ETLA Working Papers

No. 18

24 September 2013

Antti Kauhanen
Julia Salmi
Paulus Torkki

PERFORMANCE MEASUREMENT IN HEALTHCARE INCENTIVE PLANS

Suggested citation: Kauhanen, Antti, Salmi, Julia & Torkki, Paulus (24.9.2013). "Performance measurement in healthcare incentive plans". ETLA Working Papers No 18. http://pub.etla.fi/ETLA-Working-Papers-18.pdf

ETLA ELINKEINOELÄMÄN TUTKIMUSLAITOS
THE RESEARCH INSTITUTE OF THE FINNISH ECONOMY

ISSN-L 2323–2420
ISSN 2323–2420 (print)
ISSN 2323–2439 (online)

Performance measurement in healthcare incentive plans*

Antti Kauhanen†

Julia Salmi‡

Paulus Torkki§

Abstract

By using quantitative survey data and conducting a case study, we examine performance measurement of incentive plans in Finnish private sector health care organizations. We find that the performance measures used in the incentive plans are in line with recent economic theories of performance measurement. The findings from the case study emphasize the importance of choosing appropriate performance measures and designing the pay package as a whole. Inadequate performance measurement leads to incentive plans that do not help organizations reach their goals.

* We thank Tekes, the Finnish Funding Agency for Technology and Innovation for funding our study, and Essi Kujansuu for providing excellent research assistance.

† Research Institute of the Finnish Economy, Lönnrotinkatu 4 B 00120 Helsinki, Finland. E-mail antti.kauhanen@etla.fi, Tel. +358 9 6099 0244.

‡ Research Institute of the Finnish Economy. E-mail: julia.salmi@etla.fi

§ HEMA-Institute, BIT Research Centre, Aalto University, Finland. E-mail: paulus.torkki@aalto.fi

Introduction

The resources used by and the costs of the healthcare sector have risen over the past decades. In many Western countries, the money spent on publicly funded healthcare has risen faster than the level of available funding. The pressure to improve health while spending less money has forced healthcare organizations to find new ways of arranging and managing their services.

Incentive pay, or pay for performance, has been accepted as a promising way to improve performance of healthcare delivery (Institute of Medicine, 2007). The key question is whether incentive pay can effectively improve healthcare delivery costs. Many scholars have argued that incentive pay may not actually be beneficial in this respect (e.g. Golden & Sloan, 2008).

In this article, we discuss performance measurement in healthcare incentive plans. The success of incentive pay plans depends heavily on the quality of performance measurement. We emphasize the role that the goals of the organization have on the choice of performance measures and discuss why different types of plans are feasible for different organizations. We present evidence from a survey on performance measurement in private sector incentive plans and discuss a case study from the perspective of cost effectiveness. This article complements the existing literature, which deals mainly with the healthcare systems of the US and the UK. The Finnish system is organized quite differently, and the differences are reflected in the incentive systems. Additionally, following Golden and Sloan (2008), we focus on physicians and nurses, not on groups or hospitals.

Performance measurement in incentive plans

The most important, and the most difficult, part of any incentive pay plan is performance measurement (Gibbs, Merchant, Van der Stede, & Vargus, 2009). An ideal performance measure captures the impact of an employee's actions on the goals of the organization and nothing else. In practice, such measures do not exist.

Distortions

Performance measures may cause *distortions* when they do not adequately capture the impacts of all of the employee's actions (Baker, 2002). Typically, distortions arise when performance measures gauge only a part of an employee's actions (Holmström & Milgrom, 1991). In this case, the employee typically shifts efforts towards the actions that are rewarded. This does not, of course, mean that the

other actions are completely neglected but that the weight they receive is insufficient from the perspective of the organization.

A simple example can be found between the tasks of record keeping and physician services. There are many studies on performance-related pay (PRP) schemes that come with increased record-keeping demands. Such plans may lead to increased reporting at the expense of actual services. For example, Fairbrother, Hanson, Friedman, and Butts (1999) suspect that improved record-keeping rather than increased physician efforts to vaccinate patients explains the observed increase in immunization rates. Mullen, Frank, and Rosenthal (2010) report that although both dimensions were being rewarded, efforts shifted from tasks that required the time and effort of doctors toward tasks that relied more on identification and scheduling because the latter required relatively less effort. Doran et al. (2011) found that as a consequence of PRP, physicians performed more tasks for which pay was conditional and fewer tasks for which pay was unrelated. Distortions, however, can also arise when performance measurement focuses on what is easily measurable as these measures may not correspond well to the goals of the organization.

In the healthcare setting, distortions arise easily. The result of these distortions may be because the goal of the organization is difficult to define or because measuring the impact of an employee's actions on the goal is extremely difficult. In this case, the ownership of the organization plays a major role. For example, in publicly funded organizations, the goal may be to provide high-quality services to residents of a certain area, or the goal may be to ensure a certain level of health for the population being served. In many healthcare delivery systems, there are problems with reaching these goals. These problems include the overprovision of nonessential care, the overutilization of diagnostics, inadequate access to and quality of care, and insufficient or potentially harmful care (Golden and Sloan 2008, pp. 289). The key problem is that the ultimate goal (e.g., adequate access and quality of care) is difficult to measure reliably. A particular problem with the measurement of quality is that because quality can only be revealed over time, the time span to measure quality should be sufficiently long.

In privately owned organizations where customers pay for the services themselves, the goal may be to maximize profits. These goals have very different implications for incentive plans. In this case, even though the problems of distortion are less severe, they still exist.

Risk

Another important feature of performance measures is *risk*. A performance measure is risky if it captures other factors in addition to the employee's own actions. While virtually all performance measures contain some risk, a useful distinction is between uncontrollable and controllable risks (Gibbs et al., 2009). Uncontrollable risk corresponds to the classical concept of risk, and we would like to exclude these factors from the performance measure because the employees will have to be compensated for carrying such income risk. Controllable risk refers to random events to which the employee can foresee and react to some extent. These are the types of risks we include in the performance measure because we typically want the employees to react to such events.

The key problem in performance measurement is that there is a trade-off between risk and distortion (Baker, 2002). Broader measures, which capture more of the employee's actions, cause fewer distortions but are also more risky as they tend to include more uncontrollable factors (Gibbs et al., 2009). An example of a broader measure is firm value in a publicly traded firm. While this measure captures the impact of all of the employee's actions on the goal of the organization, the majority of the factors are completely outside the control of the employee. Narrower measures, on the other hand, are less risky (i.e., include fewer uncontrollable risks) but may cause more distortions as they may miss some of the employee's actions.

A typical way to balance this trade-off is to combine broader and narrower factors into an incentive system. In doing so, job design plays a significant role. If the job entails considerable discretion or multiple tasks or if controllable risk is important, then broader measures will be accorded more weight (Gibbs et al., 2009; Raith, 2008) because distortions are more important than risk in this case. However, if the job is very narrow in the sense that there is little discretion and only a few tasks and if controllable risk plays a small role, then narrow measures do not cause significant distortions and instead help decrease uncontrollable risk.

The trade-off between risk and distortion is very different in public and private healthcare organizations. This is because the goals of the two organizations are different. In the public sector, the goals of the organization are more difficult to measure or the measures include a significant degree of uncontrollable risk. Narrower measures, however, may cause distortions. For example, overprovision may be caused by fee-for-service type of incentive plans. The problem is that the performance measure (number of services provided) is too narrow compared to the goal of the organization. In the private sector, however, where the customer pays for healthcare, overprovision is not a major concern.

Manipulation

Many performance measures may also be manipulated, which means that employees take actions that improve the performance measure but have no impact or have a negative impact on the goals of the organization. Reliance on self-reported data also opens up the possibility of up-grading patient classifications to obtain increased reimbursements. The system can also be "gamed" by choosing less sick patients (see e.g. Shen, 2003). Performance, as measured, is improved but not necessarily in the way the decision makers would have expected.

Level of pay and intensity of incentives

The level of pay in incentive systems depends on the base pay and the incentive payments. The amount of the incentive payments (or the intensity of incentives) depends on the properties of the performance measures (Baker, 2002; Holmström & Milgrom, 1991; Prendergast, 1999; Raith, 2008). Problems in performance measurement decrease the intensity of incentives, so for example if there is worry about distortions or uncontrollable risk, optimal incentive intensity is reduced. In practice, this may mean that incentive pay plays only a small role in total compensation or that it is not used at all. If the problems in performance measurement are larger in the private sector, this suggests that incentive pay should play a reduced role in the public sector.

Incentives should also be more intensive when the employees are more responsive to them (e.g. Holmström & Milgrom, 1991). This means that if it is easy for the employees to increase their effort to affect the performance measure, stronger incentives should be used.

Base pay, however, should be used to set the expected level of pay to a level that is in line with the market level of pay when taking into account the risk in the performance measures and the level of effort required. For example, if the performance measures are riskier than the ones used by other competing employers and all other things are equal, the base pay should be higher to compensate for the risk.

Effectively, the base pay determines how the surplus from the employment relationship is divided between the parties: the larger the base wage, the larger the share of the surplus captured by the employee. If the base wage is set in negotiations between the employer and the employee, the division reflects their relative bargaining powers.

Typically, when incentive pay is introduced, the base pay remains unchanged, which means that the expected value of the pay package increases. If the performance measures are not appropriate or if the intensity of the incentive is too high, this may indicate that the introduction of the incentive pay plan increases costs by more than it increases output (however that is measured). Indeed, Golden and Sloan (2008, p. 300) argue that incentive systems may reward for changes that would have happened in any case.

Thus, when implementing an incentive system, the whole package (base pay and incentive payments) should be designed simultaneously. In practice, the problem is that changing the base pay may be challenging, especially if the wages are governed by collective agreements, as is the case in the Finnish public sector.

Empirical studies

There has been a surge in the number of quality, or outcome targeting, PRP schemes despite the lack of a widespread agreement on whether PRP can be used as an instrument to improve the outcomes or quality of healthcare services, in particular, without further increasing the spending on healthcare. Yet, this is exactly what PRP is supposed to achieve.

Evidence on the efficacy of performance-related pay is mixed largely because PRP is such an umbrella of concepts[5]. PRP has been found to significantly increase physician activity (see e.g. Conrad et al.,

[5] Many studies use the term “pay for performance” or P4P, but here we use performance related pay.

2002; Gaynor & Pauly, 1990; Krasnik et al., 1990; Shafrin, 2010) when capitation- or salary-based compensation systems are replaced with fee-for-service (FFS) systems. While these studies show that incentives matter, it is not clear whether the increase in physician activity is aligned with the goals of the organizations.

A common reason for replacing capitation- or salary-based compensation systems with FFS was the perception that healthcare services were being underprovided. Today, the problem with many FFS funding models is the observed overprovision of services. Resources are being wasted on unnecessary or low-value procedures or on expensive treatments that have low-cost alternatives. Most of the available PRP systems do not encourage cost-aware behavior but rather compensate for demanding and expensive operations (Golden & Sloan, 2008). Thus, these systems tend to generate distortions as they encourage actions that are not necessarily aligned with the goals of the organization.

The impact of PRP has been estimated to be smaller, insignificant, inconclusive, and sometimes even harmful (e.g. Dranove, Kessler, McClellan, & Satterthwaite, 2003; Lee et al., 2012; Mullen et al., 2010) when more complex measures such as quality or efficiency are targeted. Estimating efficiency and quality is difficult because they are the result of complementary and conflicting factors between which tradeoffs happen. Hussey et al. (2009) studied health efficiency measures in the published literature and found that 97.4 % of the measures did not account for changes in quality. The danger is that these “efficiency measures” reflect costs rather than efficiency. The study also found that reliability tests on measures were rare, as only 2.3 % of the measures had reportedly gone through a validity test, while weaker sensitivity analyses were performed on 25.3 % of the measures. Petersen, Woodard, Urech, Daw, and Sookanan (2006) and Conrad and Perry (2009) emphasize the importance of reliability analyses in understanding the connection between effort and the measurements collected. For example, it is important to know whether it is easier to improve quality than it is to “game” the performance measure. Likewise, it is valuable to know the timeframe in which the benefits can be observed. Researchers have primarily focused on acute conditions to overcome the delay problem.

The generalizability of results is generally quite weak due to the lack of control groups (as pointed out in several literature reviews, among others,Christianson, Leatherman, & Sutherland, 2007; Van Herck

et al., 2010). For example, exogenous time trends have been found to explain effects that were first assigned to the introduction of PRP (Grant, 2009; Lee et al., 2012). Overall, variations in the results are explained by heterogeneous methodologies and focuses.

Cost-effectiveness of the incentive systems is another issue. A paper by Emmert, Eijkenaar, Kemter, Esslinger, and Schöffski (2012) provides a good review of recent economic evaluations of PRP programs in healthcare. Although the majority of the studies reviewed indicated that the program in question was cost-effective and/or improved healthcare efficiency, most of the analyses were narrow in that they evaluated changes only in the targeted and rewarded categories. Fleetcroft and Cookson (2006) and Walker et al. (2010) studied how the sizes of the rewards within a multi-target PRP program were related to the economic value of the improvement required to receive a reward. Fleetcroft and Cookson (2006) find no relationship between health outcomes and pay, while Walker et al. (2010) find that in many cases even if the health gains are modest, the system is cost effective. However, that direct benefits exceed direct costs is not sufficient (and not even a necessary) condition in justifying PRP. Firstly, PRP should also be the best available policy, i.e., that similar results cannot be achieved by other, perhaps less expensive means. Secondly, the distortion effects or net externalities must not outweigh the net benefits.

These results are concerns that are reflected in wider literature, not just in the literature regarding healthcare. Studies have shown that incentive pay increases productivity in various settings (for a survey, see Lazear & Oyer, 2012). However, firms are not ultimately interested in productivity, but on the effect on profitability or other goals, and research has shown that incentive pay may increase productivity but decrease profitability (Freeman & Kleiner, 2005) and that employees may react to incentives in undesirable ways (Asch, 1990; Courty & Marschke, 2004).

Materials and Methods

Setting

The study consists of two parts. In part 1, a description of the incentive systems in Finnish private sector healthcare is provided, and in part 2, a case study of the effects of incentive system implementation on cost-efficiency is presented.

Register-based Study

We use two data sources to obtain a representative picture of PRP systems in the Finnish private healthcare sector. Both data sets are derived from the Confederation of Finnish Industries (EK), which is the central organization of employer associations. Member firms of the EK represent over 70 % of the Finnish GDP, and member companies employ over 950,000 employees. First, we use the EK's compensation systems survey to develop an outlook on the PRP systems used in the Finnish private healthcare sector in 2007[6].These surveys have been conducted every third year beginning with 2005, and they provide a good overview of the PRP system in the Finnish private sector. Second, we use the EK's wage statistics to provide us with information about the PRP component of employees' wages over time (2007-2011). The wage statistics are collected every year, and participation is mandatory for member firms[7]. These data are at the individual level and provide detailed and reliable information about wages. This enables us to more closely examine the differences in PRP between nurses and doctors.

Case Orthopedic hospital

Healthcare is not an industry with a single production concept or business model, but rather, it consists of a cluster of operating logics or modes (Lillrank, Groop, & Malmström, 2010). In this case study, we focus on elective operations. These are procedures where demand is pre-selected and sorted through a referral system, where diagnostics and corresponding procedures are reasonably precise, where production can be planned and scheduled in advance, and where outputs can be counted and evaluated against set quality criteria. Accordingly, quality-adjusted productivity is a relevant KPI and a basis for compensation.

The case hospital, located in Finland, is a third sector, not-for-profit hospital that concentrates mainly on orthopedic operations. The hospital annually performs between 2,000 and 2,500 elective operations. The demand consists of both referrals from public hospitals and patients who either have voluntary insurance or pay for the services out-of-pocket.

[6] For more details about sampling and content of the survey, see Kauhanen and Napari (2012b)
[7] For a more detailed description of the data, see, e.g., Kauhanen and Napari (2012a)

The incentive system of the hospital was investigated by interviewing the management of the hospital, which included the chief executive officer, chief surgeons and nurses. The study focused on personnel groups involved in direct patient care, that is, doctors and nurses. In addition to the interviews, the daily output data and salary data of surgeons and nurses were collected from hospital information systems for the years 2003, 2007 and 2010. The data included surgeons, who provided full-time, in-patient care during the entire measurement period.

The incentive system was implemented in 2005 to increase volume and capacity utilization rates in the hospital. The surgeons received 1 to 4 % of the standardized procedure-specific price of the operation. If the patient was referred from a public institution, the percentage was lower, but if the patient was paying out-of-pocket, the higher percentage was applied. The base wage for the surgeons remained unchanged when the incentive system was implemented. Nurses did not receive incentives based on output or productivity.

The performance measure is the number of operations performed. From the perspective of the surgeons, this is somewhat risky as they cannot control the demand for operations. The possible distortions depend on the goal of the organization. At that point in time (2005), the hospital wanted to increase the number of operations, thus this change in the plan helped to communicate that intent. However, the plan is risky for the hospital as it will only decrease costs if the number of operations exceeds the baseline.

The effects of the changes in the incentive system were studied by conducting a before-and-after analysis in productivity using surgical procedures as an output and personnel costs as an input. Both the productivity changes of the operating unit and the productivity changes in in-hospital care were assessed. We followed the outputs of seven full-time surgeons who stayed in the hospital for the whole period from 2003 to 2007. The other resources were assessed as FTEs.

Results

Description of the Finnish Healthcare system: Register-based study

Because the public sector is the predominant healthcare provider in Finland (75 % of the healthcare sector), the role of the private sector (25 %) is largely complementary. Public healthcare is financed entirely through taxes, with the exception of small visiting fees while private healthcare services are financed partially through taxes via national health insurance (NHI). However, only certain services are covered by the NHI, and those are reimbursed using fixed tariffs. Private healthcare services can and usually are priced freely above the reimbursement level. Hence, approximately one-third of all private healthcare expenditures are reimbursed by the NHI, and the remaining difference is paid by voluntary private health insurances and by the patients themselves (out-of-pocket)[8] (Vuorenkoski, Mladovsky, & Mossialos, 2008).

In the public sector, most employees have fixed salaries and incentive systems based on productivity are rare. In the private sector, the salary models of doctors vary from fixed salaries to full output-based salaries for doctors. The nurses, however, usually have fixed salaries in both the public and the private sectors.

The differentiated market for private healthcare, although subsidized, is still reflective of supply and demand or cost and benefit. In the private healthcare market, the level of service (perceived by the customer) has particular value on its own, which is why the market exists in the first place as there is always the option of (practically) free public healthcare.[9] Thus, cost-effectiveness with respect to health services is not a main concern from the providers' perspective, and therefore, "overproviding" would simply translate to more business. A particularly interesting feature in the Finnish private healthcare sector is that most PRP schemes are implemented by the firm itself rather than by a third-party financer. Thus, the purposes and goals of PRP are different as well. PRP by private companies is used to incentivize employees to put forth greater effort or to offer competitive wages. The norm in

[8] There is also occupational health care that is often provided by the private sector and covered entirely by the public sector and the employer together. It does not, however, overly complicate the principal-agent problem.

[9] The primary reasons for using private health care are shorter waiting times, freedom to choose the provider/physician, and perception of better quality. Also influential is that referrals are not required to see a specialist, etc. (Vuorenkoski et al., 2008)

the published literature is that a PRP scheme is implemented by a financier and that the goal is to enhance quality or cost-efficiency of healthcare services. This paper provides a complementary view on performance-related-pay in healthcare.

Prevalence of PRP

Panel A of Table 1 depicts the percentages of the nursing staff and physicians in employment who received performance-related-pay between 2007 and 2011. In 2007, 12 % of the nursing staff and 14 % of the physicians had a performance-related pay component. Since then, the percentage of physicians in the PRP system has increased to approximately 20 %, while the percentage of the nursing staff has stabilized at approximately 13 %. The PRP can be both performance-related-pay and profit sharing. Here it has to be noted that many doctors work as entrepreneurs in hospitals. Their whole income is based on revenue they generate and thus can be considered as performance-related pay.

Table 1 Prevalence of PRP

Panel A: Prevalence of PRP among nursing staff and doctors, %

	2007	2008	2009	2010	2011
Nurses	12.23	18	11.31	13.25	13.15
N	6792	6584	6915	6869	7477
Doctors	14.14	15.86	20.46	23.65	20.14
N	665	643	479	554	591

Panel B: Prevalence of PRP in 2007 in Finnish private sector, %

	Firms using PRP		Employees in a firm	
Private healthcare sector	2007	N	2007	N
Lower white-collar	28	155	49	14
Upper white-collar	27	155	44	17
Private sector				
Lower white-collar	69	1730	85	768
Upper white-collar	75	1730	84	837

In panel B, we have the company-based data from the EK compensation survey for the year 2007. When compared with the private sector in general, it becomes clear that performance-related systems in healthcare are much rarer than they are in the private sector. Of the private healthcare

companies, 27.3 percent had PRP schemes for their upper white-collar workers, the employee group that includes physicians, and 27.7 per cent had PRP schemes for lower white-collar workers, the group that includes the majority of the nursing staff. Slightly less than half of the employees in each group participated in the plan (49 % of lower and 44 % of upper white-collar workers). These figures are much lower than they are in the private sector. On average, 69 per cent of private firms had PRP schemes for lower white-collar workers and 75 per cent had PRP plans for upper white-collar workers. The respective average inclusion rates were 85 % and 84 %.

These results suggest either that the problems in performance measurement in the healthcare sector are more severe than in the private sector or that relying on the intrinsic motivation of the employees produces sufficiently good outcomes[10].

As accounts on the prevalence of PRP plans in healthcare are rare, finding suitable comparisons has not been fruitful. However, the impression from the literature is that the prevalence of plans has been increasing throughout most of the 2000s. For example, in the US, the number of PRP schemes increased until 2008, after which the financial crisis and the new US healthcare legislation (HITECH Act & PPACA) most likely contributed to the decline in 2010 (Med-Vantage, 2011).

Intensity of PRP

In Table 2, statistics regarding the size of PRP relative to the base salary are shown. The wage data represented in Panel A reveal that for the nursing staff, PRP was approximately 2.5 % of the base salary in 2007 and that physicians' PRP was closer to 3 % of the base salary. However, since 2007, the nursing staff's percentage has decreased somewhat while the physicians' PRP increased to almost 5 % in 2008, just before the financial crisis. Furthermore, the average physician's PRP remained above 3 % throughout the financial crisis, which undoubtedly had a negative effect on the figures.

[10] Here we also have to take into account that some doctors work as entrepreneurs.

Table 2 The intensity of incentives

Panel A: The size of PRP relative to basic salary, %

	2007	2008	2009	2010	2011
Nurses	2.51	2.5	2.45	2.32	2.06
N	831	1185	782	910	983
Doctors	2.74	4.83	4.16	3.16	3.26
N	94	102	98	131	119

Panel B: The size of PRP relative to the basic salary in the private sector, %

	PRP relative to base pay	
Private healthcare sector	2007	N
Lower white-collar	2.75	7
Upper white-collar	5.01	8
Private sector		
Lower white-collar	5.44	610
Upper white-collar	8.66	639

From Panel B, it can be deduced that the amount of PRP relative to basic earnings for nurses, physicians and the private healthcare sector in general, was much lower in 2007 than in the private sector in general. The company-based dataset reveals that PRP on average added 2.75 % to the base salary of a lower white-collar healthcare worker, while for upper white-collar workers participating in the PRP scheme, the performance-related component was approximately 5.01 % of the basic salary in the private healthcare sector[11]. For the entire private sector, the respective figures where 5.44 % and 8.66 %.

These results again suggest that, on average, the problems with performance measurement in the healthcare sector are more severe than the problems in the private sector. Theory implies that the more risky, distorted or prone to manipulation the performance measures are, the lower the intensity of the incentives.

[11] The discrepancy between the numbers in Panels A (doctors) and B (upper white-collar workers) is due to the fact that typically the incentive payments are paid in March of the following year. Thus, in panel A, the figures in the column 2008 refer, for the most part, to payments made based on performance in 2007.

In the US, the size of the incentive pay relative to basic compensation is higher by approximately 8 percentage points (Med-Vantage, 2011). However, this figure should be interpreted with caution as we do not have information from the physicians' perspective. It is likely that a physician's earnings come from several sources, not all of which include PRP. Thus the respective figure for the size of PRP relative to base salary might be considerably smaller.

The Med-Vantage (2011) survey, which also collects data on PRP for hospitals, found that in 2010, 2 % of the total reimbursements to hospitals for PRP incorporating health plans was performance-related. (In the literature regarding PRP in healthcare, the size of incentive pay ranges from less than 1 % to a maximum (but not actual) of 25 % in the British quality and outcomes framework (Trisolini, 2011).)

Characteristics of the PRP plans: performance measures and organizational level of performance measurement

The EK compensation system survey provides information on the incentive plans for different employee groups. The focus is on the organizational level of the performance measurement and the measures themselves. Table 3 reports that in 2007, on average, the performance of all healthcare was measured at two levels, the most popular being the profit center* (e.g., hospital) and the corporate group. This differs noticeably from the average figures of the private sector, where broad measures (profit center, company and group level) are relatively less common and narrow measures (own work, team) are more common. This rather broad performance measurement in the healthcare sector most likely reflects two things. The first is the joint (or team) nature of the work, and the second is the relatively small size of the healthcare providers. Accordingly, measuring performance at a narrower level may induce distortions, and given that the providers are relatively small, firm-level measures are therefore not too risky.

Table 3 Organizational levels of performance measurement

	Private healthcare		Private sector	
	Lower	Upper	Lower	Upper
Own work	0.21	0.24	0.49	0.58
Team	0.18	0.18	0.49	0.37
Establishment	0.15	0.03	0.18	0.16
Department	0.2	0.28	0.38	0.33
Project	0.12	0	0.26	0.3
Profit center	0.48	0.44	0.6	0.63
Company	0.23	0.42	0.59	0.74
Group (of companies)	0.41	0.43	0.34	0.42
Average	1.97	2.01	3.33	3.53
Number of observations	13	15	706	731

Table 4 lists the performance measures assessed in the EK questionnaire and the results for the private healthcare sector and private sector in general. Again, on average, fewer measures are used in healthcare than in the private sector. Additionally, upper white-collar workers have, on average, 0.4 measures more than lower white-collar workers. The most important performance measures are profitability and quality of service or products, followed by development goal and turnover. Interestingly, cost savings as a performance measure is used much less frequently in the Finnish private healthcare sector than in the private sector. Another measure that shows the orientation of the hospital studied here is sales targets, which are used as a performance measure in 40 % of the plans for upper white-collars. This reflects the nature of the private sector hospitals, who are not worried about over provision of services.

These performance measures reflect the nature of the organizations under study. The private sector healthcare providers aim for profitable operations while maintaining high quality. Most plans also combine broad measures such as profitability with narrower measures such as quality and productivity. This is consistent with theory, which suggests that to balance the trade-off between risk and distortion, combinations of narrow and broad measures should be used.

Table 4 The performance measures

	Private healthcare		Private sector	
	Lower	Upper	Lower	Upper
Cost savings	0.2	0.15	0.45	0.45
Quality of service or products	0.41	0.44	0.62	0.55
Productivity	0.38	0.29	0.54	0.55
Lead time	0.26	0.04	0.26	0.23
Inventory quantity or value	0	0	0.17	0.19
Development goal	0.4	0.43	0.53	0.58
Sales target	0.12	0.4	0.34	0.39
Profitability	0.94	0.92	0.81	0.92
Enhanced utilization of capital	0.12	0.12	0.12	0.19
Turnover	0.36	0.37	0.27	0.36
Market share	0	0.12	0.13	0.19
Other	0.25	0.27	0.38	0.36
Number on average	3.2	3.6	4.6	5
Number of observations	13	15	724	744

Med-Vantage (2011) surveys regarding US PRP programs offer a good reference point for these measures. In the US, clinical quality is the most heavily weighted measure, while “efficiency and cost” is the second most important measure (the numbers are 60 % and 20 %, respectively, while the rest of the measures account for the remaining 20 %). While the contrast with our results is quite considerable, it can largely be explained by the differences in the funding of healthcare.

In the US, performance has also been measured at multiple levels. When the health plans were examined to determine the number of levels on which they measure and reward performance (maximum being 3) the results were 1.69 levels in 2006, 1.85 levels in 2008, but only 1.12 levels in 2010. Possible explanations for the decline could be the changes in legislation, the financial crisis, or a better understanding of pay-for-performance and innovation (especially in the area of quality and outcome measurement). A Med-Vantage (2011) survey found that 52 % of health plans measured performance in 2010 at the physician level, 36 % at the physician group level and 24 % at the large IPO level.

Case Study: The Orthopedic Hospital

Next we study a case hospital and analyze its incentive plan in light the previous discussion. We start with the number of operations performed in the hospital. The total number of operations in the hospital was 2037 operations in year 2003, 2327 operations in 2007 and 1732 operations in 2010. Table 5 shows the number of operations for the surgeons who were employed for the entire period. As evidenced in the table, the initial increase and subsequent decrease in operations applied broadly to all surgeons. Based on interviews, the volume decrease in 2010 was partially caused by the financial crisis and partially by increased competition among hospitals. After the implementation of the incentive system, the annual output per surgeon increased by 11 %. However, by 2010, the outputs per surgeon had decreased by 41 %.

Table 5 The number of operations by surgeon

	Number of operations 2003	Number of operations 2007	Number of operations 2010	Change 2003-2007	Change 2003-2010
Surgeon 1	247	326	159	32 %	-36 %
Surgeon 2	85	167	62	96 %	-27 %
Surgeon 3	291	319	138	10 %	-53 %
Surgeon 4	224	284	80	27 %	-64 %
Surgeon 5	132	175	117	33 %	-11 %
Surgeon 6	170	212	91	25 %	-46 %
Surgeon 7	120	106	62	-12 %	-48 %

The changes in the number of operations may take place through two distinct channels. That is, the surgeons may operate on more patients per day or they may operate more days per week. As evidenced in Table 6, these channels operate differently when demand is increasing or decreasing. The number of operations per day increased by 24 % between 2003 and 2007, and the OR days per surgeon increased by 4 % for the same time period. In 2010, the daily output per surgeon decreased by 3 % while the OR days decreased by 41 % compared to 2003. These developments may reflect the fact that many of the surgeons are able operate also in other hospitals: in a downturn they shift their effort to other hospitals where demand may be more robust.

Table 6 The OR days and operations per OR day per surgeon in years 2003, 2007 and 2010

	Days in OR 2003	Operations per day 2003	Days in OR 2007	Operations per day 2007	Days in OR 2010	Operations per day 2010
Surgeon 1	111	2,2	117	2,8	84	1,9
Surgeon 2	75	1,1	95	1,8	44	1,4
Surgeon 3	120	2,4	117	2,7	75	1,8
Surgeon 4	114	2,0	119	2,4	49	1,6
Surgeon 5	111	1,2	113	1,5	71	1,6
Surgeon 6	102	1,7	105	2,0	64	1,4
Surgeon 7	88	1,4	76	1,4	51	1,2

Despite the initial increase in the number of operations, cost efficiency decreased in 2003 and 2010 by 3 % and 44 %, respectively. It is evidenced from Table 7 that while the surgeon salaries per operation initially increased by 41 %, this increase was partially offset by a decrease in nursing costs both in the operating room and in the ward. This, of course, reflects the fact that nurses were paid a flat wage throughout the period of observation. Thus, even though the change in the incentive plan decreased cost efficiency, the demand conditions were favorable. In 2010, when demand fell, the costs of all resource groups per operation increased. By then, the surgeon salaries per operation had increased by 136 %, and the total salary costs had risen by 44 %. Table 7 shows that for the hospital, the incentive system was financially weak.

Table 7 The salary costs per operation in 2003, 2007 and 2010

Cost per operation	2003	2007	Difference from 2003	2010	Difference from 2003
Surgeon salaries	400	564	+ 165 € (41 %)	942	+543 € (136 %)
OR nursing	603	559	-44 € (-7 %)	722	+120 € (20 %)
OR total salaries	1002	1123	+ 121 € (12 %)	1664	+662 € (66 %)
Ward nursing costs	913	846	-67 € (-7 %)	1094	+181 € (20 %)
Total salary costs	**1915**	**1970**	**+ 54 € (3 %)**	**2759**	**+843 € (44 %)**

The incentive plan did not work as anticipated for two reasons. First, the results imply that environmental factors played a larger role for the changes in the number of operations than the efforts of the surgeons. If the employees are not able to strongly respond to incentives, the incentive should not be very intense. In this case, the incentives were quite strong. After all, 4 % of the list price of an operation is a substantial part of the profit generated by the operation. Second, the plan held the base pay constant even when the incentives were introduced. This made the plan vulnerable to changes in demand.

Discussion and Conclusions

Performance measurement is the key challenge in any incentive system, and the appropriate performance measures depend on the goals of the organization. These, in turn, are heavily influenced by the organization of the healthcare system and the ownership of the organization. Most of the literature on incentive plans in healthcare has concentrated on the US and UK, where insurers play a large role. We focus on incentives for physicians and emphasize that distortions possibly caused by incentive pay are always defined with respect to the goals of the organization. Thus, publicly funded and privately owned healthcare providers should have very different incentive systems.

Herein we show that typical performance measurement incentive plans for healthcare providers in the Finnish private sector are based on a combination of profitability and some narrower measure such as quality or productivity. Performance is typically measured at the hospital level. The features of the incentive plan make sense in light of the theory as a combination of broad and narrow measures helps to balance the trade-off between risk and distortion, while measurement at the hospital level is consistent with the team nature of the work and the relatively small size of the hospital.

Proper performance measures help organizations achieve their goals. However, if the available performance measures are not very good, then the incentives will be low causing the plan to have a negligible effect on operations.

Whether PRP improves the cost efficiency of healthcare provision depends not only on the performance measures but on the design of the plan as a whole. The key elements are the base wage

and the intensity of incentives. These should be determined simultaneously as adding PRP on top of existing wages may make the system vulnerable in terms of costs.

We conducted a case study of a hospital that changed its compensation system. The hospital implemented an incentive plan that provided rewards based on the number of operations performed. Only doctors were included in the compensation system despite the fact that the majority of the resource consumption per care episode was derived from nurse resources (Hussey et al., 2009). The analysis shows that the cost efficiency of doctors was lower in the new model and that the cost efficiency of nurses improved. Therefore, if there had been doctors available, adding more fixed-salary doctors to the system rather than paying them more would have been a more cost-efficient way to improve productivity.

The incentive plan was determined to be problematic for the hospital. In terms of productivity, the problem was that the fixed salary was not adjusted. Therefore, if personnel productivity remains at baseline, cost-efficiency decreases. The incentive system assumes that the implementation of the system automatically increases volume and productivity. The demand of the services varied considerably, and the incentive system could not permanently increase the total demand. As the output per surgeon dropped below baseline, the cost-efficiency decreased dramatically because the fixed salary did not decrease, and in addition, the bonuses per operation were paid automatically. The study results indicate that profit or productivity-based incentives may be considered in elective surgery to ensure the benefits to the producer. In addition, limiting the incentives only to specific personnel groups may cause sub-optimization of surgery as the output is produced by a multidisciplinary-team.

References

Asch, B. J. (1990). Do Incentives Matter? The Case of Navy Recruiters. *Industrial and Labor Relations Review*, 89S-106S.

Baker, G. (2002). Distortion and Risk in Optimal Incentive Contracts. *Journal of Human Resources, 37*(4), 728-751.

Christianson, J. B., Leatherman, S. T., & Sutherland, K. (2007). *Financial incentives, healthcare providers and quality improvements: a review of the evidence*: Health Foundation.

Conrad, D. A., & Perry, L. (2009). Quality-based financial incentives in health care: can we improve quality by paying for it? *Annual review of public health, 30*, 357-371.

Conrad, D. A., Sales, A., Liang, S. Y., Chaudhuri, A., Maynard, C., Pieper, L., . . . Piland, N. (2002). The impact of financial incentives on physician productivity in medical groups. *Health Services Research, 37*(4), 885-906.

Courty, P., & Marschke, G. (2004). An Empirical Investigation of Gaming Responses to Explicit Performance Incentives. *Journal of Labor Economics, 22*(1), 23-56.

Doran, T., Kontopantelis, E., Valderas, J. M., Campbell, S., Roland, M., Salisbury, C., & Reeves, D. (2011). Effect of financial incentives on incentivised and non-incentivised clinical activities: longitudinal analysis of data from the UK Quality and Outcomes Framework. *BMJ: British Medical Journal, 342*.

Dranove, D., Kessler, D., McClellan, M., & Satterthwaite, M. (2003). Is More Information Better? The Effects of "Report Cards" on Health Care Providers. *The Journal of Political Economy, 111*(3), 555-588.

Emmert, M., Eijkenaar, F., Kemter, H., Esslinger, A. S., & Schöffski, O. (2012). Economic evaluation of pay-for-performance in health care: a systematic review. *The European Journal of Health Economics, 13*(6), 755-767.

Fairbrother, G., Hanson, K. L., Friedman, S., & Butts, G. C. (1999). The impact of physician bonuses, enhanced fees, and feedback on childhood immunization coverage rates. *American Journal of Public Health, 89*(2), 171-175.

Fleetcroft, R., & Cookson, R. (2006). Do the incentive payments in the new NHS contract for primary care reflect likely population health gains? *Journal of health services research & policy, 11*(1), 27-31.

Freeman, R. B., & Kleiner, M. M. (2005). The Last American Shoe Manufacturers: Decreasing Productivity and Increasing Profits in the Shift from Piece Rates to Continuous Flow Production*. *Industrial relations: a journal of economy and society, 44*(2), 307-330.

Gaynor, M., & Pauly, M. V. (1990). Compensation and productive efficiency in partnerships: evidence from medical groups practice. *Journal of Political Economy, 98*(3), 544-573.

Gibbs, M. J., Merchant, K. A., Van der Stede, W. A., & Vargus, M. E. (2009). Performance measure properties and incentive system design. *Industrial relations: a journal of economy and society, 48*(2), 237-264.

Golden, B. R., & Sloan, F. A. (2008). Phycisian Pay for Performance: Alternative Perspectives. In F. A. Sloan & K. Hirchel (Eds.), *Incentives and choice in health care* (pp. 289-318). Cambridge, Massachusetts: MIT Press.

Grant, D. (2009). Physician financial incentives and cesarean delivery: new conclusions from the healthcare cost and utilization project. *Journal of health economics, 28*(1), 244-250.

Holmström, B., & Milgrom, P. (1991). Multitask Principal–Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *Journal of Law, Economics, and Organization, 7*(special issue), 24-52.

Hussey, P. S., De Vries, H., Romley, J., Wang, M. C., Chen, S. S., Shekelle, P. G., & McGlynn, E. A. (2009). A systematic review of health care efficiency measures. *Health Services Research, 44*(3), 784-805.

Institute of Medicine. (2007). *Rewarding Provider Performance: Aligning Incentives in Medicare*: National Academies Press.

Kauhanen, A., & Napari, S. (2012a). Career and Wage Dynamics: Evidence from Linked Employer-Employee Data. In S. W. Polachek & K. Tatsiramos (Eds.), *Research in Labor Economics* (Vol. 36, pp. 35-76): Emerald Group Publishing Limited.

Kauhanen, A., & Napari, S. (2012b). Performance measurement and incentive plans. *Industrial relations: a journal of economy and society, 51*(3), 645-669.

Krasnik, A., Groenewegen, P. P., Pedersen, P. A., von Scholten, P., Mooney, G., Gottschau, A., . . . Damsgaard, M. T. (1990). Changing remuneration systems: effects on activity in general practice. *BMJ: British Medical Journal, 300*(6741), 1698.

Lazear, E. P., & Oyer, P. (2012). Personnel Economics. In R. Gibbons & J. Roberts (Eds.), *Handbook of Organizational Economics*. Princeton, New Jersey: Princeton University Press.

Lee, G. M., Kleinman, K., Soumerai, S. B., Tse, A., Cole, D., Fridkin, S. K., . . . Kassler, W. (2012). Effect of nonpayment for preventable infections in US hospitals. *New England Journal of Medicine, 367*(15), 1428-1437.

Lillrank, P., Groop, P. J., & Malmström, T. J. (2010). Demand and Supply–Based Operating Modes—A Framework for Analyzing Health Care Service Production. *Milbank Quarterly, 88*(4), 595-615.

Med-Vantage (Producer). (2011). 2010 National P4P Survey: executive summary. Retrieved from http://www.imshealth.com/deployedfiles/ims/Global/Content/Solutions/Healthcare%20Analytics%20and%20Services/Payer%20Solutions/Survey_Exec_Sum.pdf

Mullen, K. J., Frank, R. G., & Rosenthal, M. B. (2010). Can you get what you pay for? Pay-for-performance and the quality of healthcare providers. *The Rand journal of economics, 41*(1), 64-91.

Petersen, L. A., Woodard, L. D., Urech, T., Daw, C., & Sookanan, S. (2006). Does pay-for-performance improve the quality of health care? *Annals of internal medicine, 145*(4), 265-272.

Prendergast, C. (1999). The provision of incentives in firms. *Journal of economic literature, 37*(1), 7-63.

Raith, M. (2008). Specific knowledge and performance measurement. *The Rand journal of economics, 39*(4), 1059-1079.

Shafrin, J. (2010). Operating on commission: analyzing how physician financial incentives affect surgery rates. *Health economics, 19*(5), 562-580.

Shen, Y. (2003). Selection Incentives in a Performance-Based Contracting System. *Health Services Research, 38*(2), 535-552.

Trisolini, M. G. (2011). Quality Measures for Pay for Performance. In J. Cromwell, M. G. Trisolini, G. C. Pope, J. B. Mitchell & L. M. Greenwald (Eds.), *Pay for Performance in health care: Methods and approaches*: RTI Press.

Walker, S., Mason, A. R., Claxton, K., Cookson, R., Fenwick, E., Fleetcroft, R., & Sculpher, M. (2010). Value for money and the Quality and Outcomes Framework in primary care in the UK NHS. *The British Journal of General Practice, 60*(574), e213.

Van Herck, P., De Smedt, D., Annemans, L., Remmen, R., Rosenthal, M. B., & Sermeus, W. (2010). Systematic review: effects, design choices, and context of pay-for-performance in health care. *BMC Health Services Research, 10*(1), 247.

Vuorenkoski, L., Mladovsky, P., & Mossialos, E. (2008). Finland: Health system review. *Health Systems in Transition, 10*(4), 1-168.