impetus for joint benchmarking efforts and as a way of reducing costs. Sharing ideas, the workload and resources, and reducing duplication and the isolation of healthcare organizations mutually benefits all participants and ensures that the community benefits from the quality outcomes.

In this case, cooperation and collaborative inter-organizational learning between institutions is at the core of the approach in order to improve procedures and modes of operation. This approach requires a high level of trust and confidentiality between participating institutions. Wolfram et al. (1997) suggested this model can be utilized in order to achieve organizational effectiveness:

Benchmarking can be used as a management tool in response to a number of sector-specific factors, comprising a need to reduce costs, a need to achieve more with the same or fewer resources, in preparation for competitive tendering, a desire to improve the quality of services; to assist in changing culture or behavior.

Evidences lead us to suggest that some voluntary benchmarking is, somewhat paradoxically, undertaken for defensive reasons. Defensive benchmarking proves to an external agency that they are doing well (or are not the worst); the primary aim is not necessarily to become the best. It may be undertaken to protect the organization from potential criticism. In practice, the dichotomy between voluntary benchmarking and defensive benchmarking is not precise (Bowerman et al., 2002). It is, of course, possible to strive to be the best, but at the same time to want to prove (to other parties) how well the organization is doing. As reported Holloway et al. (1999), they gave a good illustration about it in their case study of Warwickshire Ambulance Trust: the defensive benchmarking was used to look at the underlying processes
in order to obtain greater information to refuse possible criticisms arising from compulsory benchmarking of results. From this origin, Warwickshire Ambulance Trust began to develop towards a form of process benchmarking which also allowed them to study practices of other trusts that might result in better patient care.

4. Methodology
A systematic review of literature was conducted. The course of action included the following steps:

(a) The collection of literature has been reviewed from '90 until December 2012 with the following key words “benchmarking & healthcare”. The search was conducted looking at the title and the abstract.

(b) The literature search, using EBSCO (Electronic Business Source Complete) database. Search engine available at www.scholar.google.it was specifically used to ferret the literature pertaining to benchmarking from a wide variety of sources on the internet.

(c) From the articles identified initially, I expanded my search using a snowball sampling method the snowball

(d) The development of a classification scheme was the next step.

Following these criteria 39 articles were found.
The previous Figure 1 presents a flow chart of the method used. This search generated 39 peer-reviewed articles published in English language until December 2012. After review of titles and abstracts, 27 articles excluded and 12 studies remained for a detailed reflection. Of these, I left out 5 papers because were summaries of the literature for the purpose of information/comments or written by more common authors (for example: Bowerman M., Ball A., Francis G., 2001, Benchmarking as a tool for the modernisation of local government, Financial Accountability & Management, Vol.17 No.4); in the last case I chose the article suitable of my research.

Once I have identified publications that fit my search goals, I examined the citation lists for additional 8 relevant publications. It were analyzed 15 full text in order to provide insights to the growth and development of benchmarking concept in healthcare: 6 publications belong to research paper in health, 2 articles related to theoretical framework, 4 papers pertain to case
studies in benchmarking. 3 publications come literature review in benchmarking health. The selection was made as a means to understanding the breadth and range of published applications of benchmarking in healthcare.


It is noted that one-third of the articles are published in “Benchmarking: An International Journal” as research papers.

Furthermore, the reviewed publications highlighted the importance of the maintenance of quality, the need for increased efficiencies and effectiveness of service and how this can reduce costs. There were several key themes that emerged in the review which the discussion integrates according to the following headings: what the study did, how the study was conducted, what was learnt from the experience, and what the implications were for health care in general.

5. Results

A comparison among the earlier attempts to review literature on benchmarking is made using certain attributes. They are shown in Table 2. The following step permits to identify the aim for every article.

<table>
<thead>
<tr>
<th>N.</th>
<th>TITLE, AUTHOR(S), YEAR</th>
<th>AIM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Collaborative benchmarking in a healthcare system, Robert Gift et al., 1994</td>
<td>To focus on the initial phases of the benchmarking process in the healthcare industry. Phases of collaborative benchmarking; Selecting the benchmarking project; Establishing the benchmarking collaborative; Conducting the project within the collaborative; Conducting the project with partners outside the collaborative; Lessons learned</td>
</tr>
<tr>
<td>2</td>
<td>An introduction to benchmark in healthcare, Harold Benson, 1994</td>
<td>To provide a theoretical overview of benchmarking process in healthcare</td>
</tr>
<tr>
<td>N.</td>
<td>TITLE, AUTHOR(S), YEAR</td>
<td>AIM</td>
</tr>
<tr>
<td>----</td>
<td>------------------------</td>
<td>-----</td>
</tr>
<tr>
<td>3</td>
<td>Benchmarking in health care: A review of the literature, Isabel Higgins, 1997</td>
<td>To provide a review of the 10 significant publications related to benchmarking in health care</td>
</tr>
<tr>
<td>4</td>
<td>A Case Study of Benchmarking in the National Health Service, Jacky Holloway et al., 1999</td>
<td>To critique the notion that a single approach to performance improvement can alone be responsible for significant organizational transformation.</td>
</tr>
<tr>
<td>5</td>
<td>Benchmarking in local government under a central government agenda, Amanda Ball et al., 2000</td>
<td>To explore emerging benchmarking practice in the local authority and wider public sector, in particular in UK public sector organizations that it is often driven by central government.</td>
</tr>
<tr>
<td>6</td>
<td>Creating a balanced scorecard for a hospital system, George Pink et al., 2001</td>
<td>To describe the process of creating a balanced scorecard for a hospital system, focusing on the development of performance measures in Ontario.</td>
</tr>
<tr>
<td>7</td>
<td>The evolution of benchmarking in UK local authorities, Mary Bowerman et al., 2002</td>
<td>To analyze the development of benchmarking in the public sector pre-dates its popularity in the private sector</td>
</tr>
<tr>
<td>8</td>
<td>Benchmarking in UK health: a gap between policy and practice?, Daryl Northcott, Sue Llewellyn, 2005</td>
<td>To bring greater clarity to the debate on the merits (or demerits) of relative performance evaluation through a broad assessment of current UK National Health Service (NHS) benchmarking.</td>
</tr>
<tr>
<td>9</td>
<td>What have we learned? Themes from the literature on best-practice benchmarking, Graham Francis, Jacky Holloway, 2007</td>
<td>To review the important themes in the literature on benchmarking, and to assess the contribution to knowledge provided thus far</td>
</tr>
<tr>
<td>10</td>
<td>Benchmarking in Dutch healthcare, Robbert-Jan Poerstamper et al., 2007</td>
<td>To enable its readers to make conscious decisions in benchmarking: Is it worth our while to benchmark? How do we achieve the maximum possible return?</td>
</tr>
<tr>
<td>11</td>
<td>Public sector benchmarking: an application to Italian health district activity plans, Emidia Vagnoni, Laura Maran 2008</td>
<td>To suggest an application of the benchmarking technique on the content evaluation of the health district activity plans</td>
</tr>
<tr>
<td>12</td>
<td>Benchmarking applications in public sector principal-agent relationships, Bjørn Andersen et al., 2008</td>
<td>To explore the range of benchmarking applications that can be used in a principal-agent relationship setting often found in the public sector</td>
</tr>
<tr>
<td>13</td>
<td>Benchmarking the quality of breast cancer care in a nationwide voluntary system: the first five-year results (2003–2007) from Germany as a proof of concept, Sara Y Brucker et al., 2008</td>
<td>To establish a nationwide voluntary collaborative network of breast centres with independent data analysis; to define suitable quality indicators (QIs) for benchmarking the quality of breast cancer (BC) care; to demonstrate existing differences in BC care quality; and to show that BC care quality improved with benchmarking from 2003 to 2007.</td>
</tr>
<tr>
<td>14</td>
<td>Benchmarking health care in federal systems: the Canadian experience, Patricia Baranek et al., 2011</td>
<td>To draw broader conclusions about the challenges and opportunities created by federal contexts.</td>
</tr>
</tbody>
</table>

Table 2: The aim of articles

The attributes that were considered for comparisons are:

- Type of publications covered: they are text books, journal papers, theoretical framework, case study, literature review, research paper
- Main content: this looks at the way in which the literature has been reviewed as new value of publication

- Key words: they are words or phrases often used in the article

Apart from these distinguishing attributes, certain common parameters like the name of the publication, author(s), year of publication, journal of publication are also used.

This comparison is shown in Table 3.
<table>
<thead>
<tr>
<th>YEAR</th>
<th>AUTHOR(S)</th>
<th>TITLE</th>
<th>PUBLISHED IN</th>
<th>KEY WORDS</th>
<th>TYPE OF PUBLICATION COVERED</th>
<th>MAIN CONTENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1994</td>
<td>Robert Gift et al.</td>
<td>Collaborative benchmarking in a healthcare system</td>
<td>Healthcare Management</td>
<td>Health services administration</td>
<td>Research paper</td>
<td>Collaborative benchmarking within a healthcare system offers participating organizations several benefits. It brings together organizations interested in creating breakthrough improvements in common processes.</td>
</tr>
<tr>
<td>1994</td>
<td>Harold Benson</td>
<td>An introduction to benchmark in healthcare</td>
<td>Radiology management</td>
<td>Quality Improvement</td>
<td>Theoretical framework</td>
<td>Is the first in a series of articles discussing operational benchmarking as a quality improvement tool for health care institutions</td>
</tr>
<tr>
<td>1997</td>
<td>Isabel Higgins</td>
<td>Benchmarking in health care: a review of the literature</td>
<td>Australian Health Review</td>
<td>Quality Improvement</td>
<td>Literature review</td>
<td>It highlights the importance of the maintenance of quality health care, the reduction of health care costs and the need for improved efficiency and effectiveness in providing health care.</td>
</tr>
<tr>
<td>1999</td>
<td>Jacky Holloway et al.</td>
<td>A case study of benchmarking in the National Health Service</td>
<td>The Open University Business School</td>
<td>Benchmarking improvement, performance measures</td>
<td>Case study</td>
<td>It argues that complex approaches to performance improvement such as benchmarking, however technically powerful they may be, are only as effective as the people who apply them and their compatibility with the organizational context in which they are used.</td>
</tr>
<tr>
<td>2000</td>
<td>Amanda Ball et al.</td>
<td>Benchmarking in local government under a central government agenda</td>
<td>Benchmarking: An International Journal</td>
<td>Benchmarking, Value, performance measurement, Central government</td>
<td>Research paper</td>
<td>It is a conflation of two distinct views of benchmarking: benchmarking as a rigorous and challenging scrutiny of local government processes; and benchmarking as an instrument of central government control.</td>
</tr>
<tr>
<td>2001</td>
<td>George Pink et al.</td>
<td>Creating a balanced scorecard for a hospital system</td>
<td>Journal Health Care Finance</td>
<td>Balanced scorecard, hospitals, performance measurement</td>
<td>Case study</td>
<td>In 1999, hospitals in Ontario, Canada, collaborated with a university-based research team to develop a report on the relative performance of individual hospitals in Canada's most populated province.</td>
</tr>
<tr>
<td>2002</td>
<td>Mary Bowerman et al.</td>
<td>The evolution of benchmarking in UK local authorities</td>
<td>Benchmarking: An International Journal</td>
<td>Benchmarking value, performance measurement, local government</td>
<td>Literature, Research paper</td>
<td>The reasons for benchmarking in the public sector are confused; pressures for accountability in the public sector may militate against real performance improvement; and an appropriate balance between the use of benchmarking for control and improvement purposes is yet to be achieved.</td>
</tr>
<tr>
<td>2005</td>
<td>Daryl Northcott, Sue Llewellyn</td>
<td>Benchmarking in UK health: a gap between policy and practice?</td>
<td>International Journal of Management Reviews</td>
<td>Benchmarking, government policy, NHS</td>
<td>Research paper</td>
<td>The findings are relevant to both NHS policy makers and to NHS actors who must engage with the processes and outcomes of benchmarking practices.</td>
</tr>
<tr>
<td>2007</td>
<td>Graham Francis, Jacky Holloway</td>
<td>What have we learned? Themes from the literature on best-practice benchmarking</td>
<td>Benchmarking: An International Journal</td>
<td>Benchmarking, performance management tool</td>
<td>Literature, Research paper</td>
<td>Four themes from the literature are highlighted: studies of the nature of benchmarking practice; criticisms of benchmarking; evaluating the effectiveness of benchmarking; and the notion of best practice.</td>
</tr>
<tr>
<td>2007</td>
<td>Robbert-Jan Poerstemper et al.</td>
<td>Benchmarking in Dutch healthcare</td>
<td>PricewaterhouseCoopers</td>
<td>Quality of care</td>
<td>Conceptual paper</td>
<td>It focuses on the benchmarks that PricewaterhouseCoopers has conducted with other consultancies and agencies in the Dutch healthcare and related sectors over the past decade.</td>
</tr>
<tr>
<td>YEAR</td>
<td>AUTHOR (S)</td>
<td>TITLE</td>
<td>PUBLISHED IN</td>
<td>KEY WORDS</td>
<td>TYPE OF PUBLICATION COVERED</td>
<td>MAIN CONTENT</td>
</tr>
<tr>
<td>------</td>
<td>-----------</td>
<td>-------</td>
<td>--------------</td>
<td>-----------</td>
<td>-----------------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>2008</td>
<td>Emidia Vagnoni, Laura Maran</td>
<td>Public sector benchmarking: an application to Italian health district activity</td>
<td>Benchmarking: An International Journal</td>
<td>Benchmarking, Corporate strategy, Health services sector, Public sector</td>
<td>Research paper</td>
<td>The paper highlights the role of benchmarking both in providing a comprehensive lecture key of the reality of district planning and suggesting cues from the best practices observed.</td>
</tr>
<tr>
<td>2008</td>
<td>Sara Y Brucker et al.</td>
<td>Benchmarking the quality of breast cancer care in a nationwide voluntary system: the first five-year results</td>
<td>Bio Med Central Cancer</td>
<td>Quality indicators, cancer</td>
<td>Theoretical framework</td>
<td>The benchmarking system described allows both comparisons among participating institutions as well as the tracking of changes in average quality of care over time for the network as a whole. Marked QI increases indicate improved quality of BC care.</td>
</tr>
<tr>
<td>2008</td>
<td>Bjørn Andersen et al.</td>
<td>Benchmarking applications in public sector principal-agent relationships</td>
<td>Benchmarking: An International Journal</td>
<td>Benchmarking, public sector organizations, performance measures</td>
<td>Conceptual paper</td>
<td>Whereas most work on benchmarking in the public sector tend to view only improvement-oriented, voluntary benchmarking as relevant and useful, this paper demonstrates how many different imposed benchmarking schemes with other purposes can be useful.</td>
</tr>
<tr>
<td>2011</td>
<td>Patricia Baranek et al.</td>
<td>Benchmarking health care in federal systems: the Canadian experience</td>
<td>Benchmarking health care in federal systems</td>
<td>Improving quality, safety and access, planning of health services</td>
<td>Case study</td>
<td>It highlights the characteristics (processes, outcomes, challenges and opportunities) of benchmarking experiences to date.</td>
</tr>
</tbody>
</table>

Table 3: A summary of literature on benchmarking healthcare

In order to compare two models, it was found *how many articles were voluntary and compulsory in healthcare*. As the reader can see (graph 1-2) the most of studies considered both concepts, 5 dealt about compulsory model, 4 articles were related to voluntary benchmarking in healthcare.
Benchmarking found application in the healthcare scenarios, especially used for the national assessment services of various countries (Vagnoni, Maran, 2008).
In relation to heading “WHERE and WHEN the studies were done” as figure 2 showed that the majority part of studies were conducted in Anglo-Saxon and USA countries in ‘90. Some European countries (Italy, Germany, Netherlands and Norway) developed after half 2000.

Figure 2: Timeline and countries

Until 2000 some authors speak about compulsory benchmarking where it was applied in the UK and Italy countries; and after 2000 other authors focusing on voluntary model, as in Canada and in Germany. In the last country Brucker et al. (2008) established a voluntary benchmarking network of breast centres in collaboration with an external, independent service provider commissioned to collect and analyze the Quality Assurance data and the success of the benchmarking system as a whole.

The research provides insights about the nature of benchmarking and its contribution to performance improvement in public health, and hereinafter, I reported the experiences of UK, Italy and Canada countries in order to show the differences of two models.

In the UK, for example, a compulsory model is used. The English NHS, where the development of benchmarking took place in a backdrop of fiscal or other legislative controls emanating from central government, is apparently singled out as a highly advantageous
management tool by central government for improving services: such a view is simplistic as it ignores the entanglement of benchmarking in the wider policy process.

In fact, Holloway et al. (1999), for example, suggest that: “Knowing one’s position in the league table does little to enable the organization to understand how better performers achieved their status…”.

It is often driven by central government (Holloway et al., 1999): independent reports on “best practice” have been produced for long by the Audit Commission and National Audit Office. Substantial analyses of the uses of benchmarking in the UK, in healthcare sector or public sector, (Llewellyn and Northcott, 2005, Bowerman et al. 2002) put on evidence that political context influences the use at benchmarking reducing the learning purposes of this technique.

One of the factors of failure indicated by Northcott and Llewellyn in the application of benchmarking in the English NHS, as a tool to encourage learning through best practices, is the perception that the Trust had in having to compare this as compulsory.

Also Ball et al. (2000) said that the performance monitoring in local government is a critical factor, which in turn results from the controlling nature (in fiscal and political terms) of UK central government. The result is a conflation of two distinct views of benchmarking: benchmarking as a rigorous and challenging scrutiny of local government processes and benchmarking as an instrument of central government control. Such a state of affairs would appear to offer a number of advantages to those policy makers whose plan for the reform of local government encourages the use of benchmarking across the sector.

This model showed how regional priorities involve the structure and the characteristics of the performance evaluation systems: the influence of behavioral and contingency theories can be seen in such critiques of “compulsory” public sector benchmarking policies.

Also in Italy, although there is enthusiasm about benchmarking across Regions, this technique is not commonly applied within regional boundaries as governance tool.
Even though most of the Regions of Italy declare to be willing to compare their performance with others, they show some reserve on how benchmarking should be done (Vainieri, Nuti, 2011). Some Regions declared that benchmarking should be done by National Government after having shared the selection of indicators, some say that the comparison should be run by an external benchmarking agency, others prefer having a regional supervision on how to run comparison, finally someone asks only for a comparison on methodology.

Figure 3 summarizes the regional positions, pointing out the different visions that go from a regional system (where there is maximum autonomy on measuring performance, no benchmarking across Regions) to a national system (where everything is decided and done by National Government). On one hand we can classify the first case as voluntary models, on the other hand as a mandatory form.

Source: Vainieri, Nuti, 2011

Figure 3: Different visions on benchmarking

One example of study that reports the application of voluntary benchmarking is in Ontario healthcare.

An interesting example is based on the work initiated by Baker and Pink in the 1990s (Pink et al., 2001) of how hospitals in Ontario (Canada) chose and used Balanced Scorecard measures,
in order to provide information that hospitals can use to identify areas of needed improvement.

It illustrates how hospitals perform in areas of patient satisfaction and financial performance, where they have achieved successes and where they can make improvements to enhance patient care. This study was published in “The Hospital Report” series that was born with the partnership among Ontario hospitals and the Government of Ontario. All provinces and territories voluntarily agreed to participate and all Canadian acute care hospitals (over 600) are participating. Since April 2012, the results for 30 indicators (21 clinical indicators and 9 financial performance indicators) are publicly available at www.cihi.ca (Baranek et al., 2011).

The voluntary participation of the institutions has meant that they agreed to make available to researchers their health data and financial purchase of services to achieve customer satisfaction surveys users with agreed methodologies and finally to provide much information about the operating methods they adopted. The rating system on a voluntary basis, used in Ontario, has involved 98% of acute care hospitals, the 90% of First Aid and between 70% and 85% of all other types of institutions. Despite the public funding, the government has never pushed to transform the system.

Its aim is to help people in Ontario better understand and assess the performance of their local hospital and of the province’s hospitals as a whole (all reports are available on the Hospital Reports’s website at www.hospitalreport.ca). It also supports efforts by hospitals to improve the quality of their services. This permitted to provide a concrete evaluation tool for hospital administrators to guide decision-making. They also give the public with an understanding of their local hospital’s performance compared with its peers: it makes Ontario an international leader in accountability.

Indeed the rationale of voluntary benchmarking is clearly identifiable with an emphasis on continual improvement and striving to be the best.
According to some authors, voluntary benchmarking ensures substantial change processes rather than formal ones (Ball et al., 2000).

The publication of findings and the sharing of information both formally and informally provide an understanding of a range of aspects of best practice and double approach to improve benchmarking process. Particularly interesting are the cases of Tuscany and Lombardy that both use benchmarking as learning tool among health authorities. Indeed while the former applies benchmarking to all indicators in a full transparent way, the second uses it especially for outcome indicators keeping clear the label of health authorities (Vainieri, Nuti, 2011) through the P.Re.Val.E. (Fusco et al., 2012).

6. Discussing and concluding comments

Benchmarking has to be understood in its sector-specific context.

There is strong evidence that public bodies are capable of implementing innovation and meaningful benchmarking approaches. The government’s encouragement (or compulsion) of benchmarking has to achieve an appropriate balance, however, between control, for accountability purposes, and facilitation, for performance improvement purposes.

The contents of this study will contribute to encourage managers to apply the concept of benchmarking depending on the competitive or collaborative environment. So we concluded that:

a) the research seems to show that the most typical way of employing benchmarking is as a voluntary management tool aimed at achieving improvements. Although some authors have discussed other usages, there are references to attempts to classify benchmarking in the healthcare sector as either voluntary or compulsory, and the compulsory approach is said to be flawed (Bowerman et al., 2002; Andersen et al., 2008);
b) Nevertheless the study has been to demonstrate that compulsory type of benchmarking is also required as Italian regions highlighted.

c) However even in the compulsory side it seems there can be different degree of application in any eventual framework for central government control.

The collective implications of these studies highlight the key theme: in the healthcare sector the benchmarking process developed to improve the quality of the healthcare service locally, nationally and internationally; at the same time containing costs and improving efficiency and effectiveness of the service, can only be achieved through striving for best practice. From a practical perspective, this means that all healthcare organizations should benchmark.

Further research is undoubtedly necessary to unravel the different types and purposes of benchmarking and to consider the implications for disclosure at different levels.
References

Andersen B., Henriksen B., Spjelkavik I., 2008, Benchmarking applications in public sector
principal-agent relationships, Benchmarking: An International Journal, Vol. 15 Issue 6,
pp.723-741

Anderson K., McAdam R., 2004, A critique of benchmarking and performance measurement:

Ball A., Bowerman M. and Hawksworth S., 2000, Benchmarking in local government under a

Baranek P., Veillard J., Wright J., 2011, Benchmarking health care in federal systems: the
Confederations 2011

Benson H., 1994, An introduction to benchmark in healthcare, Radiology Management,
Vol.16 No.4, pp. 35-39.

authorities, Benchmarking: An International Journal, Vol. 9 No. 5, pp 429-49

Brucker S.Y, Schumacher C., Sohn C., Rezai M., Bamberg M., Diethelm Wallwiener and the
Steering Committee, 2008, Benchmarking the quality of breast cancer care in a nationwide
voluntary system: the first five-year results (2003–2007) from Germany as a proof of
concept, Bio Med Central Cancer

Camp R., 1989, Benchmarking: the search for best practices that lead to superior
performance, part I, Quality Progress, pp. 61-8.


CHAPTER 3 – Assessment and improvement of the Italian Healthcare system: first evidences from the national performance evaluation system

Background: The Italian National Health System (NHS) follows a Beveridge model and provides universal health care coverage to the whole population. The strong policy of decentralization, that has been taking place since the early 1990s, has gradually shifted powers from the state to the 21 Italian regions. These regions now have political, administrative, and financial responsibility regarding the provision of health care while the Italian State retains limited supervisory control and continues to have overall responsibility for the NHS to assure uniform and essential levels of health services across the country. However, this devolution process contributed to accentuate the interregional disparities in the provision and quality of health care services. In this context, it has become fundamental, both for the Ministry of Health and for Regions to adopt a common performance evaluation system (PES).

Purposes: The paper reports the first evidences and results following the adoption of the PES and shows how this tool can be useful for policy makers both at the national and regional level.

Methodology: The pilot Italian National PES consists of 34 indicators of performance measured allowing benchmarking across the 21 regional healthcare systems and their healthcare organizations. A target chart was designed as a specific tool to report regional results. Moreover, statistical analyses have been carried out to find out the correlation among the selected indicators and between overall performance and costs.
**Findings and Practice implications:** The pilot Italian PES could be considered as strategic tool in the health care sector supporting the Ministry of Health to assure a uniform level of care to the population and the regional managers to facilitate strategy implementation and improvement initiatives by the benchmarking process.

**Keywords:** performance evaluation system, benchmarking, assessment, improvement.
1. Introduction
A number of factors, including rising costs, technological advancements, population aging, and medical errors, have contributed to the adoption of performance measurement systems by many industrialized countries (Smith 2002; Arah et al. 2006; Kelley and Hurst 2006). The introduction of New Public Management (NPM) principles in the 1980s promoted a number of reforms in order to drive a more efficient, effective, and accountable public sector (Hood 1995a; Lapsley 1999, Saltman and Busse 2002; Saltman and Vrangbaek 2007). OECD countries have applied these principles in different ways with different emphasis (Hood 1995b). Among the NPM principles, the one asking the public sector to adopt more explicit and measurable standards of performance measurement has motivated countries to create different performance measurement systems. However, in the 1980s the use of performance measurement systems primarily focused on financial measurements and was therefore unable to help organizations achieve multiple strategic objectives or drive change (Pollitt 1985; Ghabadian and Ashworth 1994; Guthrie and English 1997; Lorden, Coustasse, and Singh 2008). Consequently, the adoption of a sophisticated and comprehensive multidimensional performance evaluation system, such as the balanced scorecard, which looks beyond traditional financial measures, has been suggested (Linard et al. 2000; Eccles 1991; Jackson 1993; Kloot and Martin 2000; Fottler, Erickson, and Rivers 2006; Yang and Tung 2006). In particular, within the healthcare sector, a growing number of studies describe the adoption of a multidimensional performance evaluation system by a broad range of healthcare organizations, including single providers and entire national health systems (e.g., Linard et al. 2000; Aidemark, 2001; Zelman, Pink, and Matthias 2003; WHO 2003; Asbroek et al. 2004; Chang, 2007; Ba-Abaad, 2009, McLoughlin et al. 2001).

Within the Italian healthcare system, the need for performance measurement has grown in urgency since the early 1990s when the government approved the first reform of the National Health Service (Legislative Decrees 502/1992 and 517/1993) (Lo Scalzo et al. 2009).
This was a period when national reforms started transferring several important administrative and organizational responsibilities from the central government to the 21 regions, with the aim of making regions more sensitive to the need to control expenditure and promote efficiency, quality, and citizen satisfaction. This devolution process was recently enforced by the act regulating fiscal federalism (law 42/2009) delivered by the Italian parliament in 2009, which provided regions with significant autonomy in organizing healthcare services, allocating financial resources to their local health authorities (LHAs), and in monitoring and in assessing performance (Formez 2007; Censis 2008; Nuti 2008; Antonini and Pin 2009). The central government retains overall responsibility for ensuring that services, care, and assistance are equitably distributed to citizens across the country. However, this strong policy of decentralization, along with a series of rationalization measures, has contributed to accentuating the already existing interregional disparities in healthcare, especially the north–south divide, and have undermined the egalitarian principles of the National Health Service. Italian regions differ for historical and geographical reasons. The sharpest division is between north and south. The north has one of the most advanced industrial societies in the world, whereas the south, which encompasses the area of seven regions (Campania, Molise, Basilicata, Calabria, Puglia, and the two islands, Sicily and Sardinia) is, by contrast, one of the most economically depressed areas in Europe. In this context, at regional level, only a few Italian regions have adopted systems able to measure performance (Nuti 2008; Vittadini 2010; Provincia Autonoma di Trento 2011; Barone et al. 2006–2008). At a national level, the Italian NHS has had to face the challenge of creating and developing efficient systems capable of preserving both the principles of egalitarianism and high-quality services. In recent years a pilot multidimensional national PES has been developed. This program aims to reduce discrepancies and focus on the intrinsic goals of the NHS by monitoring the capacity of each
region to guarantee its citizens, regardless of their social status, equal access to essential health services while maintaining quality, efficiency, and appropriateness.

Given this background, the purpose of this article is to describe and illustrate the development and design of the first Italian health performance measurement system, which is intended to assess the dimensions of performance at both the national and regional levels.

We next present the construction and development of the pilot conceptual framework, followed by a description of the PES reporting system. Finally, we report policy implications and conclusions.

2. The Italian Performance Evaluation system: the conceptual framework

As a result of the ongoing process of devolution of power, the Italian NHS is currently organized on the basis of two levels: the central government, which has planning and funding responsibilities, while ensuring that all citizens have uniform access to healthcare, and the 21 regional governments, which organize and supervise the provision of healthcare services within their jurisdiction and allocate overall financial resources to the productive units—approximately, in 2009, 146 local health authorities and 100 independent hospitals across the country\(^1\). Each LHA, under the supervision of the corresponding regional government, is directly responsible for the provision of comprehensive care to its entire resident population, regardless of income or occupational status (Ferrario and Zanardi 2010).

In 2009, the Ministry of Health decided to take up the challenge in assessing the performance of the health services provided by the regions. The pilot national performance evaluation system was designed and implemented by the research team of the Laboratorio Management e

\(^1\) These values derive from the national health providers’ register, 2009.
Sanità (MeS Lab) in accordance with the National Agency for Regional Health Care Services (Agenas), which is in charge of its further development\textsuperscript{2}.

This choice was based on the fact that the MeS Lab has had extensive experience in evaluating the performance of healthcare services. Since 2005, it has designed and implemented, in Tuscany, a multidimensional performance evaluation system to monitor the performance of all Tuscan Health Authorities (Nuti 2008; Nuti 2010; Nuti et al. 2010; Nuti et al. 2009; Nuti et al. 2011). Moreover, since 2008 other Italian regions (Liguria, Piedmont, Umbria, Aosta Valley, Marche, Basilicata, and Trento and Bolzano autonomous provinces) decided to adopt the same system, so that today almost half of Italian regions use the same framework in order to assess the performance of their healthcare services.

In the first phase of the development of a pilot national PES, the research team started from the Tuscan experience and carefully tailored the regional PES to meet the national health system’s needs and strategies. The set of measures to be chosen should, in fact, reflect a robust picture of the healthcare performance that can be reliably reported across regions using comparable data. It should be appropriate to support the national health planning process by informing decision makers of how resources and health services are provided to a population and by highlighting inequalities within regional health systems. After having discussed the strategic goals of the national health system with national policy makers and Agenas, the selected indicators and dimensions were derived and sent to all regions. On the basis of their feedback comments, the indicators and dimensions were further modified.

\textsuperscript{2} The National Agency for Regional Healthcare (Agenzia nazionale per I servizi sanitari regionali, www.agenas.it) carries out its activities in close collaboration with the Ministry of Health and with the regions and participates in research programs funded by the MoH.
Selection of Domains of Indicators

As already stated, the main NHS goal is to ensure that the delivery of healthcare should be equally provided to the population across the 21 regions and across all the levels of care that constitute the health system. As a consequence, the national performance indicator framework should therefore include performance indicators for benchmarking within the following three domains (settings) of care: hospitals, primary care including pharmaceutical care, and public and preventative health. Indicators should then be chosen and developed to provide information about the performance of the regional systems across the three domains in terms of the following:

- Quality of the services delivered to citizens, according to the specific goal of each level of care, in order to ensure that patients receive safe, prompt, and correctly delivered services.
- Equity, which deals with potential performance differences across and within regions. Differences across providers for the same indicator should result from epidemiological issues and not management or professional capabilities.
- Appropriateness: Each patient should receive nothing more but also nothing less than what is required.
- Efficiency, which means achieving desired results with the most cost effective use of resources (Donabedian 2003).

Selection of Performance Indicators

In the second phase of the framework development, the team selected indicators reflecting the importance of the dimensions outlined above. Starting from the Tuscan PES, the initial comprehensive list was reduced to a more manageable group of performance indicators. Many performance indicators, although of relevance regarding the analysis, remain not calculable because of the reliance on only a few data sources, while some other indicators
were not considered in order to avoid information overload. This limited set of measures should be able to give the Ministry of Health (MoH) a rapid and comprehensive overview of the NHS and its differences across regional healthcare systems. This is important given the MoH’s role of supervisory control and its overall responsibility for the NHS in guaranteeing uniform and essential levels of health services across the country.

Each indicator was tested to determine its suitability for measuring NHS performance by means of iterative discussions, first with the MoH and then with regional representatives. A transparent, consensus-based process is critical, because there are frequent decisions involving trade-offs between accuracy and reliability, data availability and comparability across the different regional healthcare systems. In particular, the effort to identify the core set of performance indicators consisted of the following steps:

1) Starting from international evidence and the already developed and tested measures of the Tuscan PES, the research team identified and then calculated 38 performance measures at regional and provider levels.

2) The research team shared the list and results of the performance measures with the MoH.

3) The MoH officially showed to the 21 regional health councillors the above-mentioned list and results.

4) The research team was then asked by the MoH to collect and analyze all the comments coming from the regions.

Finally, after an intensive period of interaction between the research team and the regional representatives, a final set of 34 indicators out of the initially presented 38 was chosen (Table 1).
In particular, the following four indicators were excluded from the initial list:

1. Hospitalization rate for gastroenteritis (2–17 years old)
2. Hospitalization rate for pneumonia
3. Per capita cost for a defined daily dose (DDD) of a drug
4. Health services per inhabitants by homogeneous groups of clinical diagnosis that are provided within the confidence intervals of the national average

These indicators were not included either because they couldn’t be applied uniformly to the different regional healthcare systems (the first two indicators) or because they did not receive the consensus by all the regional representatives. Most of the 34 indicators were derived either from the framework already developed by Tuscany region or from international experience (OECD 2003; WHO 2003; AHRQ 2006; CIHI 2001; Department of Health 2008) and were selected on the basis of the following criteria (Kelly and Hurst 2006):

- The importance of what is being measured in terms of policy relevance
- The scientific soundness of the measure in terms of its validity, reliability, and the explicitness of the evidence base
- The feasibility and cost of obtaining nationally comparable data for the measure

The Italian hospital discharges database for the years 2007 and 2008 was used for all the measures belonging to the hospital and primary care dimensions, the 2008 OsMed report (Gruppo di lavoro OsMed 2009) for the indicators regarding pharmaceutical care, the 2008 national screening report, and the MoH database for the measures related to public health and prevention.

Avoidable hospitalization rates for chronic conditions from inpatient data were used as a proxy of primary care performance because of the lack of national comparable sources on territorial services (Ricketts et al. 2001). Moreover, indicators from hospital inpatient data,
where possible, have been standardized according to sex and age using the Italian residents of the year 2001 as a standard population. Finally, all the indicators based on hospitalization data were calculated at both regional and interregional levels (all the providers within each region), while pharmaceutical and public health and prevention care data were only available at aggregated (regional) level. Starting from the final indicator list, 23 indicators out of the total 34 were chosen to be evaluation measures and were assigned performance assessment ratings in order to allow for regional benchmarking. This means that the 21 regions were divided into quintile groups on the basis of the distribution of each evaluation measure in order to derive five different levels for defining the indicator performance of each region categorically from worst to best. This approach allows the MoH to benchmark the Italian regions in terms of their performance results across the selected dimensions of care.

The other 11 indicators were considered useful to provide further insights into each regional performance level and were therefore defined as observational indicators.
<table>
<thead>
<tr>
<th>Indicator code</th>
<th>Indicator description *</th>
<th>Evaluation indicators</th>
<th>Observational indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>H1.1</strong></td>
<td>Ordinary hospitalization rate - acute admissions</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td><strong>H1</strong></td>
<td>Global hospitalization rate - acute admissions</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td><strong>H1.2</strong></td>
<td>Day-Hospital hospitalization rate - acute admissions</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td><strong>H1.3</strong></td>
<td>Mean DRG weight - acute admissions</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Efficiency</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>H2</strong></td>
<td>Case-mix adjusted length of stay - surgical DRG</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td><strong>H2.1</strong></td>
<td>Case-mix adjusted length of stay</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td><strong>H2.2</strong></td>
<td>Case-mix adjusted length of stay - medical DRG</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td><strong>H13</strong></td>
<td>Pre-op LOS - planned admissions</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Surgical Appropriateness</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>H3</strong></td>
<td>Percentage of medical DRG from surgical departments</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td><strong>H4</strong></td>
<td>Percentage of laparoscopic cholecystectomies in day surgery or 0-1 day admissions</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td><strong>H5</strong></td>
<td>Surgical Essential Levels of Health Services DRG - standard percentage achieved</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Medical Appropriateness</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>H6</strong></td>
<td>Medical Essential Levels of Health Services DRG: hospitalization rate</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td><strong>H14</strong></td>
<td>Percentage of short medical hospitalizations</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td><strong>T9</strong></td>
<td>Percentage of medical DH admissions with diagnostic aim</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Clinical Quality</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>H9</strong></td>
<td>Percentage of births</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td><strong>H10</strong></td>
<td>Percentage of readmissions within 30 days for the same MDC corrected by the hosp.rate</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td><strong>H10.1</strong></td>
<td>Percentage of readmissions within 30 days for the same MDC</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td><strong>H10.1.1</strong></td>
<td>Percentage of readmissions within 30 days for the same MDC - medical DRG</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td><strong>H10.1.2</strong></td>
<td>Percentage of readmissions within 30 days for the same MDC - surgical DRG</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td><strong>H11</strong></td>
<td>Percentage of femur fractures operated within 2 days</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td><strong>H12</strong></td>
<td>Regional outflow</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Effectiveness of chronical disease management</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>T2</strong></td>
<td>Hospitalization rate for heart failure (50-74 years old)</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td><strong>T3</strong></td>
<td>Hospitalization rate for diabetes (20-74 years old)</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td><strong>T4</strong></td>
<td>Hospitalization rate for COPD (50-74 years old)</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Pharmaceutical prescription efficiency</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>AF5</strong></td>
<td>Per capita gross pharmaceutical expenditure</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td><strong>AF5.1</strong></td>
<td>Gap between the districtual use of pharmaceuticals and national median</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td><strong>AF5.2</strong></td>
<td>Percentage of the districtual expenditure for equivalent pharmaceutical on the total net expenditure</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td><strong>AF5.3</strong></td>
<td>Percentage of the districtual expenditure for equivalent pharmaceutical on the total DDD</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>PREVENTIVE CARE (P)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>P1</strong></td>
<td>Flu vaccine coverage rate</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td><strong>P2</strong></td>
<td>MPR vaccine coverage rate</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td><strong>P3</strong></td>
<td>Mammography screening extension</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td><strong>P4</strong></td>
<td>Compliance with mammography screening</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td><strong>P5</strong></td>
<td>Colorectal screening extension</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td><strong>P6</strong></td>
<td>Compliance with colorectal screening</td>
<td>x</td>
<td></td>
</tr>
</tbody>
</table>

* Indicators in italics are not evaluation

Table 1: The first set of indicators of the national PES
3. **The PES reporting system**

The data collected from the PES were translated into a color-coded reporting tool that indicates to users at a glance underachievement and exceptions of each region on the basis of only the 23 evaluation measures. The target chart, which was developed within the Tuscan healthcare system in 2005 (Nuti 2008; Nuti et al. 2009), was then chosen as a reporting tool due to its user-friendly interface features. This specific performance graph was designed to provide a visual representation of performance results across all the indicators, thus enabling managers to quickly ascertain whether the regional health system is performing up to standard. This chart is divided into five bands on the basis of the five performance levels, each with its own color, from dark green - corresponding to excellent performance - to red - poor performance.

Within each regional target, the closer the evaluation indicator is positioned to the center of the target, the higher its performance level.

Figure 1 shows, as examples, the targets of three Italian regions, one from the north of Italy (Veneto region), one from the center (Umbria region), and the last one from the south (Campania region). White circles on the target depict performance measures for each region on a particular indicator, with those on the dark and light bands indicating respectively worse and better performance. Target analysis confirms, at a glance, the existence of a clear division between the north and south of Italy: Best performances are all concentrated in northern-center regions while critical situations are in the south. The figures make clear that indicators in the targets of both Umbria and Veneto are all close to the center, indicating a good global performance, while, on the contrary, most of the southern regions’ indicators are concentrated on the target boundaries, where the worst evaluation assessments are positioned.

---

3 All the regional performance results and targets are available (in Italian) on the following website: [www.salute.gov.it/dettaglio/phPrimoPianoNew.jsp?id=273&area=ministero& colore=2](http://www.salute.gov.it/dettaglio/phPrimoPianoNew.jsp?id=273&area=ministero&colore=2).
The MoH received the target diagrams and tables and graphs reporting the values of all 34 indicators measured where it was possible, at the regional and provider level.

As an example, Figure 2 shows the 2008 case-mix adjusted length of stay (LOS) indicator for all 21 Italian regions. This indicator is given as the sum of all differences between the average regional LOS of each diagnosis-related group (DRG) and the average national LOS for the same DRG (Servizio Sanitario dell’Emilia Romagna 2009). This indicator is a measure of efficiency that enables evaluation of the potential reduction in hospital days if each region had an average LOS equal to the benchmark value (here the whole nation).

Figure 1: Tuscany, Veneto, Calabria and Campania 2008 Targets

The indicator was restricted only to surgical DRGs because the LOS of medical DRGs (and thus efficiency) can be affected by inappropriate hospitalization rate. (This phenomenon is noticeable in the southern regions, which have a very high hospitalization rate for medical DRGs.)
Results of the indicator show that in 2008, the region with the worst performance is Lazio, with LOS longer than the national average by 1.1 days, while Emilia-Romagna has the shortest LOS of all regions (indicator equal to -0.7), followed by Tuscany where LOS is shorter than the national value by about 0.6 days.

Figure 2: Case-Mix Adjusted Length of Stay Indicator at Regional Level, Year 2008

As already mentioned, the distribution of inpatient indicators is also studied at a provider level, thus also allowing the MoH to monitor discrepancies in the provision of the set of essential healthcare services within each region. As an example, Figure 3 shows the 2008 case-mix adjusted length of stay indicator for all the Tuscan healthcare providers.

Descriptive results of LOS at aggregate and provider levels highlight the presence of intra- and inter-regional differences in the distribution of the indicator values. This suggests that regions with values of the case-mix LOS index above the national average and with high internal variability have room to improve their efficiency in hospital care and thereby release funds to be invested in other healthcare services.
For instance, Tuscany has calculated that if, in 2009, all its own health units with a LOS (per DRG) higher than the regional mean (per DRG) had reduced its value to the regional mean, the region would have saved around 59 million euros, which potentially could have been reallocated to other services (Nuti, Vainieri, and Bonini 2010).

![Figure 3: Case-Mix Adjusted Length of Stay Indicator by Tuscan Health Providers](image)

**Case-mix Adjusted Length of Stay Indicator by Tuscan Health Providers**

Figure 3: Case-Mix Adjusted Length of Stay Indicator by Tuscan Health Providers, Year 2008

### 4. Policy implications

The PES, together with its reporting system, represents a powerful tool for both regional and national levels. The PES is used at a national level to monitor performance in order to guarantee essential healthcare services to the whole population and at a regional level to benchmark health authorities’ performance and thus to learn from the best practices (Johnston 2004; France 2008).
Moreover, the target chart as a reporting tool can support managers as it enables decision makers to:

1. build up a more comprehensive picture of the strengths and weaknesses of each regional healthcare system through the ability to integrate a wide range of relevant information,
2. identify problematic areas and thus allow the targeting of policy and practical interventions more effectively,
3. rapidly benchmark performance across the various regional healthcare systems, and
4. make performance reports more user friendly for nonexpert users.

Another important feature of the national PES is that it may be considered a valid tool of communication between national and regional levels.

In fact, other European experiences with the introduction of multidimensional performance systems, such as the balanced scorecard, show that multidimensional systems may represent valid tools of communication among actors (Aidemark 2001).

Given the importance of informing health system stakeholders and the public about the performance of regional health systems and their local organizations, the MoH decided to make regional targets and tables publicly accessible through the ministry website (www.salute.gov.it). The MoH has used this public disclosure to shake the regional systems and encourage them to undertake improvements. It is the first time, in Italy, that the MoH has displayed the performance evaluation of regional health services on its website using the most updated performance data. The public exposure of regional performance (and their local healthcare organizations) and the user-friendly reading of the performance evaluation through the target charts should powerfully enhance the accountability process between regions and citizens. Indeed, the PES may help citizens evaluate their local policy makers.
In addition, the PES can contribute to the current debate on how to define “the standard cost” introduced by the recent Italian law 42/2009, which enforced the fiscal federalism process. The standard cost has a central role in fiscal decentralization, as each region, in order to satisfy the residents’ estimated health needs, must count on its own internal revenues and eventually on a national solidarity fund. Each region, however, may draw on this fund only if it delivers health services in line with national standards.

The PES can provide regions with national standards in terms of quality, volumes, and appropriateness on the basis of the average of regional best performers for selected indicators. Thus the standard costs may be calculated on the basis of the performance of the best practice regions (Nuti et al. 2011).

5. Conclusions
The performance evaluation system has become a public policy tool that helps the national government evaluate its strategic policy and promote a “managed” competition among regions. The information dealt with and uniformly represented can enable an efficient and constructive benchmarking process among Italian regions and health authorities. The proposed PES seems to reach a fair equilibrium between the regional governments’ need to control the local health institutions and the local institutions’ need to control their own performance (Greener 2003).

Regions must be encouraged to measure their performance at a local level, creating an appropriate culture of evaluation and learning, focusing attention not only on cost control, but also on quality and appropriateness. The national government can support this process by coordinating a benchmarking system, both at regional and upper-regional levels. This will allow local administrations to learn from other experiences, overcome self-referencing, and improve their performance enforced by reputational pressure.
Moreover, the data accountability and transparency the PES affords help public bodies garner loyalty from citizens.

Ultimately, the PES represents a real contribution to measuring standard costs on the basis of the outputs that each region is able to provide to citizens, not on how much the region spends per habitant. This seems to be a fair approach to the deployment of fiscal federalism in Italy.

However, the national PES still has some drawbacks. They can be grouped into two kinds: data availability and assessment method. The readily available and quality national data sources allow the analysis to be focused only on selected dimensions of performance. For instance, neither individual experience nor direct primary care services can be used in the evaluation. In addition, the issue of privacy did not allow recording of linkages among different types of service data that could be useful for analyzing patient pathways. Regarding assessment method, the quintile technique obliges classification of regions into the five assessment bands. This is particularly true for indicators in which performance is quite similar. This limitation can be overcome when central and regional governments share quantitative standards.

However, what we have presented is a pilot study and can be considered as a starting point for an ongoing process that aims to refine and improve the selected indicators and domains.

Finally, it’s important to consider that the possibility of working within a Beveridge system should reduce the emphasis that policy makers place on the dynamics of the health market and reimbursement procedures and could enable them to pay more attention to the quality of treatment, well-being of citizens, and equity, which are the true keystones for keeping expenditure under control. However, the variability in the results presented here and obtained by the federalist-based Italian healthcare system is enough to question the efficacy of the Beveridge system in reducing inequalities. This geographic variability also exists in the United States (Wennberg and Gittlesohn 1973; Wennberg 2004).
The conclusion is therefore that the choice of a Beveridge-style healthcare system is important but not in itself sufficient to guarantee equity of access for citizens. In reality, what really counts in both Italy and the United States is the capacity to manage variability by means of performance monitoring and assessment, which help political decision makers and health professionals to accurately focus their actions in order to improve equity and reduce unnecessary variability. In both countries it is still essential and desirable to have a system of performance measurement using federal benchmarking criteria. Whichever health system model is adopted, by means of public disclosure of results (Mannion and Goddard 2003; Fung et al. 2008; Hibbard, Stockard, and Martin 2003) and empowerment of the patient’s role performance measurement ensures that variability is reduced and service quality is improved.

Acknowledgments

This study was supported by the Italian Minister of Health.

The authors wish to thanks the Ministry of Health, Ferruccio Fazio, top managers of the Department of Health and Agenas, in particular Filippo Palumbo and Lucia Lispi, for their insightfully suggestions.

The authors also wish to thank the researchers from Laboratorio Management e Sanità for their help in the development of the pilot performance evaluation system, in particularly Domenico Cerasuolo that realized the software, Anna Bonini for her comments and Maria Sole Bramanti for her help during the elaboration.
References


Linard K., Basset M., Dvorsky L. & Yoon J., 2000, A dynamic balanced scorecard template for public sector agencies, 19th International Conference of the System Dynamics Society


Ricketts TC., Randolph R., Howard HA., Pathman D., Carey T., 2001, Hospitalization rates as indicators of access to primary care, Health Place, Vol.7 No.1, pp.27-38


Vittadini G., 2010, La valutazione della qualità nel sistema sanitario, analisi dell’efficacia ospedaliera in Lombardia, Edizioni Guerini e Associati, Milan: CRISP


Wennberg J. E., 2004, Perspective: Practice Variations and Health Care Reform: Connecting the Dots, Health Affairs


CHAPTER 4 - The performance evaluation system of prevention: the experience of the Experimental Institutes for Animal Disease Prevention (II.ZZ.SS)

Il benchmarking, oltre ad essere un importante strumento manageriale, e quindi obbligatorio, è anche utilizzato in sistemi sanitari in cui sono le strutture a volersi mettere in discussione in maniera volontaria: per la prima volta il settore della prevenzione ha cercato di documentare e rendere trasparenti le scelte e gli esiti delle attività svolte, superando il carattere dell’autoreferenzialità.

Il presente lavoro, partendo da un ampio quadro teorico, riporta l’esperienza concreta degli Istituti Zooprofilattici Sperimentali (II.ZZ.SS.) e mette in evidenza sia il superamento delle difficoltà della rilevazione dei dati in un settore così complesso come quello della prevenzione, sia il forte impatto che ha avuto la misurazione delle performance sul livello decisionale di una Istituzione pubblica (Naddeo, 2006). In particolare, mostra come l’uso sistematico di queste informazioni possa favorire e sostenere processi continui di innovazione e miglioramento organizzativo, al fine di definire e attuare politiche capaci di rispondere adeguatamente ai bisogni della collettività (Nuti et al., 2009).

L’articolo si sviluppa in cinque sezioni: la prima descrive il contesto italiano in cui opera la prevenzione; la seconda delinea la nascita, il ruolo e l’evoluzione degli II.ZZ.SS. nel tempo; la terza riporta il modello concettuale a cui si è fatto riferimento e la metodologia utilizzata nel sistema di valutazione della performance degli II.ZZ.SS.; la quarta descrive i risultati ottenuti, ed infine la quinta contiene alcune considerazioni critiche e commenti conclusivi.

Parole chiave
Servizi sanitari, valutazione delle performance, igiene degli alimenti, produzioni zootecniche
1. **Introduzione: il settore della prevenzione nel contesto Italiano**

Alla luce dei cambiamenti avvenuti nel corso degli anni specie sul versante normativo, il settore della prevenzione è stato chiamato a svolgere sempre di più un’azione di controllo sui processi che singoli soggetti pubblici o privati devono promuovere al fine di adottare misure idonee a tutelare il benessere dei cittadini, dei lavoratori e degli animali. È perciò prioritario che il servizio pubblico eserciti la funzione di vigilanza e di controllo che la legge gli attribuisce nei vari settori come: la promozione dei corretti stili di vita, la protezione e tutela della salute e della sicurezza dei cittadini negli ambienti di vita e di lavoro tra cui la sicurezza alimentare, la protezione dai rischi per la salute, anche di origine ambientale, ed infine la difesa dalle malattie infettive e diffusive umane e altre animali mediante azioni di verifica sui percorsi decisionali propri di altri soggetti.

Da come si può intuire il settore della prevenzione è molto ampio: qui ci occuperemo della sicurezza pubblica veterinaria (SPV) come “il complesso di prestazioni che il pubblico e le pubbliche amministrazioni si aspettano (consciamente o inconsciamente) dai Servizi Veterinari (SV) pubblici”. Essa è chiamata a svolgere compiti importanti che ricadono poi su altri soggetti (Regioni, ASL, II.ZZ.SS., cittadini)\(^4\) come:

- a) sanità animale e zoonosi
- b) sicurezza alimentare
- c) ambiente
- d) ricerca e formazione

\[a) \textbf{Sanità animale e zoonosi.}\] In Italia, malattie come la tubercolosi e la brucellosi non sono ancora eradicate e anzi causano ogni anno numerosi casi umani di malattia nonostante siano

soggette a piani di profilassi obbligatoria dagli anni ‘50. Questo porta a valutare le malattie degli animali in funzione delle loro conseguenze socio-economiche in quanto possono causare rilevanti perdite all’economia italiana, sia direttamente (es. animali morti, minori produzioni zootecniche) sia indirettamente (es. mancata commercializzazione per chiusura dei mercati internazionali), alla salute umana (es. zoonosi, malattie professionali) e all’ambiente (es. abbandono di attività zootecniche a causa di malattie degli animali). Un’azione più incisiva in questo settore andrebbe pertanto perseguita con ogni sforzo, così come interventi nel settore delle malattie minori (non solo di origine infettiva) che portano a determinare notevoli diseconomie agli allevatori, oppure attuare strategie di controllo basate su principi che tengano conto sia del benessere animale, sia delle tecniche di allevamento, sia dell’ambiente.

b) Sicurezza alimentare. Come in Europa, anche in Italia la sicurezza alimentare riveste una priorità assoluta. La strategia che viene applicata è denominata “dai campi alla tavola” e segue un approccio integrato basato essenzialmente su otto pilastri attraverso la capacità di garantire lo stesso livello di sicurezza sia per gli alimenti prodotti localmente che per quelli importati; lo svolgimento dei controlli ufficiali attraverso l’accreditamento obbligatorio dei laboratori; l’importanza dell’etichetta che esplicita la composizione del prodotto, il produttore e i metodi di preparazione, anche al fine di evitare alimenti che provocano allergie; la promozione e la realizzazione di eventi formativi sulla sicurezza alimentare e di attività di educazione alimentare; il coordinamento con il sistema di sorveglianza sulle malattie trasmesse con gli alimenti, promozione di indagini e studi mirati tenendo conto delle classi più a rischio (anziani, persone con AIDS/HIV, ecc.).

---

5 Dai campi alla tavola: Prodotti alimentari sicuri per i consumatori europei, Commissione Europea, 2004
c) Ambiente. Il degrado ambientale cui si assiste in tutto il mondo colpisce anche in Italia le aree ad elevata produzione zootecnica (es. fenomeno della diffusione di alghe nel Mare Adriatico) e le aree marginali (abbandono del territorio con conseguente incremento dei disastri naturali – incendi, alluvioni, ecc.). In questi contesti la partecipazione attiva dei servizi veterinari alla programmazione e gestione del territorio risulta di particolare importanza anche per contribuire al miglioramento delle condizioni di vita della popolazione che si dedica ad attività zootecniche tradizionali. Nelle zone in cui si verificano disastri industriali (ad es. contaminazione da diossine in Campania) si sono verificate gravi ripercussioni sulla zootecnia, sugli alimenti di origine animale e inquinamento delle acque. Occorre pertanto tener presente le indicazioni della recente “Strategia dell’ambiente e la salute” lanciata dalla Commissione Europea in collaborazione con l’Organizzazione Mondiale della Sanità, indicazioni che richiedono tra l’altro a) risorse e strutture per integrare e valutare, in maniera accurata ed efficiente, le basi di dati raccolti da strutture diverse (ad es. II.ZZ.SS., ASL e ARPA); b) risorse e strutture per la messa a punto di sistemi di allerta rapidi; c) strategie per la formazione degli operatori sanitari di tutte le professioni nei campi della prevenzione; d) risorse mirate alla ricerca sulle interazioni tra ambiente, filiera alimentare e salute, umana ed animale.

d) Ricerca e formazione. Andrebbe modificata la mentalità in tema di ricerca, per renderla non patrimonio esclusivo di alcuni Enti specializzati (es. ISS, II.ZZ.SS., Università, ecc.) ma aperta alle richieste ed agli apporti provenienti dal territorio. Non negando la necessità di sempre nuove scoperte scientifiche, parte della ricerca andrebbe tuttavia indirizzata a soddisfare le esigenze pratiche e valutate in base alle ricadute sulla popolazione. L’Italia ha una presenza ancora molto limitata in seno ai progetti di ricerca europei, e questa andrebbe incentivata con tutti i mezzi, come anche lo stabilirsi di reti di collaborazione fra i diversi
Istituti nazionali. Anche il sistema dei crediti formativi andrebbe radicalmente rivisto, così come la divulgazione delle informazioni dovrebbe essere uno strumento di appoggio per tutte le attività dei Servizi veterinari (ad. es. piani di profilassi, sicurezza alimentare, controllo delle popolazioni animali) ed effettuata con tecnologie adeguate a tutte le popolazioni bersaglio.

Quindi come è stato fin qui delineato, il settore della prevenzione (D.Lgs.112/98: Legge Costituzionale 3/2001) coinvolge molti soggetti (Figura 1): le ASL, con i dipartimenti di prevenzione, l’Arpa per le prestazioni di carattere ambientale, gli Istituti Zooprofilattici Sperimentali per la sicurezza alimentare e la sanità animale, e l’Istituto Superiore per la Prevenzione e la Sicurezza sul Lavoro (Ispsl) per la vigilanza e l’ispezione (Cinquini et al., 2009). Questo articolo evidenzia come il ruolo degli II.ZZ.SS. assume particolar importanza nel nostro Paese e come la loro attività è direttamente correlata ai bisogni della collettività.

Figura 1: Il sistema veterinario e di sicurezza alimentare in Italia
2. La storia degli II.ZZ.SS.

Dal punto di vista normativo le riforme sanitarie degli anni ‘90 che hanno introdotto strumenti e logiche aziendali nella gestione delle aziende sanitarie, hanno riguardato anche gli II.ZZ.SS.; ad esempio con la Riforma Sanitaria del 1993 è stata istituita la figura del Direttore Generale (D.Lgs. 502/1992: Riordino della disciplina in materia sanitaria; D.Lgs. 270/1993: Riordinamento degli II.ZZ.SS.; D.Lgs. 229/1999: Norme per la razionalizzazione del SSN) e con il D.Lgs. 229/99 è avvenuto il riconoscimento del ruolo degli II.ZZ.SS. a supporto dei Dipartimenti di Prevenzione.

Gli II.ZZ.SS. nascono come strutture tipicamente veterinarie, con un compito prioritario di supporto diagnostico alle aziende zootecniche. Solo in un secondo momento hanno assunto un ruolo importante anche nel controllo degli alimenti di origine animale e destinati all'alimentazione zootecnica che oggi rappresenta sicuramente l'attività più significativa e che assorbe la maggior parte delle risorse in termini di personale e di attrezzature (Severini, 2011).

Un elemento importante che caratterizza l'attività degli II.ZZ.SS. è la loro distribuzione su tutto il territorio nazionale che permette, quando possibile, di ottenere delle sinergie importanti in termini di eccellenze tecniche. In particolare, per quanto riguarda le eccellenze scientifiche, una funzione importante viene svolta dai centri di referenza nazionale, individuati con particolari decreti ministeriali presso i singoli istituti, e che rappresentano dei nodi in cui si concentrano elevate competenze scientifiche in specifici settori. A questi è assegnato il compito di supporto tecnico per gli altri laboratori ma anche di verifica delle performance di tutte le strutture presenti sul territorio nazionale (attraverso, ad esempio l'organizzazione e la valutazione di ring test periodici). Si aggiunge inoltre il ruolo dell’Istituto Superiore di Sanità come laboratorio di riferimento scientifico per tutti gli altri laboratori che svolgono attività di controllo ufficiale.

Figura 2: Gli Istituti Zooprofilattici Sperimentali in Italia

In particolare le attività che svolgono queste strutture sono riconducibili a:

- sorveglianza epidemiologica: attività volte a raccogliere informazioni e dati che consentono l’identificazione precoce, la diagnosi e la risposta tempestiva alla presenza delle malattie animali. I dati raccolti possono essere sia di natura qualitativa, ossia conoscenze relative alla storia naturale delle infezioni/malattie e alle nuove acquisizioni scientifiche (ecologia, trasmissione, rapporti ospiti-agenti-ambiente), sia quantitativa, ossia dati che
derivano da studi e indagini che consentono di “misurare” determinati fenomeni (ad esempio la percentuale di animali infetti identificata nell’arco di un anno)
- La qualità: è collegata al punto precedente in quanto gli II.ZZ.SS. hanno un mandato istituzionale che li caratterizza come strutture responsabili per l’esecuzione delle attività di analisi ufficiali di laboratorio (chimiche, microbiologiche, parassitologiche, virologiche, sierologiche, di ricerca di OGM, ecc.) nei settori della sanità e benessere animale, nella sicurezza igienica e sanitaria delle produzioni zootecniche e degli alimenti di origine animale, a supporto di tutti i servizi ispettivi dei loro territori di competenza (Servizi Veterinari e di Igiene degli Alimenti delle ASL, NAS, altri organi di vigilanza territoriale). Gli II.ZZ.SS. con i presidi di osservatorio epidemiologico svolgono poi in tutti questi settori compiti di raccolta dati e monitoraggio sulle condizioni zoosanitarie delle filiere e di sicurezza alimentare nei territori regionali di competenza, in stretta collaborazione con i Servizi Regionali; ciò anche ai fini della programmazione delle attività di vigilanza.

La ricerca sperimentale: dove uno degli obiettivi primari è l’analisi quantitativa per la valutazione dei rischi correlati a specifici pericoli presenti negli alimenti e nelle produzioni animali. Tale attività va considerata accanto alle ricerche relative alla messa a punto di metodiche analitiche e tecniche di gestione delle misure di protezione della salute e benessere animale e della salute pubblica. La grande quantità di dati prodotti viene inoltre avviata da organismi nazionali o comunitari, per contribuire alla valutazione dei rischi, mentre a livello locale, nei territori di competenza, tali dati possono essere utilizzati per la variazione delle attività di controllo sulla base delle caratteristiche delle strutture produttive. Tali sono gli interventi di tipo tecnico-scientifico secondo l’approccio integrato alla sicurezza alimentare richiesto dall’UE anche in recenti provvedimenti legislativi (Brizioli, 2005).
La formazione: erogata regolarmente dagli Istituti e collegata sia alla formazione continua attraverso i crediti ECM per i medici veterinari e tecnici della prevenzione (utenti esterni), sia alla gran parte del lavoro svolto dentro le aziende dove si è contribuito ad elevare l’igiene degli standard degli operatori all’interno dei laboratori di prova (utenti interni).

Gli II.ZZ.SS. hanno quindi compiti in materia di ricerca scientifica, di accertamento dello stato sanitario degli animali, garantendo ai servizi veterinari delle Regioni e delle Asl le prestazioni e la collaborazione tecnico-scientifica necessarie all’espletamento delle funzioni di sanità pubblica veterinaria. Inoltre producono, con l’autorizzazione del Ministero della Salute, vaccini ed ogni altro prodotto necessario per la profilassi delle malattie trasmissibili degli animali, quali: peste suina africana, brucellosi animali, dell'echinococcosi, del cosiddetto morbo della mucca pazza, della lingua blu, ecc.

Per comprendere meglio il loro ruolo all’interno del SSN sono state riportate alcune caratteristiche che evidenziano le differenze sia geografiche sia organizzative. Come mostra la tabella seguente (Figura 3) le dimensioni degli II.ZZ.SS. sono molto diverse: l’I.Z.S. Puglia-Basilicata è quasi un terzo dell’Istituto più grande, l’I.Z.S. Lombardia-Emilia Romagna, sia in termini di valore della produzione sia in termini di personale complessivo. Si evidenzia inoltre la variabilità del volume di esami dovuto a differenti esigenze di territorio e alla tipologia di esami effettuati.
<table>
<thead>
<tr>
<th>II.ZZ.SS.</th>
<th>Numero esami complessivi effettuati dagli Istituti</th>
<th>Ricavi da attività commerciale</th>
<th>Valore della produzione totale</th>
<th>Totale dipendenti</th>
<th>Borse di Studio e Co.co.co.</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.Z.S. Lombardia Emilia Romagna</td>
<td>6,173,778</td>
<td>€ 8,734,443</td>
<td>€ 68,450,103</td>
<td>653</td>
<td>87</td>
</tr>
<tr>
<td>I.Z.S. Lazio Toscana</td>
<td>1,960,629</td>
<td>€ 757,108</td>
<td>€ 37,479,606</td>
<td>398</td>
<td>74</td>
</tr>
<tr>
<td>I.Z.S. Puglia Basilicata</td>
<td>1,489,938</td>
<td>€ 85,620</td>
<td>€ 21,892,508</td>
<td>162</td>
<td>51</td>
</tr>
<tr>
<td>I.Z.S. Piemonte Liguria e Valle d'Aosta</td>
<td>1,481,445</td>
<td>€ 232,367</td>
<td>€ 28,303,293</td>
<td>346</td>
<td>52</td>
</tr>
<tr>
<td>I.Z.S. Sicilia</td>
<td>4,073,951</td>
<td>€ 200,767</td>
<td>€ 28,495,562</td>
<td>291</td>
<td>78</td>
</tr>
<tr>
<td>I.Z.S. Umbria Marche</td>
<td>780,301</td>
<td>€ 1,124,538</td>
<td>€ 23,885,791</td>
<td>234</td>
<td>29</td>
</tr>
<tr>
<td>I.Z.S. delle Venezie</td>
<td>2,120,161</td>
<td>€ 4,117,569</td>
<td>€ 45,657,344</td>
<td>480</td>
<td>94</td>
</tr>
<tr>
<td>I.Z.S. Sardegna</td>
<td>693,395</td>
<td>€ 321,725</td>
<td>N.D.</td>
<td>288</td>
<td>52</td>
</tr>
<tr>
<td>I.Z.S. del Mezzogiorno</td>
<td>3,149,278</td>
<td>€ 491,961</td>
<td>€ 27,039,140</td>
<td>291</td>
<td>26</td>
</tr>
</tbody>
</table>


Figura 3: Dati di contesto degli II.ZZ.SS.

3. Il modello “toscano” verso gli II.ZZ.SS.


---

\(^6\) In particolare nell’aprile 2009 è stato rilasciato il brevetto n° 0001358839 “Sistema di Valutazione della Performance di Aziende Sanitarie” avente ad oggetto il sistema di indicatori adottato in Toscana e successivamente nelle altre regioni di cui sopra, mentre nel gennaio 2008 è stato depositato il brevetto n° PI 2008A0000008 “Metodo Per La Gestione Di Indicatori Multidimensionali Di Performance Di Aziende Di Servizi” avente ad oggetto l’architettura hardware e software per gestire gli indicatori.
- **I soggetti coinvolti**
Con il supporto del Laboratorio Management e Sanità (MeS) della Scuola Superiore Sant’Anna di Pisa, una prima valutazione della performance degli Istituti, è stata effettuata nel 2009 con due II.ZZ.SS., Lazio-Toscana e Umbria Marche, su richiesta espressa dei rispettivi Direttori Generali (Leto et al., 2009).


- **La metodologia**
Nella fase di progettazione il processo di adattamento del sistema di valutazione della performance delle aziende sanitarie agli II.ZZ.SS. ha riguardato diversi step:
  a) L’individuazione delle dimensioni da analizzare;
  b) La selezione degli indicatori ed individuazione degli standard di riferimento per la loro valutazione;
  c) La raccolta ed elaborazione di dati provenienti da fonti aziendali o da indagini ad hoc
  d) La pubblicazione dei risultati

---

7 L’unico Istituto che non ha partecipato al progetto è l’II.ZZ.SS. di Teramo
Per quanto riguarda il punto a) nella fase di sperimentazione sono state individuate 5 dimensioni di analisi Mission istituzionale, del territorio e della ricerca (MIT); apprendimento e sviluppo organizzativo (ASO); Efficienza e sostenibilità economica (ECO); Utente (U) e Salute della Popolazione Animale (SPA) dopo la fase di revisione avvenuta attraverso la partecipazione ad alcuni focus group. Pur condividendo la necessità di confronto sugli outcome, con l’ingresso degli altri sette Istituti, non è stata analizzata l’ultima dimensione in quanto richiedeva di un maggior coinvolgimento di altri soggetti appartenenti al sistema sanitario interessati alla salute umana e animale.

In relazione al punto b), nella sperimentazione per individuare gli indicatori, si sono svolti focus group tematici prima separatamente e poi in modo plenario attraverso il confronto dei professionisti e la condivisione delle direzioni. Come si evince gli indicatori sono emersi dal “basso”, cioè dai dipendenti appartenenti a diverse categorie professionali che hanno preso spunto da informazioni esistenti nel proprio sistema di budget o nei flussi ministeriali (es. rilevazione della soddisfazione dei corsi formativi ECM), da indicatori presenti nel sistema di valutazione della performance delle aziende sanitarie toscane oppure da indicatori proposti nello studio Documento di Valutazione sui determinanti di Salute e sulle Strategie del servizio sanitario regionale (DVSS) effettuato dalla Regione Umbria nel 2004 (Romagnoli, 2006), individuando così gli indicatori da inserire nel sistema di valutazione degli II.ZZ.SS. (Nuti et al., 2012).

Per quanto riguarda l’ultimo punto, vorrei evidenziare le difficoltà relative alla raccolta e all’elaborazione dei dati: ciò ha comportato sia la costituzione di gruppi di lavoro specifici sia la verifica della correttezza dei dati aziendali. Nella prima parte quindi è stato condiviso il fatto di approfondire alcune tematiche in ottica di miglioramento continuo che hanno riguardato:

- La revisione degli indicatori della formazione alla luce delle recenti modifiche che riguardano gli eventi ECM. A questo gruppo di lavoro, coordinato dal referente dell’I.Z.S. Lazio-Toscana, hanno partecipato tutti i responsabili della formazione degli II.ZZ.SS. coinvolti e un referente del Dipartimento per la sanità pubblica veterinaria, la nutrizione e la sicurezza degli alimenti del Ministero della Salute.

- La scelta dei prodotti-beni da utilizzare per confrontare gli indicatori di efficienza della gestione degli approvvigionamenti e del magazzino. A questo gruppo di lavoro, coordinato dal responsabile degli acquisti dell’I.Z.S. Lazio-Toscana, hanno partecipato tutti i responsabili degli uffici acquisti degli II.ZZ.SS.

- La riclassificazione ed il confronto dei bilanci degli I.Z.S. è stato coordinato dal Laboratorio MeS con la partecipazione dei direttori amministrativi e/o i responsabili di bilancio di tutti gli Istituti.

Nella seconda parte, vista la criticità emersa per la raccolta dati, sono state attivate importanti visite presso gli II.ZZ.SS. La difficoltà del reperimento dati consisteva nel comprendere cosa era stato “contato”: nonostante le schede di calcolo degli indicatori, l’estrazione dei dati inevitabilmente comporta delle scelte di interrogazione dei database non contemplate nelle riunioni di condivisione. Per limitare questi fenomeni dovuti alle modalità di estrazione dei dati ed interpretazione dei termini (il glossario creato per condividere la terminologia tra gli Istituti infatti non è esaustivo e non comprende tutte le fattispecie che nella realtà si presentano agli operatori) ciascun I.Z.S. ha avuto la possibilità di analizzare la struttura del
proprio sistema informativo. Inizialmente sono state individuate le relazioni generali fra i diversi dataset (se in outsourcing, se collegate) e successivamente, in modo specifico, il flusso delle informazioni relativo alla verifica del calcolo della tempestività di risposta di alcuni esami più critici. Man mano che si verificava il modo in cui potevano essere calcolati tali indicatori, si delineava il livello di informatizzazione e formalizzazione del flusso dati per la costruzione degli altri indicatori in carico all’istituto (Nuti et al., 2012).

Alla visita, effettuata prima del calcolo degli indicatori, hanno partecipato un ricercatore del Laboratorio MeS, i referenti degli II.ZZ.SS. Umbria-Marche e Lazio-Toscana che l’anno precedente avevano incontrato ed affrontato le problematiche di estrazione dei dati, due/tre referenti di altri II.ZZ.SS. ed un referente del Ministero della Salute. Oltre alla collaborazione e solerzia dei referenti degli II.ZZ.SS., le visite in loco sono state indispensabili per la prosecuzione del lavoro permettendo di superare le barriere della prevenzione costituite da flussi dati non consolidati o strutturati. E’ stato un ulteriore punto di forza che ha consentito di comprendere quali tipologie di informazioni erano disponibili e utili per il calcolo degli indicatori e quali rappresentavano delle problematiche di rilevazione dati.

4. I risultati
Sono stati identificati 80 indicatori di cui 21 di valutazione e 59 di osservazione. Questi ultimi servono a contextualizzare un fenomeno o ad integrare informazioni non contenute negli indicatori di valutazione. Un esempio è fornito dall’indicatore “giorni medi di ritardo per esami relativi ai piani istituzionali” che integra l’indicatore di valutazione “giorni mediani di ritardo per esami relativi a piani istituzionali”: la relazione fra i giorni di ritardo medi e mediani evidenzia in modo immediato la presenza di outlier ossia di fenomeni di ritardo sporadici ma rilevanti. Nella rappresentazione di sintesi del sistema di valutazione della performance degli II.ZZ.SS. sono presenti solo 14 “pallini” perché alcuni indicatori sono stati raggruppati in un unico indicatore sulla base della comunanza dell’attività o del fenomeno

Figura 5 - L’indicatore U6 qualità percepita degli eventi formativi ECM

Nella figura 6 si riporta l’elenco degli indicatori e sottoindicatori di valutazione e di descrizione (riportati in corsivo) per ciascuna dimensione e sottodimensione.
### SISTEMA DI VALUTAZIONE DELLA PERFORMANCE DEGLI ILZZ.SS.

**Elenco indicatori:**

<table>
<thead>
<tr>
<th>CODICE</th>
<th>INDICATORE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MIT</strong> - Prospettiva Mission istituzionale, Territorio e Ricerca</td>
<td></td>
</tr>
<tr>
<td><strong>Mission Istituzionale ed esigenze del territorio</strong></td>
<td></td>
</tr>
<tr>
<td><strong>MIT1</strong></td>
<td>Rispetto dei tempi</td>
</tr>
<tr>
<td><strong>MIT1.1.1</strong></td>
<td>Percentuale di esami che rispettano i tempi per piani istituzionali-PNR</td>
</tr>
<tr>
<td><strong>MIT1.1.2</strong></td>
<td>Giorni MEDI di ritardo per esami relativi a piani istituzionali-PNR</td>
</tr>
<tr>
<td><strong>MIT1.1.2m</strong></td>
<td>Giorni MEDIANI di ritardo per esami relativi a piani istituzionali-PNR</td>
</tr>
<tr>
<td><strong>MIT1.1.3</strong></td>
<td>Percentuale di esami che rispettano i tempi per piani istituzionali di profilassi</td>
</tr>
<tr>
<td><strong>MIT1.1.4</strong></td>
<td>Giorni MEDI di ritardo per esami relativi a piani istituzionali di profilassi</td>
</tr>
<tr>
<td><strong>MIT1.1.4m</strong></td>
<td>Giorni MEDIANI di ritardo per esami relativi a piani istituzionali di profilassi</td>
</tr>
<tr>
<td><strong>MIT1.2.3</strong></td>
<td>Tempo medio di risposta degli esami (per 11 tipologie di esami)</td>
</tr>
<tr>
<td><strong>MIT1.2.3a</strong></td>
<td>Tempo medio esecuzione degli esami (per 11 tipologie di esami)</td>
</tr>
<tr>
<td><strong>MIT2</strong></td>
<td>Qualità di processo</td>
</tr>
<tr>
<td><strong>MIT2.1</strong></td>
<td>Percentuale di prove accreditate</td>
</tr>
<tr>
<td><strong>MIT2.2</strong></td>
<td>Percentuale di prove con procedure standardizzate</td>
</tr>
<tr>
<td><strong>MIT3_2</strong></td>
<td>Percentuale di esiti favorevoli relativamente ai ring test</td>
</tr>
<tr>
<td><strong>MIT5_3</strong></td>
<td>Percentuale di piani istituzionali pubblicati</td>
</tr>
</tbody>
</table>

**Formazione**

<p>| <strong>MIT7.1.1</strong> | Crediti medi per evento ECM |
| <strong>MIT10</strong> | Programmazione dell'offerta formativa |
| <strong>MIT10.1</strong> | Offerta formativa programmata |
| <strong>MIT10.2</strong> | Offerta formativa extra programma |</p>
<table>
<thead>
<tr>
<th>CODICE</th>
<th>INDICATORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIT16</td>
<td>Capacità di attrazione delle risorse attraverso l'attività di ricerca</td>
</tr>
<tr>
<td>MIT16.1</td>
<td>Percentuale di fondi di ricerca nazionali assegnati all’I.Z.S. per la ricerca corrente</td>
</tr>
<tr>
<td>MIT16.3</td>
<td>Percentuale di fondi di ricerca nazionali assegnati all’I.Z.S. per la ricerca finalizzata</td>
</tr>
<tr>
<td>MIT16.4</td>
<td>Percentuale di altri fondi di ricerca assegnati all’I.Z.S.</td>
</tr>
<tr>
<td>MIT17</td>
<td>Produzione scientifica</td>
</tr>
<tr>
<td>MIT17.1.1</td>
<td>Attività scientifica posizionamento rispetto ad altri istituti</td>
</tr>
<tr>
<td>MIT17.1.2</td>
<td>IF medio per dirigente sanitario, medico e veterinario</td>
</tr>
<tr>
<td>MIT17.2</td>
<td>Presentazioni medie per dirigente sanitario, medico e veterinario in convegni internazionali</td>
</tr>
<tr>
<td>MIT19</td>
<td>Partecipazione a gruppi internazionali per dirigente sanitario, medico e veterinario</td>
</tr>
<tr>
<td>MIT21</td>
<td>Rispetto dei tempi dei progetti di ricerca</td>
</tr>
<tr>
<td>ECO3_1</td>
<td>Rotazione del magazzino</td>
</tr>
<tr>
<td>ECO4_1</td>
<td>Tempo medio di evasione della richiesta d’acquisto di beni consumabili (sottoposti a contratto)</td>
</tr>
<tr>
<td>ECO6</td>
<td>Efficienza degli approvvigionamenti (9 beni)</td>
</tr>
<tr>
<td>ECO11</td>
<td>Tempo di evasione delle fatture</td>
</tr>
<tr>
<td>ECO12</td>
<td>Equilibrio Economico-Reddituale</td>
</tr>
<tr>
<td>ECO12.1</td>
<td>Equilibrio Economico Generale</td>
</tr>
<tr>
<td>ECO12.2</td>
<td>Equilibrio Economico Gestione Caratteristica</td>
</tr>
<tr>
<td>ECO12.3</td>
<td>ROI</td>
</tr>
<tr>
<td>CODICE</td>
<td>INDICATORE</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>ECO13</td>
<td>Equilibrio Patrimoniale – Finanziario</td>
</tr>
<tr>
<td>ECO13.1</td>
<td>Indice di disponibilità</td>
</tr>
<tr>
<td>ECO13.3</td>
<td>Indice di elasticità finanziaria</td>
</tr>
<tr>
<td>ECO - Prospettiva dell'efficienza e della sostenibilità economica</td>
<td></td>
</tr>
<tr>
<td>ECO13.4</td>
<td>Costi di finanziamento</td>
</tr>
<tr>
<td>ECO13.4.2</td>
<td>Dilazione dei debiti</td>
</tr>
<tr>
<td>ECO14</td>
<td>Produttività dei dipendenti</td>
</tr>
<tr>
<td>ASO - Prospettiva dell'apprendimento e dello sviluppo organizzativo</td>
<td></td>
</tr>
<tr>
<td>ASO2</td>
<td>Formazione del personale</td>
</tr>
<tr>
<td>ASO2.1.1</td>
<td>Ottenimento crediti formativi interni ed esterni</td>
</tr>
<tr>
<td>ASO2.1.2</td>
<td>Ottenimento crediti formativi interni</td>
</tr>
<tr>
<td>ASO2.2</td>
<td>Partecipazione del personale ad eventi non ECM</td>
</tr>
<tr>
<td>Formazione</td>
<td></td>
</tr>
<tr>
<td>Indicatori di struttura</td>
<td></td>
</tr>
<tr>
<td>ASO4</td>
<td>Percentuale di assenza</td>
</tr>
<tr>
<td>ASO5</td>
<td>Tasso infortuni</td>
</tr>
<tr>
<td>U - Prospettiva dell'utente</td>
<td></td>
</tr>
<tr>
<td>U6</td>
<td>Qualità percepita degli eventi formativi ECM</td>
</tr>
<tr>
<td>U6.1</td>
<td>Rilevanza degli argomenti</td>
</tr>
<tr>
<td>U6.2</td>
<td>Qualità educativa</td>
</tr>
<tr>
<td>U6.3</td>
<td>Efficacia dell'evento</td>
</tr>
</tbody>
</table>

Figura 6 - Elenco degli indicatori
Come nel sistema progettato dalla Regione Toscana, la rappresentazione grafica di sintesi è espressa attraverso un bersaglio a cinque fasce concentriche (Nuti, 2005; Nuti, 2008) e permette di capire subito sia i punti di forza che di debolezza di ciascuna azienda fornendo una fotografia sintetica dell’andamento aziendale ed al contempo informazioni dettagliate per ciascuna delle dimensioni di analisi.⁸

Il modello è semplice e complesso allo stesso tempo. Semplice, perché la metafora del “bersaglio” (Figura 7) che è stata utilizzata è di immediata chiarezza, ma anche complesso, perché, come in un gioco a scatole cinesi, partendo da un dato di sintesi, permette di analizzare con passaggi di sempre ulteriore dettaglio i dati di origine e le loro determinanti. Ha alcune caratteristiche fondamentali: è trasparente e condiviso, è capace di monitorare non solo i risultati economico finanziari, ma anche le modalità con cui queste strutture si organizzano e ottengono risposte efficaci e adeguate ai bisogni della collettività.

⁸ La scelta di associare ad ognuna delle cinque fasce di valutazione della performance una scala di colori ed un range numerico deriva dal sistema toscano: si va da una performance molto scarsa con colore rosso ed un range di valori da 0 ad 1 ad una performance ottima con colore verde scuro ed un range di valori da 4 a 5 (vedi S. Nuti, Il sistema di valutazione della performance delle aziende sanitarie toscane, ETS, 2009)
Figura 7: La rappresentazione di sintesi: il bersaglio

Dalla fotografia del percorso intrapreso è emerso che gli II.ZZ.SS. hanno delle peculiarità e dei punti critici da rafforzare. Dal punto di vista economico gli Istituti sono economicamente sani con bilanci in pareggio, anche se ci sono spazi di miglioramento per quanto riguarda sia l’indicatore ECO3.1 sulla gestione del magazzino che registra una rotazione di 4 giorni dell’I.Z.S. Piemonte-Liguria-Valle d’Aosta e di 25 giorni dell’I.Z.S. Venezie, sia l’indicatore ECO4.1 sulla tempestività di approvvigionamento. In particolare per quest’ultimo, per alcuni beni selezionati, alcuni II.ZZ.SS. riescono ad approvvigionarsi più velocemente a costi inferiori (esempio il costo medio dell’alcool varia da 10 a 38 euro in tempi che vanno da 1 a 70 giorni).
In relazione alla qualità, gli Istituti presentano buoni livelli di performance come testimoniamo i dati sulla tempestività dei Piani di Profilassi (esempio: leucosi e brucellosi) con percentuali superiori al 70%. Ampi margini di miglioramento invece si evidenziano per gli esami che rispettano i tempi per i Piani Istituzionali – PNR (indicatore MIT1.1.1): per l’I.Z.S. Puglia-Basilicata solo il 36% degli esami riesce ad essere nei tempi contro una media del 70%.

La dimensione della ricerca, che comprende sia la capacità di attrazione delle risorse sia la produzione scientifica, è uno dei settori più critici in quanto presenta un’elevata variabilità. Rispetto al primo punto i dati della ricerca corrente (indicatore MIT16.1) vanno dal 5%
dell’I.Z.S. Sardegna al 18% dell’I.Z.S. Venezie; per il secondo, l’Impact Factor medio dei dirigenti (sanitari, medici e veterinari) capaci di svolgere attività scientifica di qualità varia dal 2,38 a 9,57 contro una media superiore a 4,5.

Una delle attività principali è la formazione erogate a utenti esterni: si denota che tutti gli IZS riescono ad erogare buoni livelli di formazione con un numero di crediti riconosciuti in media di 9.

Tutti gli utenti hanno un’ottima valutazione dei corsi con una percentuale superiore all’80%.

In relazione all’impegno che gli Istituti sostengono per lo sviluppo delle competenze del proprio personale tramite interventi formativi ci sono ampi margini di miglioramento: l’ottenimento di crediti ECM che i dipendenti maturano all’interno dell’azienda registrano
percentuali che vanno dal 10% dell’I.Z.S. della Sardegna al 85% dell’I.Z.S. Umbria-Marche. Anche la percentuale di assenza denota un clima diverso tra gli II.ZZ.SS. con valori dal 4% al 8%.

5. Conclusioni
La scommessa della prevenzione è una sfida che riguarda da vicino il presente sulla salute umana e animale. Prevenire vuol dire anticipare i problemi attraverso l’aggiornamento, la ricerca, la costruzione di modelli predittivi, l’avvio di campagne di sensibilizzazione; significa fornire un servizio pubblico a tutela del cittadino, equamente orientato alla responsabilizzazione individuale (acquisizione di conoscenze, cambiamento di atteggiamenti e comportamenti) e istituzionale (costruzione di reti di prevenzione, stanziamento di risorse, programmazione, definizione di strategie, ecc.). Altrimenti succede che ci accorgiamo della prevenzione solo per difetto, nel bel mezzo di un’emergenza, quando il rischio si è già trasformato in evento negativo (Andrighetto, 2010). La conoscenza dei fenomeni e l’epidemiologia sono due elementi fondamentali che servono per adeguare la prevenzione alle nuove dinamiche, fornendo le informazioni necessarie per rilevare le criticità in modo...
sistematico, perfezionare gli strumenti e andare in una direzione che ci permette di aumentare il livello di previsione di eventi negativi.


Nella costruzione del sistema di valutazione della performance degli II.ZZ.SS., la fase più complessa è stata quella dell’omogeneizzazione dei dati. La mancanza di un flusso unico come avviene per la dimensione ospedaliera con le SDO, ha reso la raccolta dei dati e delle informazioni la fase più lunga del processo di misurazione della performance degli Istituti. Definire indicatori condivisi tali da rendere misurabili le performance ha fortemente impattato con la consuetudine di una pubblica amministrazione autoreferenziale e scarsamente incline a favorire una lettura esterna ed interna del proprio essere azienda.

Pubblicare i risultati, far emergere le diversità territoriali, geografiche e produttive, discutere le determinanti dei risultati ai vari livelli del sistema ha comportato l’introduzione di nuove azioni volte al miglioramento continuo: infatti gli istituti hanno attivato percorsi formativi; avviato processi di revisione e riorganizzazione delle attività in un’ottica di total quality management in modo mirato; usato gli indicatori per l’attribuire gli obiettivi annuali alle singole unità operative; svolto la mappatura dei processi e delle fasi produttive volti a individuare le criticità, rispetto alle quali attivare soluzioni favorevoli al cambiamento.
Il sistema ha rappresentato quindi un salto culturale soprattutto in Enti poco orientati all’attivazione di processi di miglioramento, e ha permesso di rafforzare l’intenzione della pubblica amministrazione a coinvolgere il cittadino/utente in merito ai risultati raggiunti e all’attività svolta dagli Istituti. La determinazione nel voler superare le resistenze ed il coraggio di mostrare, anche pubblicamente, le criticità del proprio istituto è stato un atto di responsabilità in primis nei confronti dei clienti istituzionali, ministero, regioni ed asl ma anche nei confronti del personale che opera all’interno dell’istituto.
References

Angrighetto I., 2010, Prevenzione e ricerca, la scelta obbligata

Battelli G., 2006, Bozza di documento programmatico sulla sanità pubblica veterinaria, (www.progettarelaveterinaria.it)

Brizioli N.R., 2005, Gli Istituti Zooprofilattici Sperimentali nel campo della sicurezza alimentare, Workshop Residui negli alimenti di origine animale: analisi del rischio, prevenzione e sorveglianza, ISS, Roma

Barretta A., Vagnoni E., 2005, Il benchmarking in ambito sanitario, Cedam

Commissione Europea, 2004, Dai campi alla tavola: Prodotti alimentari sicuri per i consumatori europei, pp.1-22


D.Lgs. 30.06.1993 n. 270: Riordinamento degli Istituti Zooprofilattici Sperimentali a norma dell’articolo 1, comma 1, lettera h), della legge 23 ottobre 1992 n. 421

D.Lgs. 31.03.1998, n. 112: Conferimento di funzioni amministrative alle regioni e agli enti locali

D.Lgs. 19.06.1999, n. 229: Norme per la razionalizzazione del SSN a norma dell’articolo 1 della legge 30 novembre 1998, n. 419


Legge costituzionale 18 ottobre 2001, n. 3 “Modifiche al titolo V della parte seconda della Costituzione” pubblicata in Gazzetta Ufficiale n. 248 del 24 ottobre 2001

Legge 5 maggio 2009, n. 42 “Delega al Governo in materia di federalismo fiscale, in attuazione dell’articolo 119 della Costituzione” pubblicata in Gazzetta Ufficiale n. 103 del 6 maggio 2009

Settore igiene pubblica

Naddeo A, 2006, Misurare per decidere: la misurazione delle performance per migliorare le politiche pubbliche e i servizi, Rubbettino Editore Srl


Nuti S. (a cura di), 2008, La valutazione della performance in Sanità, Il Mulino, Bologna

Nuti S., Vainieri M., 2009, Fiducia dei cittadini e valutazione della performance nella sanità italiana, Edizioni ETS, Pisa


Romagnoli C., 2006, La salute in Umbria: stato di salute e principali determinanti, Programmazione Socio-sanitaria dell’assistenza di base ed ospedaliera, Osservatorio Epidemiologico Regione Umbria, pp.1-306

Severini S., Il sistema dei laboratori di prevenzione a tutela della salute pubblica, spvet, n. 64, 2011.
Vainieri M., Calabrese C., Campanale C., Panero C., Nuti S., 2009, Strumenti di governance del sistema sanitario toscano: la valutazione della performance degli Estav, Mecosan, Vol.69
Final Conclusions

The benchmarking, most often used for the comparison of indicators, is a recent concept in the healthcare system. Benchmarking is a powerful tool because it overcomes “paradigm blindness” that can be summed up as the mode of thinking, “the way we do it is the best because this is the way we have always done it”. Benchmarking opens organizations to new methods, ideas and tools to improve their effectiveness. It helps crack through resistance to change by demonstrating other methods of solving problems than the one currently employed, and demonstrating that they work, because they are being used by others.

The objectives of this work were to define the concept of benchmarking and its evolution in the healthcare sector, to provide some insights on the relationships among compulsory and voluntary models and to analyze the application of its at national level.

In conclusion the literature review highlighted how benchmarking approaches have evolved in the healthcare sector. This evolution produced numerous definitions, whose common theme is continuous measurement of one’s own performance and comparison with best-performers to learn about the latest work methods and practices in other organizations.

The Paper 1 suggests that a voluntary system is the best because it contribute to generate benefits from both the organizational and healthcare points of view. In fact the first article “Which are positive and negative effects of voluntary and compulsory benchmarking models? A review of literature in the healthcare sector” focused on a review of published benchmarking articles in the healthcare on the basis of Bowerman’s (2002) classification of voluntary and compulsory models in the healthcare system. The review was obtained through a systematic search about benchmarking in healthcare including access to only published examples of benchmarking approaches and models used, considering some web databases.

This supported identification of how benchmarking approaches have developed and been used, remaining true to the basic benchmarking principles of continuous improvement.
through comparison and sharing (Camp, 1989; Watson, 1993; Bhutta and Huq, 1999; Dossi and Patelli, 2008, 2010).

Comparing data (as written in the second paper “Assessment and improvement of the Italian Healthcare system: first evidences from a pilot national performance evaluation system” ), within or between healthcare systems at national level raises the question of how such comparisons affect performance improvement. In fact the compulsory system of Italian National Healthcare System, by using first evidences of a pilot Performance Evaluation System (PES), was showed that it can be viewed as a governance mechanism supporting the national health system (Nuti et al, 2011). In particular, this type of compulsory comparison was necessary to ensure uniform levels of care for the population and assisting regional managers to evaluate performance in benchmarking.

It has shown that in some cases it is necessary that this is directed from above to help regions in difficulty and the role of the NHS to ensure the same “Essential Health Benefits Package” (LEA –Livelli Essenziali di Assistenza).

Indeed, the voluntary system has highlighted that if it’s accepted, it will be immediately inserted in the current practice (paper 3). Only if the benchmarking process is incorporated into existing policies, it will have positive effects on the healthcare system. About this aspect the last paper “The performance evaluation system in the prevention sector: the results of the Experimental Institutes for Animal Disease Prevention (II.ZZ.SS.)” was concerned the use of benchmarking as a management tool in order to manage performance. This voluntary models focused on the role played by PES (Nuti, 2008; Nuti et al., 2012) in the resources reallocation process within the Italian services Public Health, in particular veterinary medicine area (surveillance of animal stock health, hygiene of food production and animal).
Finally both cases described in the second and third paper have been discontinued.

On one hand, even though the role of the central government is limited because regions have substantial autonomy and are responsible for the planning, financing, monitoring and control of health care within their territories, is necessary to have strong political endorsement by Ministry of Health on how to change the current system within the national territory in which it operates. In the National Healthcare System this condition is lacked because it was appointed the new Ministry of Health, producing “weak central direction” for continue national exchange of views.

On the other hand, II.ZZ.SS. Institutes took a “cultural leap” especially for public bodies not oriented to change processes improvement. It helped to reinforce the intention of the government to involve the citizens on the results achieved and the activities carried out by the Institutes. This process was not been sufficiently strong for II.ZZ.SS. in order to accept the results of the comparison.

Even if there often is a wide gap between the public’s support for a set of principles concerning what needs to be done about the overall problems facing the nation’s health care system and their support for specific policies designed to achieve those goals, I conclude that is very significant to have both powerful of government’s principal advisor on health supporting the system, and not be frightened to “get in the game” in order to learn from other experiences and to improve quality health services to citizens.
Acknowledgements

During these years of my PhD, in several moments, I have been about to quit, but thanks of the contributions and participation of several people, I have been able to finalize my research and present my PhD.

For this, I would like to acknowledge and express my appreciation for my tutor, Prof. Sabina Nuti, because she has always encouraged me to pursue my research interests and for her guidance in carrying out my research. I feel privileged to have been her student.

I am deeply grateful to my advisor, Dr. Milena Vainieri, for her useful and precious suggestions.

I would like to thank all the people who contributed (ex. managers, researchers, practitioners, veterinarians) to this research for giving me a part of their knowledge and passion both at work and in life.

I appreciate all my colleagues of the Laboratorio Management e Sanità (my second family) for sharing this experience, listening, providing sound advice. Their kindness will never be forgotten.

I would also like to thank all my friends, in particular Linda & Martina, for their patience and with affection to stand my ups and downs.

Moreover my sincere thanks go to my family, my pillar of strength: Marta, Guido (who looks at me from up there), Simone, Paola and Francesco!!!