

UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

PERFORMANCE MEASUREMENT INDICATORS
IN THE HEALTHCARE INDUSTRY: A SYSTEMATIC REVIEW

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF HUMAN RELATIONS

By

NIDA'A ABU JBARA

Norman, Oklahoma 2016

PERFORMANCE MEASUREMENT INDICATORS
IN THE HEALTHCARE INDUSTRY: A SYSTEMATIC REVIEW

A THESIS APPROVED FOR THE
MASTER OF HUMAN RELATIONS

By

Dr. Jody Worley, Chair

Dr. Chan Hellman

Dr. Brenda Lloyd-Jones

Acknowledgements

I would like to express my sincere gratitude to my advisor Dr. Jody Worley, the completion of this study could not have been possible without his expertise and guidance. His continuous support and motivation helped me in all the time of working on this research. Also, I would like to extend my thanks to Drs. Lloyd-Jones and Hellman for sitting on the panel of my thesis committee and taking the time to read my thesis and provide positive recommendations for the future development of this research.

Last, a special thanks goes to my mother for supporting me spiritually throughout my study time and life in general. Without her prayers and encouragement I would not have completed this milestone in my life.

Table of Contents

Abstract	viii
Introduction	1
Performance Evaluation in Healthcare	5
Statement of the Problem	7
Purpose of the Study	10
Common Performance Measurement Systems	10
Pay for Performance (P\$P)	10
360-Degree Feedback	11
Balanced Scorecard	12
Management by Objectives	13
Research Questions	14
Method	14
Search Strategy	14
Study Selection	16
Results	17
Search Results	17
Data Extraction	18
Characteristics of Studies	19
Quality of Studies	21
Data Analysis	21
Key Findings	22
Problems and gaps of measuring performance of medical practitioners ...	22
Performance measurement systems	26
Pay-for-Performance (P4P)	27
360-Degrees / Multisource Feedback (MSF)	32
Balanced Scorecard	36
Evaluations and developments	41
Evaluations of appraisal systems	42
Developments of appraisal methods	47
Limitations	52
Discussion	53

Conclusion	57
References	59
Appendices	64
Appendix (1)	64

List of Tables

Table (1) Inclusion and Exclusion Criteria	16
Table (2) Summary of the Characteristics	19
Table (3) Classification of Included Studies	20
Table (4) Strengths and Weaknesses of Performance Measuring Tools	39
Table (5) Example of the Behavioral Competencies	45

List of Figures

Figure (1) PRISMA Flow Diagram	18
Figure (2) Diagram of problems in clinical nurse performance appraisal ...	24
Figure (3) Healthcare Applications of Balanced Scorecard	37
Figure (4) Selected additional Balanced scorecard perspectives used by healthcare	38
Figure (5) Employee evaluation criteria	50
Figure (6) Indicator Example	52
Figure (7) Soft Skills Definition	55

Abstract

Objectives: The purpose of this study is to explore and investigate the types of employee performance measurement systems applied in healthcare services and assess their efficiency in providing accurate measurement of the performance of healthcare workers across the various job types with focus on performance indicators to measure soft skills.

Study Design and Methodology: A systematic search in discipline specific databases included PsychInfo, Medline, ABI/Inform, and Business Source Elite; and in multidisciplinary databases included Academic Search Elite, Health and Psychosocial Instruments, ProQuest, and Science Direct. The used search terms were “employee performance indicators”, “employee performance appraisal”, and “healthcare”. The search was limited to publications in English language without any restrictions on year of publication. The search was supplemented with an independent manual search of references of relevant studies and bibliographies of review articles.

Results: A total of 23 articles met the inclusion criteria. Articles were classified into 3 categories: Performance measurement systems and programs comprise of 10 studies, evaluation and development of measuring tools includes 7 studies, and problems and gaps of measuring performance of medical practitioners included 6 studies. Majority of the studies (65.22%), 15 studies, are healthcare related.

Conclusions: There was wide diversity in the applicable methods. However, there appears to be no comprehensive approach to performance evaluation in healthcare. The systematic review of published literature does not provide or identify a clear

solution to the weakness of applicable employee performance appraisal systems in healthcare organizations. There is a literature gap in covering performance measurement systems for non-medical employees working in healthcare organizations and in discussing performance indicators for soft skills. The findings have supported the statement of problem and provided direction for future research that is needed to address the existing gap in performance measurement literature and contribute to evidence-informed decision-making in healthcare performance management.

Introduction

Assessing employees' performance is a process that is commonly practiced in most organizations all around the world. Employees' performance assessment relies on using a strategic tool known as "performance appraisal". The increasing importance of performance appraisal system is reflected in a recent survey about human resource professionals' perception about performance management effectiveness conducted by the Society for Human Resource Management (SHRM) in (2014), included 391 human resource professionals from a randomly selected sample of SHRM's membership, reported that 72% of organizations conduct performance appraisals annually and 16% semi-annually. Only 3% of organizations reported they did not conduct formal performance appraisals.

The performance appraisal is a periodic process that is primarily assessing the performance of employees over a past period in comparison to pre-set standards identified as "performance indicators". The performance appraisal tool uses numerical measures to provide business managers with quantitative outputs that would allow them to make judgmental decisions concerning employees' productivity. This was the preliminary purpose for introducing the performance appraisal system to organizations, which have increased over time to include motivating and improving employees' performance and determining fair allocation of pay increases and promotions.

Consequently, performance appraisal has become an integral part of the performance management process of any organization because it gives managers the ability to determine an employee's efficiency and make decisions on how to properly compensate and reward employee efforts, link organizational goals with individual

goals that assist in achieving organizational objectives with the right amount of resources, provide managers with indicators of employees' behaviors that requires corrections and capabilities that need to be developed, and yield quantitative data that can be used by managers to assess training needs, and appropriate utilization of human resources skills and abilities.

Although every organization tries to implement a “best practice” system, from their perspective, for its performance appraisal process, many of them are not utilizing the system correctly or are not satisfied with the outcome of their system. Watson Wyatt and WorldatWork (2005/2006) conducted a survey of 265 large U.S. companies across all industries and 1,100 workers. The survey responses from employers and workers agreed that the performance appraisal systems of their organizations need improvement. The survey found that about 98% of employers have adopted “best practices” including providing a formal yearly review, but they have been less successful in practice where 79% of employers say that managers at their organization are moderately or greatly effective in linking pay to performance, and only 52% of employees indicating that their managers tie pay to performance. In another recent survey conducted by Mercer (2013) about Global Performance Management, participants included 1,056 performance management leaders representing 53 countries around the globe varied in size, industries and structures, the results confirmed perspectives of many scholars and human resource professionals that currently adopted performance appraisal and performance management systems are ineffective and needs development. As evidenced by Mercer's survey, 51% of respondents reported that their performance management planning process needs work, 42% said their linkage to compensation decisions need

work, and 48% said their overall approach needs work. On enquiring about the most important outcome companies seek through performance management, 43% of respondents reported to drive employees to higher levels of performance, 21% said to provide performance feedback, and only 15% said to focus employees on the right things. These reports highlighted that organizations that are currently applied performance appraisal systems have to first determine the purpose of using performance appraisal systems, then to select the appropriate system for the organization and decide about the method of application.

A number of studies illustrated that when performance appraisal is conducted correctly, the process can provide managers with a number of valuable results (Carroll & Schneier, 1982, p.3) and (SHRM, 2010, p.289). On the other hand, scholars and professionals in addition to many surveys reported an existing gap between research and practice that inaccurate process of evaluating employee's performance leads to negative results such as employees' dissatisfaction and lower levels of performance (Carroll & Schneier, 1982, p.4), (DeNisi & Pritchard, 2006, p.254) and (SHRM, 2010, p.289). The number of companies reported that their overall performance management system delivers exceptional value is only 3% of respondents according to Mercer's survey of (2013).

Additionally, performance appraisal is frequently criticized by managers and employees as an unwelcomed and time consuming task that focuses on reviewing past work. The appraisal interview is becoming a perceived burden for many people whether as supervisors giving the appraisal and reluctant to provide frank feedback or as employees receiving them. This was evidenced by Buckingham and Goodall (2015), in

their process of redesigning the performance management system of Deloitte – a large global auditing firm. Through tallying the number of hours the organization was spending on performance evaluation process, Buckingham and Goodall (2015), found that completing the forms, holding the meetings, and creating the ratings consumed close to 2 million hours each year. Large amount of this time was primarily spent on appraisal meetings and leaders' discussions about the outcomes focusing on past performance. Researchers, as well, have criticized performance appraisal systems for being biased and subjective, especially in dealing with the ratings of the performance measurement indicators that report performance level quantitatively and link rewards to performance.

Obviously many of the prior studies have agreed that performance appraisal is a difficult process and established a direct relationship between effectiveness of the appraisal system, managers and employees' satisfaction, organizational fairness, and work performance. However, the performance appraisal process remains an integral function for organizations, although number of advocates have gone so far as to state that “employee evaluation process rarely accomplishes anything except create a paper trail used to discipline employees” and recommend to eliminating the practice of performance appraisal entirely if value is not added to the activity (Orr & Orr, 2014, p.168).

It is important for the organization and the individual that the task still be performed as effective as possible (Smither, 1998, p.132). The success of performance management system depends on choosing a credible performance measurement tool. There are many performance measurement tools that have been introduced and used by

organizations. Selecting the tool and process that best work for a company depends on the size and type of the business. Therefore, it is very important when choosing a system to consider level of complexity and objectivity. Implementation of complex systems requires higher financial resources and special human expertise. Objectivity is essential to gain credibility and ensure consistency of the system. It might be difficult to completely eliminate subjectivity but some tools are more objective than others. An effective system should be applicable to all levels of an organization from the least junior staff to the chief executive officer. According to Mercer's survey (2013), some performance measurement practices that are known to be valuable for development were found less prevalent in formal performance evaluation decisions. Globally, about 55% of organizations use 360-degree feedback / multisource feedback and 89% of companies use pay-for-performance philosophy.

Performance Evaluation in Healthcare

Healthcare is becoming one of the fastest growing and highly dynamic industries in the current time. Patient flow and intensity are also on the rise, which further increases the pressure on managements of healthcare organizations for a higher competition. This progress in the healthcare services has increased the demand for distinctive and individualized services to be delivered with utmost quality care, as well led to higher specialization in types and categories of healthcare jobs. Therefore, consistent with other types of businesses, healthcare organizations need to measure and manage the performance of the healthcare workers to guide the decision making process and motivate employees. In many ways, this is much more complex in healthcare

organizations as the decision making in healthcare involves a larger number of stakeholders with different and sometimes contradicting perspectives and priorities.

The quality of care delivered by healthcare organizations is a main driver for decision making in healthcare development that has become a common strategic objective for all healthcare organizations and workers in this field. With the large number of stakeholders in healthcare organizations, quality of care is often subject of debate that reflects the wide variation in stakeholders' perspectives. For decades, quality of care for healthcare management and medical practitioners lies in offering good practice of medicine. This perception has changed with the increasing importance of customer satisfaction for the success of businesses, as greater attention is given in recent days to the soft skills that formulate the way the service is delivered. Many studies and surveys have been conducted to evaluate the importance of soft skills in the workplace that found 85% of employee's success on the job is due to soft skills, whereas only 15% referred to hard skills (Watts & Watts, 2008; as cited in John, 2009, as cited in Robles, 2012).

Therefore, quality of care for patients is affected by the soft and hard skills used in delivering individualized services. On the other hand, quality of care for shareholders focused on the total image and reputation of the healthcare organization that could attract more patients. Those different variables in addition to others that are influencing the performance of healthcare organizations, intensify the need for more precise and reliable performance evaluation tools to guide the increasingly complex decision-making processes (Swaminathan et al., 2008; as cited in Traberg, Jacobsen, and Duthiers, 2014).

The literature shows that scholars in business management have introduced several methods and systems to assess the performance of healthcare workers and the quality of their services with continuous developmental efforts in this area purporting to identify what could be considered as the best practice for performance evaluation in healthcare.

It is important to note that the term “healthcare workers” in this study is used to cover employees working in healthcare services in all job categories including medical and non-medical jobs.

Statement of the Problem

Establishing an effective performance appraisal system is challenging for most organizations, and so the search for best practice solution for this core function. While researchers and professionals have identified several characteristics for effective performance measurement systems, many have agreed that the starting point is to determine job targets and key performance indicators that will help an organization to implement its business strategy. Setting clear and measurable performance indicators is considered one of the most critical factors to ensure accuracy of the outcomes. For example, using pay-for-performance seems easy and direct for goal oriented jobs where employee’s performance is measured upon achieving a pre-set numerical target such as sales and marketing jobs. The case is harder with jobs that offer soft services that are not linked with a numerical target. Further, small businesses usually use a standard measure that applies to all employees, with a higher dependence on self-assessment process. This tool often contributes to higher subjectivity and less credibility. Alternatively, the appraisal tool in larger organizations is generally customized to fit

different job categories. In most cases, however, performance appraisal remains a measuring tool that depends on reviewing the past results and behaviors and discussing how it was done versus how it should have been done. The effectiveness of this process is subject to how much the output data contributes to implementing the organizational strategy and achieving its goals. Too often performance measurement systems fail, in both small business and large organizations, because they were short sighted and unfocused. They lack the imperative integration with performance management. The major drawback in these systems is weakness in setting performance measurement metrics and lack of alignment of individuals' goals with organizational objectives from the start.

My experience in the last ten years involved working on implementing and managing performance appraisal processes in several organizations in varied industries. I have been challenged with many obstacles that affected the credibility of appraisal systems in these organizations, and the most challenging was during my work in the healthcare industry. Healthcare services are business facilities subject to unique operational boundaries. Other than public health facilities, all private healthcare providers operate for financial profit under high liability in providing quality care to the society. For this purpose, healthcare organizations depends on high level of soft skills in offering their services to patients who are usually emotionally influenced by the unhealthy feeling and their need for cure that requires higher level of sensitivity from employees in dealing with patients. Defining quality care service in such situations is debatable. Is it only to provide the service to the patient correctly and professionally according to the medical practice standards, or does it include the expression of feelings

in offering compassion and sympathy with patients' illness? Robles (2012) noted that soft skills as an indicator of job performance are as good as hard skills. Soft skills are not defined in the traditional sense of skills. The Collins English Dictionary defines the term soft skills as "*desirable qualities for certain forms of employment that do not depend on acquired knowledge: they include common sense, the ability to deal with people, and a positive flexible attitude*" (as cited in Robles, 2012). Soft skills are not limited to a single profession. People skills are a core component of soft skills (Cafasso, 1996; Klaus, 2010; as cited in Robles, 2012).

Therefore, soft skills requires more supervision and special experience to detect the difference between star performer who can take this extra unique step and balance between personal qualities and professionalism and the good performer who does what was expected from him or her to do correctly up to the standard limits.

The appraisal systems in such environment failed to capture the small difference between employees' performance when at end of the day they did what they were supposed to do. Quality care is usually controlled by the physicians who are the main service provider in such facilities, but when it comes to performance appraisals, all employees – medical and non medical – are judged by the same standards! The performance values that are associated with metric scales become more complicated and impractical when healthcare companies try to customize them into wider ranges of job categories. Using more than one performance appraisal approach in small companies or tailoring programs to the needs of a distinct business unit seems not feasible from the financial and human aspects. Mercer's survey (2013) indicated that three out of four global survey participants say that their performance management

practices are similar across business units and levels of leadership. Only 23% of companies indicate that the design of the executive performance measurement process differs from that of the rest of the workforce (Mercer, 2013).

Further, in the last few years, the trend of pay-for-performance incentive compensation have reached pay structures of physicians as main income generators for healthcare organizations; if this system is working as a measuring tool for performance of physicians, how is the performance of other medical supporting staff affected and measured?

Purpose of the Study

The purpose of this research is to investigate types of employee performance measurement systems currently applied in healthcare services and assess their efficiency in providing accurate measurement of the performance of healthcare employees' across their various job types with specific focus on performance indicators to assess soft skills.

Common Performance Measurement Systems

The study will review and analyze published literature on the application of selected performance appraisal tools, in addition to assess the strengths and weaknesses of these performance appraisal systems. For this purpose, the main performance appraisal tools that are commonly used and will be discussed in this research are:

Pay-for-Performance (P4P)

Although this is an incentive system that links performance to rewards, it is included in this research for its wide implementation as a motivator and performance

improvement tool, in addition to its application as a performance appraisal system. The pay for performance system based on linking performance outcomes through achieving departmental and organizational goals to rewards and punishments such as pay, promotion, or discharge. The system aims primarily at enhancing motivation for the direct impact of pay on performance (Rynes, Gerhart, & Parks, 2005).

In a survey of Fortune 500 companies, Lawker (2003; as cited in Smither, London, & Manuel, 2009, p.602), found that respondents thought that performance management systems are more effective when there is a strong connection between appraisals and rewards. According to Mercer's survey (2013), 89% of companies use pay-for-performance philosophy. The Mercer survey (2013) further reported the wide expansion of using this incentive system where 7 in 10 organizations in education and healthcare use pay for performance program.

Within healthcare, the current trend in pay for performance links clinical quality outcomes of hospitals and physicians to reimbursement by payers. The most common application affects physicians' performance and pay evaluation for their control over the provided quality of care (Helm, Holladay, Tortorella, & Candio, 2007).

360-degree Feedback

This method was first used by the DuPont Company in 1973 (SHRM 2010). The 360-degree feedback refers to managers collecting anonymous performance evaluations from more than one source that may include supervisors, direct reports, subordinates, colleagues, and customers, it could be anyone with whom the employee interacts frequently (Smither & Manuel, 2009, p.543). This approach allows rating the employee by multiple sources on a number of work-related behaviors that the organization

considers important for the job. The feedback is then aggregated to compare with the employee's self-rating. (Smither, 1998, p.345). The 360-degree feedback is also known as multisource feedback (Atwater, Brett, & Charles, 2007, p.285).

In a performance management survey conducted by SHRM Foundation in (2000), based on the responses of 480 human resource professionals, the 360-degree feedback was reported to be used by 32% of respondents' organizations and was claimed to be the only specific performance management area where companies planned to increase their activity during the following year. The Mercer survey (2013) found that about 55% of responding organizations use multisource feedback or 360-degree feedback.

Balanced Scorecard

This tool was introduced in 1992 by Harvard Business School professor Robert S. Kaplan and management consultant David P. Norton. The Balanced Scorecard based on financial metrics as a traditional measure for company success, linked with other metrics from three additional perspectives that are customer, internal process, and learning and growth. This approach started as a measuring tool for a company's performance and quickly developed into a total management strategic framework (Kaplan, 2010, p.3 - 4).

The balanced scorecard links strategies to measurable targets and actions. The balanced scorecard is considered one of the most important management innovations in the 20th Century and has been adopted by a wide range of healthcare organizations (Zelman, Pink, & Matthias, 2003, p.1).

Management by Objectives

In this method the managers set objectives for the employees in advance and monitor their performance periodically. The rewards are pre defined based on results and level of goal achievement. Management by objectives is also referred to as “goals management” and “management by results” (SHRM, 2010, p. 297). The theory of management by objectives has been introduced to the business science by Peter Drucker in 1954 (Kyriakopoulos, 2012).

In a survey conducted about twenty years after the concept of management by objectives has been introduced and adopted, almost 50% of America’s largest industrial firms of the Fortune 500, reported their attempt to utilize the system and that it has been adopted by a large number of the surveyed companies (Schuster & Kindall, 1974). Other several recent researches and articles confirmed that the approach of management by objectives is still in use and discussed its application in the healthcare industry.

This systematic review will analyze previous research literature and academic publications that have focused on implementation of the above selected performance measurement systems with concentration on applications in healthcare organizations. Specific attention will be given to find if their systems were effective in providing accurate measure of employees’ performance. The study will contribute to the existing literature by presenting key considerations for a best practice model of performance measurement system based on scholarly recommendations and practical experiences.

Research Questions

As the purpose of this research is to investigate types of the currently used performance measurement systems in healthcare organizations and assess their efficiency in providing accurate measurement of the performance of healthcare workers in their various job types with specific focus on soft skills, the following research questions were formulated to direct the focus of this systematic review towards the main objectives of this research:

- RQ1. Are the implemented performance appraisal and measurement systems effective in evaluating the performance of healthcare workers?
- RQ2. What are the key performance indicators for jobs based on soft skills?
- RQ3. Can a single performance appraisal system produce a valid and reliable measure for all job categories?
- RQ4. How is the performance appraisal output linked to the reward system?
- RQ5. Is there any best practice model for performance measurement in healthcare industry?

Method

Search Strategy

A systematic search in discipline specific databases and in multidisciplinary databases was conducted to identify the existing work published about employees' performance indicators and appraisals in the different disciplines with focus on healthcare related databases. The systematic search used the advanced search option in

all of the databases using the terms “employee performance indicators”, and “employee performance appraisal”, and “healthcare”. The search was limited to publications in English language without any restrictions on year of publication.

The literature search was performed electronically using the portal of the Bizzell Library of the University of Oklahoma. The search was conducted in different databases that were classified as: Discipline specific databases include PsycInfo, Medline, ABI/Inform, and Business Source Elite; and Multidisciplinary databases include Academic Search Elite, Health and Psychosocial Instruments, ProQuest, and Science Direct.

The use of Google Scholar was excluded for its limitation in exporting bulk citations to other database management software. Another challenge was encountered in using the portals of ABI/Inform and ProQuest because these search databases have restrictions on the number of the references that can be exported to external database management software. Only up to 4000 search result can be viewed and exported from these two databases. Therefore, although the search result shows a higher number of references on ABI/Inform and ProQuest, only the allowed maximum number of references were exported and indicated in the PRISMA flow chart for the search process. The order of appearance of search results in all databases were listed on the basis of the most recent to the oldest; therefore the maximum allowed number of references that was exported from ABI/Inform and ProQuest covered the most recent publications resulted from the search on these databases. The citations of the search results of every database were saved and exported to EndNote X7.5 for screening.

Study Selection

The screening process on EndNote started in eliminating the duplicated references. The remained references after the elimination of the duplicates were screened for relativity through examining the title of each reference. After eliminating all the references that were unrelated by the title, the full text for the residual studies was downloaded. The studies that its full text is not available were excluded. At this stage, articles were excluded if one of the following conditions applies:

- The title is not related to the purpose of this study.
- The context of the article is not directly related to the purpose of the study (performance measures in a context other than measuring employees' performance, or performance measures in a specific industry other than healthcare).
- The full text is not accessible.

The inclusion and exclusion criteria applied in this review is summarized in **Table (1)**:

Table 1 - Inclusion and Exclusion Criteria		
Factor	Inclusion Criteria	Exclusion Criteria
Language	English	All other languages
Publication design / Type	published and unpublished, research articles, research studies, thesis and dissertations, peer reviewed studies, chapters of books, and evaluations and development of performance measures	editorial articles, marketing publications, commentary publications, perspectives and survey reports, and papers on history or theory of performance and measurement tools
Time / dates	No restrictions	No restrictions
Industry	Healthcare organizations, healthcare professionals, and general (not for a specific industry)	specific industries other than healthcare
Application focus	Individual performance measurement	organizational performance measurement, organizational performance management, relationships between performance and other factors in organizational behavior domain

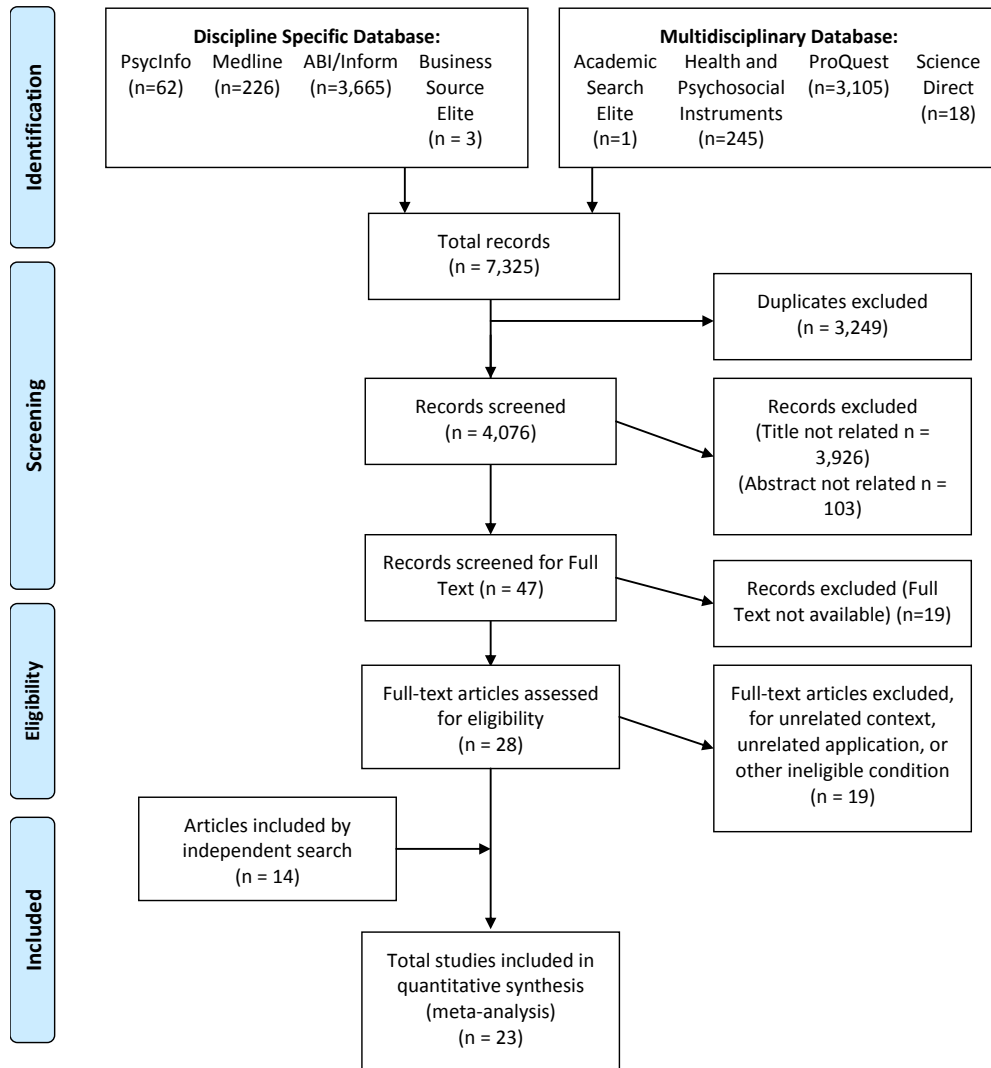
The remaining filtered articles were screened through reviewing the full text for eligibility and relevance to the purpose of this study and according to the inclusion and exclusion criteria that were set.

Results

Search Results

In total, the search terms initially yielded 7,325 potentially relevant articles for the purpose of this study. After eliminating duplicates and title and abstract review for irrelevancy, 47 studies remained for a more detailed screening. Screening for full text availability yielded 28 articles, with further screening for context irrelevancy yielded 9 articles that meet the inclusion criteria. Additional external independent search was conducted and references assessed that yielded 14 more articles added to the reference list that meets the inclusion criteria. Total number of eligible articles for inclusion in this systematic review is 23 articles. **Figure (1)** presents the PRISMA flow diagram detailing the systematic process of searching databases, screening and selecting studies for inclusion.

Figure 1 - PRISMA Flow Diagram



Data Extraction

Following the study selection and screening process, details of the selected articles were listed in a consolidated developed template on Excel worksheet for data extraction that include the following specific information: author(s), year of publication, title of the study, purpose of the study, and measures. **Appendix (1)** includes the Excel

worksheet list for details of the selected articles resulted from the systematic review and the added articles resulted from the independent search.

Characteristics of Studies

In this section, the included studies are described and compared on their characteristics and results. The total of 23 studies was identified of relevant contents, of which 11 studies (47.83%) were published between 2009 and 2016, 10 studies in the years from 2001 to 2008, and 2 studies were published between the years 1989 to 2000. Eleven studies, almost half of the studies (47.83%) were conducted or published in the USA and 5 studies (21.74%) were conducted or published in the UK. Two studies took place in the Netherlands, and one study in each of Greece, Denmark, Germany, Canada, and Iran. A summary of the characteristics of each of the included studies is provided in **Table (2)**.

Table 2 - Characteristics of Included Studies	
Characteristic	Number / Percentage (n = 23)
<u>Year of Publication</u>	
1986 - 2000	2 (8.70%)
2001 - 2008	10 (43.48%)
2009 - 2016	11 (47.83%)
<u>Country</u>	
USA	11 (47.83%)
UK	5 (21.74%)
Netherlands	2 (8.70%)
Iran	1 (4.35%)
Greece	1 (4.35%)
Denmark	1 (4.35%)
Germany	1 (4.35%)
Canada	1 (4.35%)
<u>Performance Relativeness</u>	
Pay-for-performance (P4P)	5 (21.74%)
360 Degree Feedback / Multisource feedback	5 (21.74%)
Balanced Scorecards	0 (0%)
Evaluation and Development	7 (30.43%)
Performance problems	6 (26.09%)
Helathcare relativeness	15 (65.22%)

The studies were classified into 3 categories: performance measuring systems and programs of which 10 studies fall under this category, evaluation and development of measuring tools includes 7 studies, and problems and gaps of measuring performance of medical practitioners that includes 6 studies. Five of the 10 studies classified under the performance measuring systems discuss the Pay-for-performance (P4P) measuring tool; the other 5 studies are about the 360 Degree / Multisource feedback (MSF) measuring system. A majority of the studies (n = 15, 65.22%), were directly related to healthcare. **Table (3)** shows the classification and distribution of studies.

Table 3 - Classification of Included Studies			
Category	Study	Industry / Focus	Purpose
performance measuring systems / program	Kirschner, K. et al. 2012	Healthcare / Primary care	Pay-for-performance (P4P)
	Miller, G. & Babiarz, K. 2013	Healthcare	Pay-for-performance (P4P)
	Emmert, M. et al. 2011	Healthcare	Pay-for-performance (P4P)
	Overeem, K. et al. 2012	Healthcare Professional / Physicians	Multisource feedback (MSF)
	Taylor, S. 2011	General	360 Degree Feedback
	ATWATER, L. et al. 2007	General	360 Degree Feedback / Multisource feedback
	Carson, M. 2006	General	360 Degree Feedback
	Helm, C., Holladay, C., & Tortorella, F., 2007	Healthcare Professional / Employees	Pay-for-performance (P4P)
	Elwyn, G. et al. 2005	Healthcare / Primary care	Peer Assessment
	Rynes, S., Gerhart, B., & Parks, L., 2004	General	Pay-for-performance (P4P)
Evaluation & Development	Payne, S. et al. 2008	General	Compare online performance appraisal (PA) system to the traditional paper-and-pencil (P&P) approach
	Carlos, V. & Rodrigues, R. 2015	General	To develop a self-reported measure that might be applicable across jobs and cultures
	Decker, J. 1999	Healthcare	Competence model in healthcare
	Grigoroudis, E. & Zopounidis, C. 2012	Healthcare	To develop multicriteria analysis that considers the complexity of the different job profiles
	Traberg, A. 2011	Healthcare	To design Management by objective (MBO) framework
	Wiese, D. & Buckley, M., 2016	General	evaluation to the performance appraisal processes and tools
	Sturman, M., Cheramie, R., & Cashen, L., 2005	General	examine measurement type (i.e., subjective and objective measures) and their reliability
Medical professionals' Performance problems / gaps	Cohen, D. & Rhydderch, M. 2006	Healthcare	To recognize doctors' underperformance
	Nikpeyma, N. et al. 2014	Healthcare	To explore problems of clinical nurses appraisal system
	Rowe, A. et at. 2005	Healthcare	determinants to the performance of health workers
	Overeem, K. et al. 2007	Healthcare	Methods to assess the performance of individual doctor
	Greenfield, J. 2015	Healthcare	the process of annual appraisal for practice nurses
	Stalker, M. et al. 1986	Healthcare	Performance appraisal process and indicators for nurses

Quality of Studies

An in-depth review of the selected studies revealed substantial variation in the quality of the studies in terms of content. Although the majority of studies had strongly demonstrated weaknesses of the existing appraisal systems, they poorly discussed details of performance indicators, with only one study suggesting a list of performance indicators for medical practitioners. None of the studies covered the problem of measuring performance of non-medical individuals' who work in healthcare organizations. Also, none of the studies explored performance indicators by job category or accounting for soft skills.

Generally, the content of the studies are of poor quality relative to the purpose of conducting a systematic review to explore the published researches and articles that discussed employee performance appraisal and indicators in the healthcare industry. The inadequate quality of the content of the studies could be attributable to an existing gap in covering the application of performance appraisal systems in healthcare organizations. A large number of studies focused on performance management at organizational level with limited attention given to measure employees' performance across the various job categories.

Data Analysis

Studies were classified and organized into three categories based on identifying a common factor in the main purpose of the study in order to create congruence between the studies that are classified under each category. This classification will enable eliciting a context relative to the purpose of this study from each category.

Key Findings

This research focused on the micro level of performance evaluation systems measuring individual performance. The study aimed to investigate and identify the types of the employees' performance measurement systems in use in the healthcare organizations and assesses their efficiency in providing accurate measurement of the performance of healthcare employees', among the various job categories with specific focus on soft skills. In reviewing the eligible studies included in this systematic review, a literature gap was identified in investigating detailed performance measurement indicators for the various job categories of workers in healthcare organizations and in the application of employee performance evaluation systems.

Analyses were undertaken by reviewing each study within the context of the category in which the study is classified. Specifically, six of studies included in the review explored problems of measuring performance of medical practitioners, ten studies were about performance measuring systems, and seven studies concerned with evaluation and development of performance measuring tools. Outcomes in general were limited.

Problems and gaps of measuring performance of medical practitioners

Measuring the performance of healthcare organizations requires understanding the contribution of healthcare workers and the impact of their underperformance on the organizational quality of care objective. Underperformance in healthcare is not limited to certain category of workers. As doctors are considered the central actor in healthcare organizations (Overeem et al., 2007), it is important to assess the reasons for their

underperformance like other healthcare workers. Two studies were identified eligible in this systematic review that discussed the assessment of doctors' performance, although the assessment and subsequent management of doctors' underperformance are not so well researched or documented.

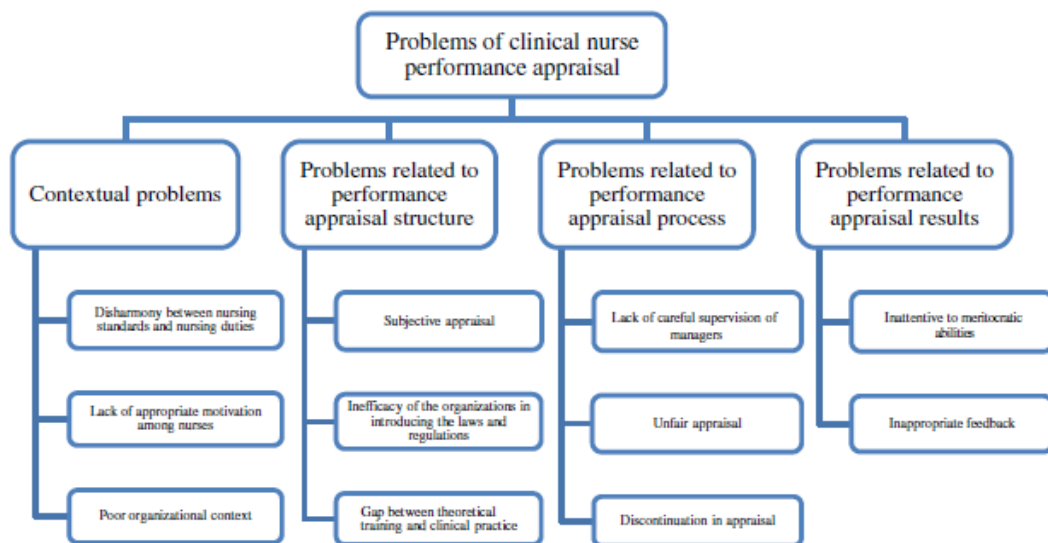
Cohen and Rhydderch (2006), investigated the phenomena of doctors' underperformance in the United Kingdom and found that stress in health professionals is high, with 28% showing above threshold symptoms compared to 18% of workers as a whole in the UK. In addition to the psychosocial factors that influence doctors' performance such as depression, burn out, alcoholism and drug addiction, the authors find it essential to evaluate other behavior and organizational elements and understand the individual motivation of doctors to perform.

Overeem et al. (2007), in conducting a systematic review to evaluate the feasibility of methods and instruments used in routine practice to assess the performance of individual doctors, identified 6 different methods that assess doctors' performance in real practice, extracted from 64 eligible articles. The methods are: simulated patients; video observation; direct observation; peer assessment (360-degree feedback/ multisource feedback); audit of medical records, and portfolio or appraisal. The methods were further classified to direct or indirect, summative (concerned with validity and reliability) or formative (effective in improving performance), with considerable differences in feasibility of the methods in terms of time and costs. Overeem et al. (2007) noted that *"Although the need for regular performance assessment of individual doctors is clear, the best way to do it is not"* (p. 1040). The systematic review revealed that the peer assessment method is most convenient in terms

of time, while portfolios and appraisals methods are most time-intensive for doctors with a significant variation in the methodological quality of the studies.

Performance of healthcare practitioners remain the focal point in this category. For example, Nikpeyma et al. (2014) explored the problems of clinical nurse performance appraisal system in large metropolitan teaching hospital in Tehran, Iran, and discovered that the most important reasons for the lack of efficacy in nurses performance appraisal system was the absence of fairness, objectivity, appropriate feedback, and staff participation in performance appraisal, and the absence of trained managers. Nikpeyma et al. (2014) suggested that there are four themes within the problem set facing clinical nurse performance appraisal systems: contextual problems, problems related to performance appraisal structure, problems related to performance appraisal process, and problems related to performance appraisal results. Details of the four themes are illustrated in **Figure (2)**.

Figure (2) Diagram of problems in clinical nurse performance appraisal



Even with the limited findings of this study, Nikpeyma et al. (2014), has addressed the problem of performance measures and indicators in healthcare

organizations when discussing the second theme – structural problems. The participants in this study considered the appraisal subjective and unfair because the items indicated in the appraisal form “performance indicators” are not properly defined and congruent with the working conditions, therefore they cannot evaluate the real nursing performance including professional information, clinical and communicative skills. However, Nikpeyma et al. (2014), concluded that in order to achieve high quality patient care, some dimensions of the appraisal system must be subject for revisions and modifications. Nikpeyma et al. (2014), did not present a detailed solution to the performance indicators problem.

Two other studies contributed to the importance of measuring the performance of nurses, with significant difference between the two studies in terms of context. Greenfield (2015) demonstrated the process of appraisal in a stepwise form clarifying what has to be done before, during and after the appraisal. Greenfield (2015) was very general in most of the details of the appraisal process and made only a limited contribution to the subject of measuring performance of workers in healthcare organizations.

On the contrary, Stalker et al. (1986), were thorough in describing the process of changing and improving the nursing performance appraisal system. The project was conducted in 1981 at Memorial Sloan-Kettering Cancer Center in New York. The main objective was to modify the nursing performance appraisal system to be used as a developmental tool in addition to its role of evaluation. The step-by-step process aimed at reducing the length of the appraisal instrument and provide clarity to the items of performance evaluation through the use of quantitative data and computer programs to

measure and evaluate the competencies and rating system. Although in the last twenty years, the performance measurement and management systems have developed significantly beyond the systems and processes that were applicable in the early 80's, Stalker et al. (1986) were successful in describing the process and the specific competencies and performance indicators that are fundamental to the philosophy of nursing practice. They recommended continuous examination to some categories in the performance evaluation system of nurses to reflect realistic performance standards that can still be used to-date.

The inadequate health-worker performance in low and middle-income countries was explored by Rowe et al. (2005). The study presented an overview of the factors that might influence the performance of health-workers and their motivation, and strategies needed to improve health-workers' performance. Lack of coverage to the components of performance appraisal system and their role in reflecting an actual picture of health-workers' performance was observed. Definition and analysis of the needed performance competencies and the validity of performance measurement methods were part of the recommendations of this study for further investigations to get over the knowledge gap in published literature. To ensure practical value of the findings, Rowe et al. (2005) emphasized the importance of making certain that findings for one setting in health area can be applied to other settings beyond low and middle income countries.

Performance Measurement Systems

Performance evaluation is conducted by organizations for many reasons; the most common objective is for performance improvement through salary administration, performance feedback, and identification of employee strengths and weaknesses

(Aguinis, 2005). Therefore, objective feedback and accurate administrative decisions that link performance evaluation results to rewards system are core factors for the success of a performance measuring system.

Most running organizations have a form of employee performance evaluation system. Some organizations incorporated one system and modify it to fit the organizational needs, where others tried more than one system aiming at finding the best fit. In this systematic review research, the findings of the 10 eligible studies that were classified under performance measuring systems will be discussed in this section.

Description to the concepts and implementation experiences of Pay for Performance (P4P), 360 Degree / Multisource Feedback (MSF), and balanced scorecard systems will be demonstrated with some of the advantages and disadvantages of those measuring tools.

Pay for Performance (P4P)

In 2013, a survey conducted by Mercer reported that 89% of companies use pay-for-performance philosophy, and in a reflection to the growing interest in pay-for-performance programs, Kirschner et al. (2012), indicated that 7 in 10 organizations in education and healthcare use pay for performance program to improve quality of care. In this systematic review, 4 studies out of 5 were about the context and application of pay for performance in healthcare. The use of performance incentives in health programs started in the 1990's by rewarding both the process indicators and measures of clinical quality (Miller & Babiarz, 2013).

Scholars have identified monetary rewards as one of the main motivating factors that influence individuals' behavior in work settings, as Locke et al. (1980), stated "*Money is the crucial incentive . . . no other incentive or motivational technique comes even close to money with respect to its instrumental value*" (p. 379) (as cited in Rynes et al., 2005). On the basis of this concept, the pay for performance (P4P) program aims primarily to enhance and motivate performance using the direct impact of pay on performance (Rynes, Gerhart, & Parks, 2005), through linking the performance appraisal results to rewards and punishments. Emmert et al. (2011), found that P4P in healthcare operates on the same conceptual basis that physicians' behavior in practicing medicine can be influenced by the financial incentives (Hillman et al. 1989; Hellinger, 1996; Gosden et al., 2000; and Town, 2004; as cited in Emmert et al., 2011).

Within healthcare, the current trend in P4P links clinical quality outcomes of hospitals and physicians to reimbursement by payers. This application has an effect on physicians' performance and pay evaluation for the control of physicians to the provided quality of care (Helm, Holladay, Tortorella, and Candio, 2007). Miller and Babiarz (2013), measured the effectiveness of P4P in improving the performance of healthcare workers and consequently enhance the quality of care with the introduction of monetary rewards as extrinsic motivators. The researchers assessed the potential consequences on the performance of individual healthcare providers. The revealed psychological effect demonstrated that the use of financial incentives as an extrinsic motivator works well on the short run, but may lead to unintended consequences overtime such as demoralization (Oxman & Fretheim, 2008), reductions in intrinsic motivation - such as social or self-image (McDonald et al. 2007; Ashraf, Bandiera, and

Jack, 2012), less trust between patients and providers (Ellingsen & Johannesson, 2008), and decline in quality of individuals entering the public health workforce if the use of financial incentives selects against intrinsically motivated health care workers (Witter et al., 2012; as Cited in Miller & Babiarz, 2013).

Rynes, Gerhart, and Parks (2004) also investigated the psychological consequences of using P4P in different work settings based on the psychological relation between pay and motivation. In a meta-analysis, Rynes et al. (2004) presented several examples of the incentive effect of P4P in improving performance of individuals. Lazear (1986; as cited in Rynes et al. 2004) for example, found a 44% increase in productivity when a glass installation company switched from salaries to individual incentives. Roughly about 50% of this increase was due to increased productivity of existing workers, while the other 50% attributed to new productive workers who replaced underperformers. Although, Rynes et al. (2004) identified a significant role for P4P in improving performance, it was still difficult to decide on the performance measures that can objectively identify individual contributions. Accordingly, Rynes et al. (2004) investigated the advantages and shortcomings of the alternative choices to determine which performance measures are more appropriate. In this regard they compared Behavior-Based (Subjective) measures versus Results-Based (Objective) measures, Incentive Intensity (strength) measures, and Individual versus Group (or Collective) Performance measures. On this issue, Miller and Babiarz (2013) commented that although very few P4P programs have rewarded good health, rewarding health outcomes rather than health input provide strong incentives for providers to exert effort, and encourages them to use their knowledge creatively to

innovate in developing new, meaningful delivery strategies. In practice, this could be possible for upper-level managers who possess greater flexibility for innovation in service delivery where rank-and-file health workers must follow detailed, highly prescriptive protocols from which they are not allowed to deviate. Further, the efforts of the rank-and-file health workers eventually have a straight impact on organizational performance because they have the direct contact with target populations. On the other hand, rewarding health workers for their own individual performance may create disincentives for teamwork or cooperation; conversely, rewarding providers for group performance creates unjustified incentives for underperforming individual health workers who may privilege (free-riding) be rewarded among co-workers. Therefore, deciding about incentive programs to improve performance of healthcare workers, such as P4P program, is a complicated issue that requires continuous balancing between financial compensations and patient well-being to avoid unintended and perverse consequences (Miller & Babiarz, 2013).

Helm, Holladay, & Tortorella (2007), evaluated the effectiveness of implementing P4P program in aligning employee's goals to institutional goals and linking performance to rewards. Joinson (2001; as cited in Helm et al., 2007), noted that many companies are getting close to alignment of objectives by adopting an annual performance appraisal system in which employees' performance is evaluated at the same time every year. When reporting appraisal results at a time closer to company's annual budget and business plan, the individual performance results are aligned with organizational objectives and the allocation of financial rewards. For the purpose of their study, Helm et al. (2007), administered an employee evaluation questionnaire with

healthcare workers at the University of Texas M. D. Anderson Cancer Center that demonstrated improvement in employees' performance when employees objectives are linked with the organizational objectives. The study also presented that employees who are high performers usually expect higher financial rewards and recognition than an average performers. The study confirmed a strong positive relationship between employee's performance and perception with the compensation and rewards system. However, Helm et al. (2007) did not measure the impact of performance improvement on the quality of care.

The economic efficiency of the P4P program assessed by Emmert et al. (2011), who carried out a systematic review to explore evidence for the efficiency of the system. Nine studies were identified and revealed a majority that is in line with the potential of P4P programs to improve quality of care but with higher costs (Nahra et al., 2006, Kouides et al., 1998, An et al., 2008, and Lee et al., 2010; as cited in Emmert et al., 2011). The targeted quality measures in the systematic review conducted by Emmert et al. (2011), was reported to have varied widely across studies with no study focused on the same quality measure that was evaluated in another study. A number of studies focused only on process quality measures that examine whether desired steps are being taken. Emmert et al. (2011), concluded that P4P will continue to be a popular improvement strategy in health care, despite that difference among studies held up conducting a meaningful comparison of results.

With an experience in different P4P programs, Kirschner et al. (2012), identified three framework components for a distinguished P4P program: performance measurement, appraisal and reimbursement. Performance measurement consists of valid

and reliable indicators that make sense to the target group. Appraisal includes described analysis and interpretation of data based on well defined unit of assessment and performance standards that in turn contributes to formulating the reimbursement structure. In application, Kirschner et al. (2012), clarified that performance measures do not cover all aspects of general practice resulting in the unavailability of accurate analysis and interpretation of data.

Another remarkable feature identified by Kirschner et al. (2012), is that policy makers usually follow a top-down strategy to design P4P that is also the same strategy followed by managers on implementation. To ensure effectiveness and improvement of this system, a bottom-up procedure that involves target users should be applied. Further, the sustainability of the program is at risk if the performance indicators are not revised constantly to reflect the core functions of the evaluated standards. This will also assist in avoiding a narrow focus on the standards of quality of care in healthcare practice.

A definitive conclusion about P4P efficiency cannot be made with the available evidence. Emmert et al. (2011), indicated some factors that may improve the P4P efficiency include increasing incentive size, rewarding absolute performance and performance improvement, and minimizing time gap between care delivery and payout.

360-Degree Feedback / Multisource Feedback (MSF)

360-Degree Feedback, also known as Multisource Feedback (MSF), is a performance appraisal methodology that is used to evaluate employee's performance through acquiring information from employee's supervisors, colleagues, subordinates, and, possibly, customers. A self-evaluation report is also obtained from the person being appraised. All ratings are added up in a consolidated report (Carson, 2006). The

MSF approach became more popular in the early 1990s, when organizations realized the limitations of one-on-one appraisal system in terms of time, effort, and quality of results. The MSF appeared to offer more holistic and/or realistic evaluation that might also be more objective by involving more reviewers for a single employee (Carson, 2006).

Atwater et al. (2007) reported that about 55% of respondents to Mercer's survey (2013) use multisource feedback (MSF) or 360-degree feedback, (Church, 2000; as cited in Atwater et al., 2007). According to human resources consulting firm William M. Mercer, the number of American companies reported using the 360-degree feedback has grown from 40% in year 1995 to 65% in year 2000 (Pfau & Kay, 2002; as cited in Carson, 2006). This number is still on the rise with expectations for the MSF system to spread to other parts of the world (Atwater et al., 2007).

In healthcare applications, MSF allows an external evaluation to the performance of medical practitioners by medical peers performing similar scope of responsibilities, non-medical co-workers, and patients reflecting customers' view. Having this multisource broader evaluation compared to the employee's internal self-evaluation provides a better picture of the actual performance of the medical practitioner that can assist in achieving the highest quality care (Overeem et al., 2012). Therefore, since 1993, the healthcare organizations are increasingly using the MSF system for its key feature in assessing multiple components of professional performance from different perspectives (Overeem et al., 2012).

In this systematic review, five studies met the eligibility criteria for inclusion in this section. Two of the studies are healthcare related that used the MSF program to

evaluate the performance of physicians: Overeem et al. (2012), performed in a non-academic hospitals in Netherlands that involved the physician's self-rating, in addition to ratings of peers, co-workers and patients, and Elwyn et al. (2005), conducted in a primary care setting in UK that used The Peer Assessment Questionnaire (PAQ) for evaluating the performance of doctors. The study of Elwyn et al. (2005), is considered the first study that proof feasibility of peer assessment system in a UK primary care setting with using an approach similar to MSF in involving peers and co-workers such as nurses and management. The peer assessment review is conceptually similar to the MSF but does not include patients or customers. No explanation was given for not involving patients in the UK evaluation although of the important view of patients in reflecting the humanistic elements such as integrity, respect and compassion.

Both studies confirmed the reliability and validity of MSF system in evaluating the performance of Physicians. The significant correlations between ratings of peers, co-workers and patients, presenting different perspectives from three independent groups of raters, and the large similarity to the ratings obtained in a US studies, support the conclusion reached by Overeem et al. (2012), and Elwyn et al. (2005). Further, this approach facilitates a structured system for collecting information about the performance of healthcare workers. It also demonstrates evidence for practicing the principles of Good Medical Practice that includes good working relationships with colleagues and patients based on aspects of health and integrity (Overeem et al., 2012).

Some of the identified weaknesses and limitations of using MSF system in evaluating the performance of physicians included the tendency of Physicians to rate other members of their physician group more positively (Overeem et al., 2012), and the

potential creation of organizational ‘cluster’ effect and rating bias when most of the raters work in the same organization (Elwyn et al., 2005).

An alternative perspective was reflected in the other three studies included in this systematic review (Taylor 2011, Atwater 2007, and Carson 2006) from experiences on applications in different sectors other than healthcare with recognized high similarity in the outcome implications. In order to achieve successful implementation of MSF system, all studies confirmed the importance of having a clear plan and set of individual goals that are linked to organizational objectives before starting the process. As feedback is crucial for the success of this system, the raters must be well trained on the process in general, and on dealing with the negative feedback in particular. An environment of trust has to prevail to protect anonymity and confidentiality of the gathered information and performance feedback. In addition, despite that MSF system can be used for developmental and evaluation purposes, a collective recommendation from all studies was to start using the MSF system for developmental purposes first before the evaluation purpose. This recommendation aims to familiarize raters and appraised individuals on the process and evaluation scales, encourage the process of feedback and follow up on performance developmental action plans, promote environment of trust in the system and its process, and ensure confidentiality of information. Once raters and employees trust the system and become familiar with the rating scales, their behavior will be influenced with higher honesty that results in a more accurate evaluation.

In conclusion, authors of the included studies confirmed that MSF process can be a meaningful multi-rater method for performance appraisals that demonstrated its

effectiveness in different work settings. Carson (2006) is the only researcher who suggested a need to customize the performance indicators to directly reflect the skills and behaviors required to fulfill the responsibilities of the position subject for evaluation. Such customization would bring more meaningful results through instilling trust when individuals who are being appraised recognize the connection between their responsibilities and the performance measures.

Balanced Scorecards

A significant growth in the implementation of the balanced scorecard system has been witnessed across many industries since it was introduced in 1992 as a performance measurement tool developed by Harvard Business School professor Robert S. Kaplan and management consultant David P. Norton. The balanced scorecard overcomes the limitations of using traditional financial measures to provide an accurate picture of a company's performance. The balanced scorecards link the overall corporate strategies to measurable targets and actions through a framework that uses four perspectives: the customer's perspective; an internal business perspective; an innovation and learning perspective; and the financial perspective (Kootanaee et al., 2013).

Although Zelman et al. (2003), reported finding 142 articles on the balanced scorecard published in the period from 1999 to 2001, this systematic search revealed many articles concerning applications of the balanced scorecards that measures organizational performance in different industries, none of the articles met the criteria of inclusion in this systematic review. This could be attributed to the holistic strategic nature of applications in which the balanced scorecard is used. Most applications of the balanced scorecards methodology in healthcare organizations are designed for the entity

in total aiming at a long-term adaptation to achieve organizational mission (Zelman et al., 2003). **Figure (3)** summarizes the different types of applications of the balanced scorecards in healthcare organizations (Zelman et al., 2003):

Figure (3) Healthcare Applications of the Balanced Scorecard

Type of Application	Examples	Source
Organizational Performance	Organizations Listed in Figure 1	See Figure 1
Public Information (see n.36)	Patient Care Report Cards	Lowe and Baker Badger
Clinical Pathway (see n.37)	Cardiac Prevention	Levknecht <i>et al.</i> Schriefer <i>et al.</i>
Hospital Department Performance (see n.38)	Operating Rooms	Mathias
	Information Technology	Niss Gordon and Geiger
	Medical Rehabilitation	Cohen <i>et al.</i>
Quality of Care and Outcome Measurement (see n.39)	Breast Cancer	West <i>et al.</i>
	Mental Health	Rosenheck
	Renal Transplant	Colaneri
	Renal Dialysis	Peters and Ryan
	Post-Op Nausea	Graumlich <i>et al.</i>
Managed Care Evaluation (see n.40)	HEDIS	Kenkel
Performance Measurement of a Consortia of Hospitals (see n.41)	CRISP	Bergman

This systematic search could not identify any study that discusses individual performance measurement through implementation of the balanced scorecard in healthcare organizations. However, with the recognition to the value of the balanced scorecard in strategic management in healthcare industry, a number of organizations have applied modifications to the original formulation and framework that were introduced by Kaplan and Norton. **Figure (4)** presents examples for some of the balanced scorecards modifications in the healthcare organizations that were sourced by Zelman et al. (2003):

**Figure (4) Selected Additional Balanced Scorecard Perspectives
Used by Healthcare Organizations**

Domain Modification	Health Care Organization	Source
Development and Community Focus Human Resources Quality of Care and Services (see n.44)	Long-Term Care Industry—Ebenezer Society and Board of Social Ministries	Potthoff <i>et al.</i>
Clinical Productivity and Efficiency Mutual Respect and Diversity Social Commitment External Environmental Assessment Patient Characteristics (see n.45)	Outpatient Operations—Mayo Clinic	Curtwright <i>et al.</i>
Outcomes (see n.46)	Health Network—Carondelet Health Network Behavioral Health Care	Santiago
Shifted Customer Perspective to Top of Scorecard (see n.47)	Physician Practices—Yale Faculty Practice (Yale University School of Medicine)	Rimar
Revision of Domain Questions (see n.48)	Academic Medical Centers	Zelman <i>et al.</i>

Kootanee et al. (2013), stated that a key advantage of using the balanced scorecard method is getting a balanced view of company performance that covers the four principle business perspectives on a current and long-term vision. On the other hand, some of the disadvantages are the requirement for a long planning and implementation process and the obligation to be part of a bigger strategy for successful implementation.

In conclusion, even if the Balanced Scorecard seems a holistic strategic system that cover four dimensions of the business with intense focus on the processes and finance, the system’s practical implementation is complicated and present less importance for people and the organizational culture which could impact the individual and organizational performance (Traberg, 2011).

Table (4) provides insight into the main strengths and weaknesses of the different performance measuring tools that were assessed in this systematic review: Pay-for-performance, 360-degree / Multisource feedback, and Balanced scorecard:

Table 4 - Pros and Cons of Performance Measurement systems

Pay for Performance (P4P)	360 Degree Feedback / Multisource feedback	Balanced Scorecards
Advantages (Pros)		
<ul style="list-style-type: none"> – May provide accurate analysis and interpretation of data when designed with clear defined unit of assessment and a linked performance standards.¹ – Monetary rewards as extrinsic motivator have direct impact on improving performance.² – P4P in healthcare operates on the same conceptual basis that physicians’ behavior in practicing medicine can be influenced by the financial incentives.⁴ – Rewarding health outcomes may increases motivation to use knowledge creatively to innovate in developing new delivery strategies.³ 	<ul style="list-style-type: none"> – Used more by healthcare for its key feature of assessing multiple components of professional performance from different perspectives.⁵ – More holistic and/or realistic than one-on-one that also being more objective with involving more number of reviewers for a single employee.⁶ – Facilitates a structured system for collecting information about the performance of healthcare workers.⁵ – Because of the feedback factor, individuals generally improve their performance following 360-degree feedback.² 	<ul style="list-style-type: none"> – Links the overall corporate strategies to measurable targets and actions.¹⁰ – Provide a balanced view of company performance that covers the four principle business perspectives on a current and long-term vision.¹⁰
Disadvantages (Cons)		
<ul style="list-style-type: none"> – Could be more prone to deficiency in performance measures (e.g., paying for quantity without adequate attention to quality).² Stimulating the incentivized parts of the performance can result in a possible decline in quality of 	<ul style="list-style-type: none"> – Requires a minimum number of participants to maintain some level of anonymity among the reviewers.⁷ 	<ul style="list-style-type: none"> – Each organization must engage in the full range of strategic management activities, from defining its mission to the selection of goals and strategies, in order to develop its own unique scorecard and to assist progress

care of the non-incentivized aspect. ¹		toward the selected goals. ¹⁰
– Mostly designed and implemented top-down by policy makers and managers. A more bottom-up procedure in designing a P4P program may improve its future implementation and its effectiveness. ¹	– Managers "often look only at the last few months" of data. ⁷	– The theory and concept of the balanced scorecard requires significant modification to reflect the realities of different industries and organizations. ¹¹
– Rewarding health workers for their own individual performance may create disincentives for teamwork cooperation. Alternatively, rewarding providers for group performance creates incentives for free-riding. Both individual and group-based pay plans have potential limitations. ^{2 & 3}	– Tendency of Physicians to rate other members of their physician group more positively. ⁵ Create possibility of "political coalitions", ⁷ and organizational 'cluster' effect and rating bias when most of the raters work in the same organization. ⁸	– Most applications of the balanced scorecards methodology in healthcare organizations are designed for the entity in total aiming at a long-term adaptation to achieve organizational mission. ¹¹
– Financial incentives as extrinsic motivator works well on the short run but may lead to unintended consequences on the long run. ³	– Different viewers define the scales differently, and that skews the results. ⁷	– requirement for a long planning and implementation process and the obligation to be part of a bigger strategy for successful implementation. ¹⁰
– Leads to cost increases unless improvements in quality are large enough. ⁴	– Time and effort associated with MSF may deter continued use. ⁹	

Sources:

¹(Kirschner, 2012)

²(Rynes, Gerhart, & Parks, 2004)

³(Miller & Babiarz 2013).

⁴(Emmert, M. et al., 2011).

⁵(Overeem et al., 2012)

⁶(Carson, 2006).

⁷(Taylor, 2011)

⁸(Elwyn et al., 2005)

⁹(Atwater et al., 2007)

¹⁰(Kootanaee et al., 2013)

¹¹(Zelman et al., 2003).

Evaluations and Developments

This section encompasses two domains: (a) evaluation of existing appraisal systems or processes and (b) development of processes or methods. This systematic review identified 7 studies under this section with 3 studies in each domain and one meta-analytic study that evaluated the performance measures and ratings based on the nature of the individual performance.

The meta-analytic research of Sturman, Cheramie, and Cashen (2005), has a specific importance to this systematic analysis review for the identification of the nature of the individual job performance and the factors that may impact rating of the individual performance and affect the reliability of performance measurement indicators. The authors studied the dynamic performance theories and assessed previous findings showing that past performance predicts future performance, and how this relationship is moderated by time, job complexity, and the methods of performance measurement. The current performance evaluations as discussed by Sturman et al. (2005), are based on assessing either the behaviors of employees (subjective methods) or results of their actions (objective methods). The performance dimensions evaluated by the supervisor influence the ratings of subjective measures of performance (Rotundo & Sackett, 2002; as cited in Sturman et al., 2005). Nevertheless, this influence allows the rater to consider factors outside of the employee's control when evaluating performance. On the other hand, the ratings of objective measures do not account for circumstances outside of the individual's control and may ignore the factors that may influence the consistency of performance over time. Accordingly, the research has illustrated that objective and subjective measures of job performance are not

interchangeable (Bommer et al., 1995; Heneman, 1986; as cited in Sturman et al., 2005). Also, Sturman et al. (2005), demonstrated in their review that the individual performance may vary over time depending on internal and external factors such as individual characteristics and job complexity. The test-retest reliability assessment that they performed provided evidence that existing performance measurement systems do not provide a true measure of individual's performance beside that using objective measures versus subjective measures of performance may not be useful in all job types. Consequently, the results of this study provide valuable guidance to the professional practitioners to consider this dynamic nature of individual performance in designing performance measurement and rating systems and in carefully considering more than one single measure in selecting performance measures and indicators.

Although the research of Sturman et al. (2005), is not specifically discussing performance measurement method and systems in healthcare, it has a great implication when applied on the workers of healthcare organizations in terms of the wide diversification of their job types and complexity, beside the nature of the individual job performance that in many cases requires subjective and objective measures.

Evaluations of Appraisal Systems

The three studies included under this section are similar in terms of evaluating components of individual performance appraisal systems. However, the studies are varied in their content and the element of the appraisal system that each study has analyzed.

Through summarizing the historical evolution of the performance appraisal process, Wiese and Buckley (2016), listed the main implemented performance appraisal

tools and critically analyzed their strengths and weaknesses. The study explored the efficacy of each tool in measuring the performance of employees and achieving the organizational goals. In the process of this evaluation, Wiese and Buckley (2016) distinguished between organizational goals, rater (appraiser) goals, and ratee (appraised individuals) goals. Based on a description of Cleveland et al. (1989; as cited in Wiese and Buckley, 2016), organizational goals are such as workforce planning and determining organizational training needs. Rater goals are referred to as administrative purposes such as recognition of individual's performance to make decisions regarding salary administration, promotions, retention, termination, and layoffs. Ratee (appraised) goals are those identified by Cleveland et al. (1989), as “within person” such as feedback on performance strengths and weaknesses to identify development needs. The legal requirements are the documentation purposes that fall under the organization and rater goals such as documenting personnel decisions and conducting validation research on the performance appraisal tools. The organizational continuous attempts to achieve all these goals while using tools that were designed for one type of purpose yields dissatisfaction in the existing tools and increase demand to solve the complications of the performance appraisal systems. Therefore, Wiese and Buckley (2016) criticized the researchers’ focus on reducing errors of the existing tools and neglecting the actual measuring objective of the tool. Wiese and Buckley (2016), also pointed to the emerging social, political and technical changes in organizational environments that consequently change the workforce characteristics and bring new jobs and roles. Wiese and Buckley (2016), recommended that researchers take into account all these factors

when designing an appraisal form where a single tool would not reflect the important aspects of the work performed in different scopes of jobs.

Payne et al. (2008), measured the efficacy of using an online performance appraisal (PA) system compared to the traditional paper-and-pencil (P&P) approach through assessing employees' reaction to the different mechanisms. Traditionally, the effectiveness of PA has been measured with rater errors that reflect rater accountability, rating accuracy that reflects quality of evaluation, and perceived security and confidentiality of information (Cardy & Dobbins, 1994; as cited in Payne et al., 2008). In addition, the qualitative aspects of the appraisal, including employees' reactions that captures the extent to which the appraisal process and outcome met the employees' expectations are also contributing to the effectiveness of PA and influence employee motivation and productivity (Dickinson, 1993, Ilgen et al., 1979, Larson, 1984, Pearce and Porter, 1986, and Wexley & Klimoski, 1984; as cited in Payne et al., 2008).

This quasi-experiment carried by Payne et al. (2008) demonstrated that the online PA gained higher levels of rater accountability and employee participation with no difference in perceived rating security. Although that using the technology and online sources is currently utilized in most human resource practices, this study adds value to the literature in realizing the impact of the approach and the importance of the content of the appraisal system when designing and modifying performance measurement systems.

Decker (1999) discussed one of the main gaps in measuring performance of healthcare workers. While the main objective of healthcare organizations is to provide quality care, most performance measuring systems evaluate the practice of medical

practitioners to assess their level of quality of care, not giving equivalent importance to customer satisfaction as the end user of service. This weakness in healthcare performance measurement systems could be referred to two common elements: 1) healthcare standards are usually derived from the policies and standards of the Joint Commission on Accreditation of Healthcare Organizations (JCAHO), or the equivalent authorities in countries other than USA, that only focus on the competencies of the medical operation, and 2) recruitment and selection competencies are based on job skills and knowledge rather than the attitudinal and interpersonal skills that stimulate customer satisfaction, (Decker, Strader, and Wise, 1997a; as cited in Decker, 1999).

In this study, Decker (1999) describes the competencies that need to be incorporated by healthcare facilities and improved upon. The competencies that Decker (1999) identified as “hidden” competencies are self-esteem, accountability, professionalism, emotional self-control, teamwork and customer service. To further clarify the connection of these competencies to the customer-outcome expectations, Decker (1999) provided samples of behaviors for worker characteristics that would reflect each competency and demonstrate superior performance from the customer perspective. **Table (5)** illustrates examples of the behavioral competencies:

Table 5: Example of the behavioral competencies

Competency	Behavioral measure
Professionalism / Accountability Respect	<ul style="list-style-type: none"> – Does not talk down to others or over a patient of customer. – Does not cause defensiveness in others. – Acknowledges customer’s presence with eye contact or statement within 5 seconds of customer’s entering
Accountability	<ul style="list-style-type: none"> – Does not complain to customers or blame administration in front of customers for decreased staffing – Demonstrate emotional self-control – Solves unit problems independently without waiting for

	supervisor
Personal / Professional Development	<ul style="list-style-type: none"> - Participates in continuing education - Continually improves knowledge of job - Accountable for one's actions to customers and coworkers
Decision Making / Problem Solving	<ul style="list-style-type: none"> - Uses outcome data to make decisions - Uses performance improvement process in decision making - Demonstrates critical thinking / judgment skills in performance of duties and tasks
Support for the Organization's Values / Goals	<ul style="list-style-type: none"> - Does not spread rumors - Corrects misinformation - Encourages and support others
Customer Service / Communication Skills	<ul style="list-style-type: none"> - Speaks Courteously to customers - Offers and accepts constructive criticism - Practice active listening
Customer Service	<ul style="list-style-type: none"> - Provides timely and clear information and follow-up to requests from customers / patients - Offers assistance without being prompted - Holds in high regard the dignity and respect of patients

Decker (1999) elaborated on how healthcare organizations can build a culture of professionalism and accountability when this competency model is incorporated into the recruitment system through selection and assessment criterion for all job categories of healthcare organizations. These competencies provide a guideline for solving the problem of measuring soft skills. However, more research is required to identify the appropriate rating and scaling system that was not covered in this study.

In summary, the performance appraisal components that were discussed in those three studies have implications on the healthcare workers in different aspects. Combining the findings could set the foundation for a new performance measurement structure that meet the specific needs of the healthcare workers and contribute to better measurement of their performance.

Developments of Appraisal Methods

The importance of having an effective employee performance appraisal system is broadly recognized as an essential tool that enables organizations to assess the performance of its employees and their contribution to the organizational long-term objectives. However, most existing performance measurement systems were designed for general application with no consideration given to the particular needs of different organizations. Few studies that aim to improve individual performance measurement systems have focused on the particular aspects of the healthcare organizations; even fewer offer a holistic approach. This research has identified three studies that demonstrated three different approaches to improve individual performance measurement systems with two of the studies focused on healthcare organizations.

From defining the concept of job performance as evaluative and episodic behaviors that an individual adopts towards her/his work and job, as a result of the dynamics between cognitive abilities, personality and learning experiences, that aggregate value to the organization¹, Carlos and Rodrigues (2015) measured job performance using two dimensions and eight sub dimensions: task performance includes three sub dimensions (job knowledge, organizational skills, efficiency) and contextual performance includes five sub dimensions (persistent effort, cooperation, organizational consciousness and interpersonal and relational skills). Their objective was to develop an individual performance measurement system that might be applicable

¹ Job performance (JP) is characterised as a dynamic (e.g., Motowidlo et al. 1997; Sonnentag and Frese 2002), multidimensional (e.g., Campbell et al. 1990a, b; Motowidlo et al. 1997; Viswesvaran 2001; Sonnentag and Frese 2002; Cheng et al. 2007), behavioural (e.g., Campbell et al. 1990a, b; Motowidlo et al. 1997; Viswesvaran 2001), episodic (e.g., Motowidlo et al. 1997) and evaluative (e.g., Motowidlo et al. 1997) concept.

across jobs and cultures. Similarly, in considering the complexity of different job profiles in healthcare organizations, Grigoroudis and Zopounidis (2012) proposed an approach for an employee evaluation system that is based on multi-criteria analysis to evaluate employees on a set of different but specific job dimensions using a variant of the UTA² method.

Upon an extensive literature review, Carlos and Rodrigues (2015) included a detailed description of the process to identify the items that shall be used to measure each dimension and the new scale of evaluation that was developed following Churchill's (1979), Aguinis, Henle and Ostroff's (2001) and Viswesvaran's (2001) guidance. From the two identified dimensions: task performance and contextual performance, the authors developed a pool of individual items to measure each sub dimension. Carlos and Rodrigues (2015), measured content validity and filtered the items by implementing revisions and suggestions from 12 experts in job performance, 14 journal editors and 18 academics who reviewed and tested the pool of items. The instrument developed by Carlos and Rodrigues (2015), presented good psychometric properties. Also, the self-reported performance measure was developed to be useful for different job contexts in higher education and to complement the subjectivity of supervisor ratings. However, the measure still needs cross-validation with cultures and professional areas other than higher education.

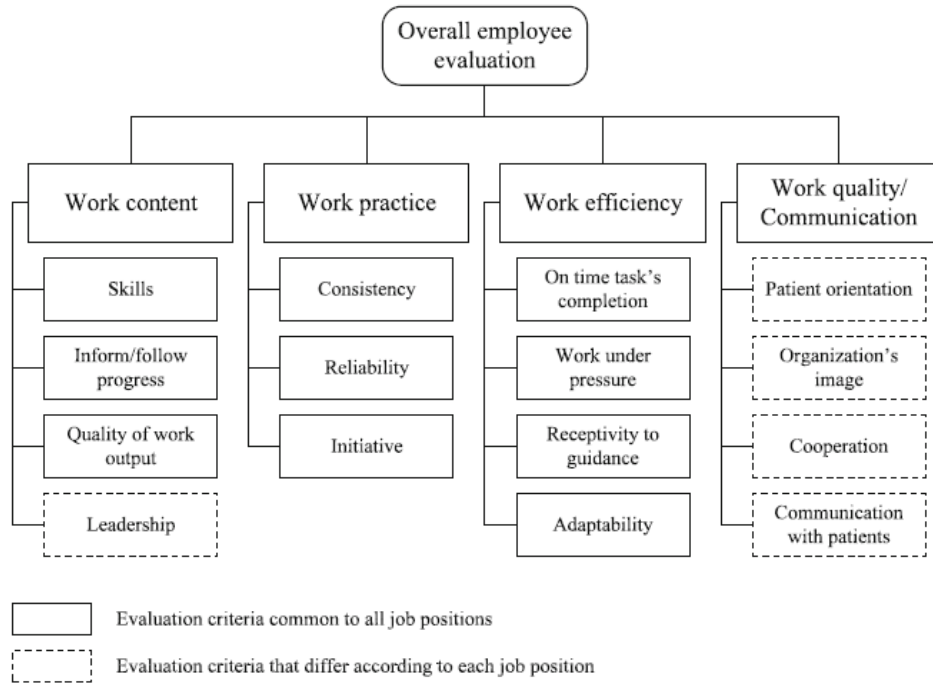
Grigoroudis and Zopounidis (2012) applied a multi-criteria model based on the UTASTAR method that is a regression based approach that adopts the aggregation-

² UTA method (Jacquet-Lagrange and Siskos 1982), which aims at inferring a set of additive value functions from a given ranking on a reference set of actions (alternatives).

disaggregation principles, but takes in consideration alternative criteria preferences. A real-world application for this system applied in a private general hospital in Greece that evaluated 8 different job positions: 1) Financial department (managers), 2) Nursing department (managers), 3) Customer services (CS) department/Secretariat (managers), 4) Financial department (personnel), 5) Nursing department (personnel), 6) Lab personnel, 7) Customer services (CS) department/Secretariat (personnel), and 8) Technical personnel.

In the process of describing the proposed employee evaluation system, Grigoroudis and Zopounidis (2012) presented step-by-step details for the developed model and results of the real-world application. The assessment of the evaluation criteria is also detailed with weights and scores of items. The implementation process maintained direct communication between the evaluator and the employee supervised by the management of the organization to control for any biases. The evaluation criteria illustrated in **Figure (5)** were grouped in four main dimensions taking into account the different job contexts in the healthcare organizations: 1) work content, 2) work practice, 3) work efficiency, and 4) work quality/Communication. However, there are some criteria that differ according to each job position. For example, the attribute of leadership concerns only the heads of the departments.

Figure (5) Employee Evaluation Criteria



In this study, Grigoroudis and Zopounidis (2012), presented a good approach for an evaluation model that accounts for the different job contexts of healthcare organizations with the ability to measure quality of work output based on the criteria of preference to decision makers for the specific job category. However, this system neglected other important issues for employees' productivity such as evaluation of communication and training for improvements.

A holistic approach to develop employee performance measurement system in healthcare organizations was developed by Traberg (2011), in a dissertation composed of five scientific articles submitted to DTU Management Engineering, Technical University of Denmark, in fulfillment of the requirements for acquiring the PhD degree. The design of this approach was based on Management-By-Objectives framework that uses organizational objectives as guidelines for the management of operations.

Although the concept of management-by-objectives seems straightforward, its implementation in healthcare organizations reveals the challenge. As management-by-objective depends on determining which objectives to measure in order to decide about the proper measuring system to achieve the results, the multi-perspectives of healthcare stakeholders brings a significant challenge to decide about priorities of what to measure and what to consider as organizational objectives. For example, standards of quality care from a customer perspective might be different from medical practitioner perspective or authorities view. Traberg (2011), realized that this difficult task is one of the reasons that many healthcare organizations use multi-methods.

To introduce a holistic framework, Traberg (2011), proposed that measurement indicators are designed in relation to a specific purpose in a specific context that provides logical representation of the performance measure. Traberg (2011), adopted the concept of balanced scorecards in designing a detailed process to link each measuring indicator to a specific job position and specific purpose to meet a specific goal. This model incorporated the Analytical Hierarchy Process (AHP) method that allows for subjective and objective assessments of elements (Saaty, 1982; as cited in Traberg, 2011). The research discusses extensively the measurement framework process presenting an example of the appearance of an indicator and its use to assess an objective in **Figure (6)**. Each measurement indicator is extracted from an objective and must have a standard template of indicator information that provides detailed description added to the complete set of indicators to construct the model performance indicators that will be applied in the organization.

Figure (6) Indicator Example

	<i>Description</i>
<i>Indicator name</i>	Waiting List
<i>Purpose</i>	Continuously monitor the maximal waiting time for a non-acute patient, distributed on modalities
<i>Responsible</i>	Head of department
<i>Field of application</i>	Each of four radiology sections of the hospital
<i>Indicator description</i>	Waiting time to the next open examination slot in the booking system for each modality
<i>Displaying guidance</i>	Y-axis: Waiting time in days;; Calendar days 6 month back
<i>Data foundation</i>	Data is collected from RIS (Radiology Information System)
<i>Indicator goal</i>	Waiting time below 20 days, complying with National Treatment assurance (4 weeks)
<i>Timeframe</i>	At all times
<i>Guiding documents</i>	The Danish Quality model (www.ikas.dk) The National Indicator Project (www.nip.dk)
<i>Benchmark</i>	Monthly benchmarked internally between all four locations Bi-annual, waiting time is benchmarked externally between Danish hospitals
<i>References</i>	The Danish Quality Model, Standard 3.1.1- Standard 3.2.1- Standard 3.6.1 - Standard 3.8.1- Standard 3.11.1

This is the distinct feature of this model especially when applied in healthcare where prioritization of functions and activities does not usually follow a specific rationale and may be different from one healthcare organization to the other.

Although, this research demonstrates a valuable contribution to solve performance measurement problems in healthcare, the rating scales were not measured and the scientifically developed model seems very demanding for a huge amount of information and involvement from management in addition to intense training for supervisors and employees on the system. The lengthy application may also hinder its implementation on smaller scale healthcare facilities.

Limitations

Prior to conducting the systematic review, it was expected that the main limitation will be in finding credible relative resources considering that the focus of the

body of knowledge on exploring individual performance appraisal systems in healthcare is still insignificant. The search investigation confirmed this limitation in addition to observing other considerations such as restrictions of some database sites on exporting bulk references. Further, the wide diversity between studies in the quality of outcomes and contexts prevented synthesizing the contents and conducting a systematic comparison between all the results. Lastly, as this study is presented to fulfill the requirements for the Master's degree in Human Relations, involving professional reviewers to establish consistency between the results and confirm validity of findings was not possible.

Discussion

The categorized analysis that explored the studies discussed the problems of measuring performance of medical practitioners and assessed the currently applicable performance measuring systems has considered finding an answer for the first research question: ***RQ1: Are the implemented performance appraisal and measurement systems effective in evaluating the performance of healthcare workers?***

The review demonstrated a general dissatisfaction with the ability of the implemented systems in providing accurate evaluation of the performance of healthcare workers. The 360-Degree / Multisource feedback (MSF) gained more credibility than other systems for its ability to reflect feedback from different perspectives; however, no evidences were presented of the MSF's ability to measure soft skills precisely. Further, the use of the term "healthcare workers" in this research included medical and non-medical categories of workers in healthcare industry, while the reviewed studies focused on measuring the performance of medical practitioners with an obvious gap in

evaluating the performance of non-medical employees who work in healthcare organizations. Each measurement tool has strengths and weaknesses that were presented in Table (4) but there were no evidence if any of the implemented performance appraisal and measurement systems is more effective in evaluating the performance of healthcare workers compared to others. When researchers identify a weakness or error in application, their focus goes to reduce the error in the measurement system but do not give equivalent importance to measure the effectiveness of the tool in the purpose it was designed for. This study could not provide a clear cut answer to the first question in identifying a complete effective system or instrument that can cover all stakeholders' perspectives and provide decision makers with accurate quantitative data about performance of healthcare workers.

The second research question is derived from the core purpose of this research: ***RQ2: What are the key performance indicators for jobs based on soft skills?*** In this regard, several studies such as Decker (1999) and Traberg (2011) discussed the process of developing key performance indicators for healthcare practitioners as part of developing or designing a performance measurement system. Decker (1999) has provided an example for the needed "hidden competencies" that were defined as attitudinal and interpersonal skills that stimulate customer satisfaction, but a link between the hidden competencies and soft skills was not made clear. Robles (2012) presented a clear definition for soft skills in **Figure (7)** illustrates that soft skills are more than interpersonal skills for its inclusive to the natural personal attributes that an individual possess.

Figure (7) Soft skills are more than interpersonal skills

Soft Skills = Interpersonal (People) Skills + Personal (Career) Attributes

Based on this definition, the behavioral based performance indicators that Decker (1999) proposed to measure the “hidden competencies” do not present accurate performance measurement indicators for jobs based on soft skills. In addition, all the studies that discussed performance indicators in this review were concerned with identifying performance indicators that measure hard skills. Soft skills were not mentioned in any of the review studies.

As the purpose of this research was to investigate types of the currently used performance measurement systems in healthcare organizations and assess their efficiency in providing accurate measurement of the performance of healthcare workers in their various job types with specific focus on soft skills, the third research question is formulated successively on the basis of the assumed literature gap in finding answers for the first and second questions. ***RQ3: Can a single performance appraisal system produce a valid and reliable measure for all job categories?***

The term “healthcare workers” in this research, as indicated earlier, covered medical and non-medical categories of workers in healthcare industry. This review revealed that all included studies focused on measuring the performance of medical practitioners with an obvious gap in evaluating the performance of non-medical employees who work in healthcare organizations. None of the reviewed studies also intended to evaluate the performance measurement system for non-medical employees in a healthcare organization. This research illustrated the organizational need to use

multiple performance measures to balance the paramount and conflicting objectives of the diverse stakeholders more aptly.

The failure to provide definite answers to the previous research questions has resulted to an inability of addressing the fourth and fifth questions that are: ***RQ4: How is the performance appraisal output linked to the reward system?*** And ***RQ5: Is there any best practice model for performance measurement in healthcare industry?***

As performance measurement indicators are the focal point for any performance measurement system, it is important to ensure that performance indicators are accurate and reliable. Successful linking of the performance appraisal output to the reward system depends highly on the ability of the measurement system of providing accurate and reliable quantitative data for decision makers. The implemented performance appraisal and measurement systems are using generic performance measures that do not reflect behavioral and contextual performance of all job categories in the healthcare industry. Consequently, the performance system outcomes are mostly subjective and therefore the rewarding system is biased.

To date, the research demonstrated that there is no best practice model to measure performance of all healthcare workers. Although this is a clear cut answer to the fifth research question, development to existing models or design of a new framework was presented in few studies. Generally, these studies were well documented, however, as many factors can affect the performance of employees in healthcare organizations, a development of an error-less employee performance evaluation system appear to be neither easy nor straightforward and requires more investigation.

Conclusion

This systematic review of the literature exploring employee performance appraisal systems in healthcare organizations presented diversification in the applicable methods and revealed absence of a comprehensive approach to performance evaluation in healthcare. There are more differences among the 23 articles reviewed in this research than similarities, even within the articles classified in any single category. Therefore, the intention to compare and synthesize the results in a systematic way could not be precisely achieved.

Although this systematic review could not identify a clear solution to the weakness of applicable employee performance appraisal systems in healthcare organizations, the findings have supported the statement of problem and provided direction for future research that is needed to find answers for some of the research questions.

The implemented performance appraisal and measurement systems presented good examples of the currently applicable systems in healthcare organizations. Several perspectives were covered by using the 360-degree feedback / multisource feedback (MSF) system, although most healthcare applications to the MSF system were focused on measuring only medical practitioners' performance. The complexity and lengthy process of the balanced scorecard created limitation on implementation especially in small and medium healthcare organizations. Another concern about balanced scorecard is the structural demand for a large amount of information to design and customize the process and indicators to a specific strategic goal. Such requirement increases the predicament of implementation. On the other hand, most pay-for-performance

applications in healthcare focused on motivating doctors' performance neglecting the value added by other healthcare workers who might be in a direct contact with patients and impact the organizational image and reputation in offering quality care.

This research confirmed that the wide diversity in job categories of workers in healthcare organizations undermines the effort to use one set of performance measures or one program for all jobs. Performance indicators have to be crafted for each job type to directly measure the behaviors and actions that were performed to fulfill a specific organizational goal. The process must set clear goals and design measurement tools that directly link outcome results to goals. Generic performance measures will not provide meaningful results.

In conclusion, a literature gap exists in covering the problem of accurately measuring the performance of healthcare workers across all their job categories. Further research and insight is needed to explore other possible solutions to this vital problem. As the key objective for every healthcare organization is the provision of reputable quality of care, researchers should give an equal importance in performance measurement to all workers in the healthcare industry where each employee is contributing from a distinctive role to draw the big picture.

Future research needs to explore answers for the following questions: What are the key performance indicators for jobs based on soft skills and behavioral actions in the healthcare industry? Is measuring the outcomes of a particular individual's job a valid and reliable measure of that person's performance? What could be considered a best practice model for performance measurement in healthcare industry?

References³

- Aguinis, H., (2005). Performance management. *Edinburgh Business School, Heriot-Watt University*. Great Britain in 2005. PE-A1-engb 1/2011 (1033). Retrieved from <https://www.ebsglobal.net/documents/course-tasters/english/pdf/h17pe-bk-taster.pdf>
- *Atwater, L. E., Brett, J. F. and Charles, A. C., (2007). Multisource feedback: Lessons learned and implications for practice. *Human Resource Management*, 46(2), 285-307. doi: 10.1002/hrm.20161. Retrieved from <http://onlinelibrary.wiley.com.ezproxy.lib.ou.edu/doi/10.1002/hrm.20161/pdf>
- Buckingham, M. and Goodall, A., (2015). Assessing performance. Reinventing performance management. *Harvard Business Review*. April 2015 Issue. Retrieved from: <https://hbr.org/2015/04/reinventing-performance-management>
- *Carlos, V. S. and Rodrigues, R. G., (2015). Development and validation of a self-reported measure of job performance. *Business Media Dordrecht 2015. Soc Indic Res*, 126:279–307. DOI 10.1007/s11205-015-0883-z
- Carroll, S. J. & Schneier, C. E., (1982). Performance appraisal and review systems: the identification, measurement, and development of performance in organizations. Scott, Foresman and Company (1982). Glenview, Illinois. ISBN: 0-673-16006-8. Retrieved from: <http://www.tamu.edu/faculty/payne/PA/Carroll%20&%20Schneier%201982.pdf>
- *Carson, M., (2006). Saying it like it isn't: The pros and cons of 360-degree feedback. *Business Horizons*, 49(5), 395. doi:10.1016/j.bushor.2006.01.004. Retrieved from: <http://www.sciencedirect.com/science/article/pii/S0007681306000206>
- *Cohen, D. and Rhydderch, M., (2006). Measuring a doctor's performance: personality, health and well-being. *Oxford University Press on behalf of the Society of Occupational Medicine. Occupational Medicine*; 56:438–441. doi:10.1093/occmed/kql076. Retrieved from <http://occmed.oxfordjournals.org/>
- *Decker, P. J., (1999). The hidden competencies of healthcare: why self-esteem, accountability, and professionalism may affect hospital customer satisfaction scores. *Hospital Topics* 77.1 (1999):14.
- DeNisi, A., & Pritchard, R., (2006). Performance Appraisal, Performance Management and Improving Individual Performance: A Motivational Framework. *Management and Organization Review*, 2(2), 253-277.

³ References marked with an asterisk (*) are the studies included in the systematic review.

- *Elwyn, G., Lewis, M., Evans, R., & Hutchings, H., (2005). Using a 'peer assessment questionnaire' in primary medical care. *The British Journal of General Practice: The Journal of the Royal College of General Practitioners*, 55(518), 690-695. Retrieved from <http://www.ncbi.nlm.nih.gov.ezproxy.lib.ou.edu/pmc/articles/PMC1464068/pdf/bjpg55-690.pdf>
- *Emmert, M., Eijkenaar, F., Kemter, H., Esslinger, A. S., Schoffski, O., (2011). Economic evaluation of pay-for-performance in health care: a systematic review. Springer-Verlag 2011. *Eur J Health Econ* (2012) 13:755–767. DOI 10.1007/s10198-011-0329-8
- *Greenfield, J., (2014). Why every practice nurse should have an annual appraisal. *Practice Nurse*, 44(9), 32-36 5p. Retrieved from: <http://web.b.ebscohost.com.ezproxy.lib.ou.edu/ehost/detail/detail?vid=3&sid=43e91e54-7031-4280-9128-b5087093bef8%40sessionmgr120&hid=115&bdata=JnNpdGU9ZWwhvc3QtbGl2ZQ%3d%3d#db=rzh&AN=107831061>
- *Grigoroudis, E., Zopounidis, C., (2012). Developing an employee evaluation management system: the case of a healthcare organization. *Operational Research*, 12(1), 83-106. doi:<http://dx.doi.org/10.1007/s12351-011-0103-9>. Retrieved from: <http://search.proquest.com.ezproxy.lib.ou.edu/docview/927712624?OpenUrlRefId=info:xri/sid:primo&accountid=12964>
- *Helm, C., Holladay, C. L., Tortorella, F. R., and Candio, C., (2007). The performance management system: Applying and evaluating a pay-for-performance initiative. *Journal of Healthcare Management*, 52(1), 49-62; discussion 62-3. Retrieved from <http://search.proquest.com.ezproxy.lib.ou.edu/docview/206739422/fulltextPDF?accountid=12964>
- Kaplan, R. S., (2010). Conceptual foundations of the balanced scorecard - working paper. *Harvard Business School, Harvard University*. Retrieved from <http://www.hbs.edu/faculty/Publication%20Files/10-074.pdf>
- *Kirschner et al., (2012). Design choices made by target users for a pay for performance program in primary care: an action research approach. *BMC Family Practice* 2012, 13:25. Retrieved from <http://www.biomedcentral.com/1471-2296/13/25>
- Kootanaee, H. J., Kootanaee, A. J., Hoseinian, H., & Talari, H. F., (2013). The balanced scorecard, alphabet of the modern management: From concept to implement. *Advances in Management and Applied Economics*, 3(1), 47-59. Retrieved from: <http://search.proquest.com/docview/1312296896?accountid=12964>

- Kyriakopoulos, G., (2012). Half a century of management by objectives (MBO): A review. *African Journal of Business Management* Vol. 6(5), pp. 1772-1786,8 February, 2012. Available online at <http://www.academicjournals.org/AJBM>. DOI: 10.5897/AJBM10.1585. ISSN 1993-8233 ©2012 Academic Journals
- Mercer LLC, (2013). Global Performance Management Survey Report – Executive Summary. Retrieved from: <http://www.mercer.com/content/dam/mercer/attachments/global/Talent/Assess-BrochurePerfMgmt.pdf>
- *Miller, G. and Babiarz, K. S. (2013). Pay-for-performance incentives in low- and middle-income country health programs. NBER Working Paper No. 18932, Revised April 2013. JEL No. H51,I12,O12,O17,Z18
- *Nikpeyma, N., Abed_Saeedi, Z., Azargashb, A., and Alavi_Majd, H., (2014). Problems of clinical nurse performance appraisal system: a qualitative study. Korean Society of Nursing Science. Elsevier, *Asian Nursing Research* 8 (2014) 15e22. Retrieved from <http://www.sciencedirect.com/science/article/pii/S1976131713000650>
- Orr, L.M. and Orr, D. J., (2014). *Eliminating waste in business: run lean, boost profitability*. Berkeley, CA: Apress. ISBN-13 (electronic): 978-1-4302-6089-9
- *Overeem, K., Faber, M. J., Arah, O. A., Elwyn,G., Lombarts, K. M., Wollersheim, H. C., and Groll, R. P., (2007). Doctor performance assessment in daily practise: does it help doctors or not? A systematic review. *Medical Education*, 41(11), 1039-1049 11p. doi:10.1111/j.1365-2923.2007.02897.x. Retrieved from <http://web.a.ebscohost.com.ezproxy.lib.ou.edu/ehost/pdfviewer/pdfviewer?sid=1453d146-654c-462d-a9c5-5ce9295cf084%40sessionmgr4001&vid=1&hid=4201>.
- *Overeem, K., Wollersheim, H. C., Arah, O. A., Cruijsberg, J. K., Grol, R., and Lombarts, K. M., (2012). Evaluation of physicians' professional performance: An iterative development and validation study of multisource feedback instruments. *Bio Med Central (BMC). Health Services Research* 2012. Retrieved from <http://www.biomedcentral.com/1472-6963/12/80>
- *Payne, S. C., Horner, M. T., Boswell, W. R., Schroeder, A. N., and Stine-Cheyne, K. J., (2008). Comparison of online and traditional performance appraisal systems. *Journal of Managerial Psychology*, 24(6), 2009, pp. 526-544, 0268-3946. DOI:10.1108/02683940910974116
- Givan, R. K., (2005). Seeing stars: Human resources performance indicators in the National Health Service. *Personnel Review*, 34(6), 634-647. Retrieved from: <http://www.emeraldinsight.com/doi/full/10.1108/00483480510623439>

- Robles, M. M., (2012). Executive perceptions of the top 10 soft skills needed in today's workplace. *Business Communication Quarterly* XX(X)1–13, <http://www.sagepub.com/journalsPermissions.nav>. DOI:10.1177/1080569912460400, <http://bcq.sagepub.com/460400BCQXXX10.1177/1080569912460400Busi>
- *Rowe, A. K., Savigny, D., Lanata, C. F., and Victora, C. G., (2005). How can we achieve and maintain high-quality performance of health workers in low-resource settings? *The Lancet*; Sep 17-Sep 23, 2005; 366, 9490; ProQuest, pg. 1026.
- *Rynes, S. L., Gerhart, B., and Parks, L., (2005). Personnel psychology: performance evaluation and pay for performance. *Annual Review of Psychology*, 56(1), 571-600. doi:10.1146/annurev.psych.56.091103.070254. Retrieved from <http://web.b.ebscohost.com.ezproxy.lib.ou.edu/ehost/pdfviewer/pdfviewer?sid=39c2840a-5c7c-4fa0-89d9-38130d315c10%40sessionmgr111&vid=1&hid=106>
- Schuster, F., & Kindall, A., (1974). Management by objectives where we stand - a survey of the fortune 500. *Human Resource Management*, 13(1), 8-11. Retrieved from <http://search.proquest.com.ezproxy.lib.ou.edu/docview/223827370?OpenUrlRefId=info:xri/sid:primo&accountid=12964>
- Smither, J.W., (1998). *Performance appraisal: state of the art in practice*. San Francisco: Jossey-Bass.
- Smither, J., & Manuel, L., (2009). *Performance management: Putting research into action (Professional practice series)*. San Francisco: Jossey-Bass.
- Society for Human Resource Management (SHRM) HR Content, (2010). Chapter 8: Performance management and appraisal. Retrieved from http://www.sagepub.com/sites/default/files/upm-binaries/45674_8.pdf
- Society for Human Resource Management (SHRM Foundation), (2000). *Performance Management Survey*. Retrieved from <http://www.shrm.org/research/surveyfindings/documents/performance%20management%20survey.pdf>
- Society for Human Resource Management (SHRM), (2014). *Survey about HR professionals' perception about performance management effectiveness*. Retrieved from www.shrm.org/research/documents/2014-performance-managment.pptx#sthash.BnCAMLRI.dpuf
- Society for Human Resource Management (SHRM), (2015). *Managing Employee Performance*. Retrieved from

<http://www.shrm.org/templatestools/toolkits/pages/managingemployeeperformance.aspx>

- *Stalker, M. Z., Kornblith, A. B., Lewis, P. M., and Parker, R., (1986). Measurement technology applications in performance appraisal. *JONA*. Vol. 16, No. 4. April 1986.
- *Sturman, M. C., Cheramie, R. A., and Cashen, L. H., (2005). The impact of job complexity and performance measurement on the temporal consistency, stability, and test–retest reliability of employee job performance ratings. *Journal of Applied Psychology*, 90(2), 269-283. doi:10. 1037/0021-9010.90. 2.269. Retrieved from <http://web.a.ebscohost.com.ezproxy.lib.ou.edu/ehost/pdfviewer/pdfviewer?sid=95bc6042-1f2e-4d54-a555-03e779f6f0cc%40sessionmgr4005&vid=1&hid=4201>
- *Taylor, S., (2011). Assess pros and cons of 360-degree performance appraisal. The Society for Human Resources Management (SHRM). Retrieved from <http://www.shrm.org/hrdisciplines/employeerelations/articles/pages/360degreepformance.aspx>
- *Traberg, A., (2011). Management by objectives in healthcare. Technical University of Denmark, DTU Management Engineering. Denmark. ISBN: 978-87-92706-07-2. Retrieved from http://orbit.dtu.dk/fedora/objects/orbit:87325/datastreams/file_5881039/content
- Traberg, A., Jacobsen, P., & Monique Duthiers, N. (2014). Advancing the use of performance evaluation in health care. *Journal of Health Organization and Management*, 28(3), 422-36. Retrieved from <http://search.proquest.com/docview/1651290751?accountid=12964> Also Retrieved from <http://search.proquest.com.ezproxy.lib.ou.edu/docview/1651290751?OpenUrlRefId=info:xri/sid:primo&accountid=12964>
- Watson Wyatt and WorldatWork, (2005/2006). Strategic Rewards Survey. Performance Management Programs Need Improvement. Retrieved from: <https://www.worldatwork.org/waw/adimLink?id=17179>
- *Wiese, D. S., & Buckley, M. R., (1998). The evolution of the performance appraisal process. *Journal of Management History*, 4(3), 233. Retrieved from <http://search.proquest.com/docview/211064856?accountid=12964>
- Zelman, W. N., Pink, G. H., & Matthias C. B., (2003). Use of the balanced scorecard in health care. *Journal of Health Care Finance*, 2003: 29, 4. Retrieved from <https://jsofian.files.wordpress.com/2006/12/use-of-bsc-in-health-care.pdf>

Appendices

Appendix (1)

Excel Data Sheet of Studies selected in the Systematic Review

Sr #	Author(s)	Year	Title	Purpose of the Study	Measure(s)
1	Debbie Cohen & Melody Rhydderch	2006	Measuring a doctor's performance: personality, health and well-being. Published by Oxford University Press on behalf of the Society of Occupational Medicine. Occupational Medicine 2006;56:438-441. doi:10.1093/occmed/kql076	Doctors, trusts and educational bodies have become increasingly accountable for their actions and a clear need has arisen to recognize underperformance earlier and in a structured and transparent manner	Case studies
2	Kirsten Kirschner, Jozé Braspenning, JE Annelies Jacobs and Richard Grol	2012	Design choices made by target users for a pay-for-performance program in primary care: an action research approach.	to design a P4P program using a bottom-up procedure, in which the different options for performance measurement, appraisal and reimbursement were discussed by the target users in a systematic consensus procedure .	Pay for performance (P4P)
3	Grant Miller & Kimberly Singer Babiarz	2013	PAY-FOR-PERFORMANCE INCENTIVES IN LOW- AND MIDDLE-INCOME COUNTRY HEALTH PROGRAMS	outlines key conceptual issues in the design of pay-for-performance contracts and summarizes existing empirical evidence related to each. In doing so, it focuses on four key conceptual issues: (1) What to reward, (2) Who to reward, (3) How to reward, and (4) What perverse incentives might performance rewards create.	Pay for performance (P4P)
4	Nasrin Nikpeyma, Zhila Abed_Saeedi, Eznollah Azargashb, Hamid Alavi_Majid	2014	Problems of Clinical Nurse Performance Appraisal System: A Qualitative Study	explore problems of clinical nurse performance appraisal system.	Assessment to the process of performance appraisal
5	Karljin Overeem, Hub C Wollersheim, Onyebuchi A Arah, Juliette K Cruisberg, Richard PTM Grol and Kiki MIMH Lombarts	2012	Evaluation of physicians' professional performance: An iterative development and validation study of multisource feedback instruments	this paper addresses three core aims, namely: (1) the initial psychometric properties of three new instruments based on existing MSF instruments and the influence of potential sociodemographic variables, (2) the correlation between physician self-evaluation and other raters' evaluations, (3) the number of evaluations needed per physician for reliable assessments.	Multisource Feedback (MSF). The MSF system in the Netherlands consists of feedback from physician colleagues (peers), co-workers and patients.
6	Stephanie C. Payne, Margaret T. Homer, Wendy R. Boswell, Amber N. Schroeder, and Kelleen J. Stine-Cheyne	2008	Comparison of online and traditional performance appraisal systems	The purpose of this paper is to compare employee reactions to the use of an online performance appraisal (PA) system to the traditional paper-and-pencil (P&P) approach.	Traditional performance appraisal questionnaire
7	Vera Silva Carlos and Ricardo Gouveia Rodrigues	2015	Development and Validation of a Self-Reported Measure of Job Performance	Considering the fact that most individual performance measures are developed in order to be applied in particular job-related contexts or cultures, our goal is to develop a job performance measure that might be applicable across jobs and cultures.	a new scale was developed to assess current measures
8	Alexander K Rowe, don de Savigny, Claudio F Lanata, Cesar G Victoria	2005	How can we achieve and maintain high-quality performance of health workers in low-resource settings?	Discuss performance of health workers and how to improve it. What are the determinants to their performance.	Review of research
9	Decker, Phillip J.	1999	The Hidden Competencies of Healthcare: Why Self-Esteem, Accountability, and Professionalism may Affect Hospital Customer Satisfaction Scores	Discuss competence model in healthcare and the need for customer focused competence model rather than the Licensing commission	HR competencies

Excel Data Sheet of Studies Added to the Systematic Review

Sr #	Author(s)	Year	Title	Purpose of the Study	Measure(s)
1	Steve Taylor	2011	Assess Pros and Cons of 360 Degree Performance Appraisal - SHRM	to consider carefully whether the use of 360 degree feedback is appropriate for performance evaluation	360 Degree Performance Appraisal
2	Russell Mannion	2014	Take the money and run: the challenges of designing and evaluating financial incentives in healthcare; Comment on "Paying for performance in healthcare organisations"	Discuss problems facing <u>healthcare</u> systems in controlling costs while improving quality and performance through P4P system	P4P (Pay for performance)
3	E. Grigoroudis and C. Zopounidis	2012	Developing an employee evaluation management system: the case of a healthcare organization	To present the development of an employee evaluation system in a <u>healthcare</u> organization	multicriteria analysis that considers the complexity of the different job profiles
4	Karlijn Overeem, Marjan J Faber, Onyebuchi A Arah, Glyn Elwyn, Kiki M J M H Lombarts, Hub C Wollersheim & Richard P T M Grol	2007	Doctor performance assessment in daily practise: does it help doctors or not? A systematic review	to systematically evaluate the feasibility of methods, the psychometric properties of instruments that are especially important for summative assessments, and the effectiveness of methods serving formative assessments used in routine practise to assess the performance of individual <u>doctors</u>	observed 6 different methods of evaluating performance: simulated patients; video observation; direct observation; peer assessment; audit of medical records, and portfolio or appraisal.
5	Andreas Traberg	2011	Management-By-Objectives in Healthcare	to design a holistic Management-By-Objectives framework that can enable managers and operational personnel to assess performance in relation to the organizational expectations (in <u>healthcare</u>)	Developed
6	LEANNE E. ATWATER, JOAN F. BRETT, AND ATIRA CHERISE CHARLES	2007	MULTISOURCE FEEDBACK: LESSONS LEARNED AND IMPLICATIONS FOR PRACTICE	to outline recent studies on MSF in order to inform practice and increase the likelihood that more leaders and organizations will benefit from this developmental process.	360 Degrees / Multisource feedback
7	Sara L. Rynes, Barry Gerhart, and Laura Parks	2004	PERSONNEL PSYCHOLOGY: Performance Evaluation and Pay for Performance	briefly trace the origins of the general separation of Performance Evaluation (PE) research from Pay-for-performance (PFP) research in psychology. Then review recent research on the relationship between PE and performance improvement, particularly with respect to multisource or 360-degree evaluation. Then turn to research on various PFP systems, such as merit pay and individual and group incentives.	Pay for Performance
8	Mary Carson	2006	Saying it like it isn't: The pros and cons of 360-degree feedback	This small case study showcases key research findings on best practices for using the 360-design feedback process, especially regarding intended use of feedback, aligning performance measures with organizational goals, and selecting and training participants.	360 degrees / Multisource feedback
9	Michael C. Sturman, Robin A. Cheramie, and Luke H. Cashen	2005	The Impact of Job Complexity and Performance Measurement on the Temporal Consistency, Stability, and Test-Retest Reliability of Employee Job Performance Ratings	the authors define and distinguish between the concepts of temporal consistency, stability, and test-retest reliability when considering individual job performance ratings over time. Furthermore, the authors examine measurement type (i.e., subjective and objective measures) and job complexity in relation to temporal consistency, stability, and test-retest reliability.	Performance measurements and ratings
10	Corey Helm, Courtney L. Holladay, and Frank R. Tortorella	2007	The Performance management System: Applying and Evaluating a Pay-for-Performance Initiative	To determine the effectiveness in aligning individual performance goals to institutional goals and linking performance to rewards (i.e. distinguishing high performers from low performers)	Pay for Performance
11	G Elwyn, M Lewis, R Evans and H Hutchings	2005	Using a 'peer assessment questionnaire' in <u>primary medical care</u>	To test the feasibility of using a peer assessment questionnaire in a primary care setting, and consider the related issues of validity and reliability and compare the results to previous studies.	Peer assessment
12	Greenfield, Jenny	2015	Why every <u>practice nurse</u> should have an annual appraisal	look at the process of appraisal from both the appraisee's and the appraiser's perspectives, to allow you to have a clear understanding of what appraisal is, what it is not, and how it should be conducted.	Annual appraisal (interview)
13	Wiese, Danielle S; Buckley, M Ronald	2016	The evolution of the performance appraisal process	The purpose of this paper is to outline the historical evolution of the performance appraisal process. The goal is to synthesize the progress (or lack thereof) which has been made in this process, while critically analyzing collective contributions to increasing the effectiveness with which behavior is both observed and evaluated.	performance appraisal process
14	Martha Z. Stalker, Alice B. Kornblith, Patricia Mazzola Lewis, and Roger Parker	1986	Measurement Technology Applications in Performance Appraisal	This article raises issues that relate to changing the performance appraisal system and gives a clear, step-by-step description of how computerized, statistical procedures were applied in improving <u>nursing performance appraisal</u> .	Performance appraisal process and indicators for nurses