

Performance Evaluation Systems and Unwarranted Variation

In many industrialized countries, a number of factors – rising costs, technological advancements, population aging, and medical errors – have contributed to the decision to adopt performance management tools (Smith [2002](#); Arah et al. [2006](#); Kelley and Hurst [2006](#)). Since the 1980s, the implementation of “new public management” principles has promoted the use of performance measurement to drive a more efficient, effective, and accountable public sector (Hood [1995](#); Lapsley [1999](#)). Originally, performance measurement focused primarily on financial measurements and was unable to support organizations’ efforts to achieve multiple strategic objectives and to drive changes (Pollitt [1985](#), [1986](#); Ghobadian and Ashworth [1994](#); Guthrie and English [1997](#); Lorden et al. [2008](#)). The adoption of comprehensive multidimensional performance measurement frameworks, such as the balanced scorecard, has been suggested (Eccles [1991](#); Jackson [1993](#); Kloot and Martin [2000](#); Fottler et al. [2006](#); Yang and Tung [2006](#)). Several studies led to the development of conceptual frameworks and models to help countries build effective tools (Arah et al. [2003](#), [2006](#); Smith [2002](#); Chang et al. [2002](#)). Simultaneously, many authors presented papers on the design of performance evaluation systems to be adopted at the organizational, regional, or national level (e.g., Linard et al. [2000](#); Aidemark [2001](#); Zelman et al. [2003](#); WHO [2003](#); Ba-Abaad [2009](#)).

While there is a broad range of literature on performance evaluation systems at the hospital level, only a few papers provide a regional focus, the majority of which concern design and implementation processes (Hilarion et al. [2009](#); Pink et al. [2001](#)). Yet the analysis of performance evaluation systems at the regional level and the effects on system performance are of particular interest given the devolution process that has occurred in most OECD countries. The need for a transparent performance evaluation system has also grown recently in the USA where the concept of accountable care organizations (ACOs) has been enacted in national health reform legislation, as a strategy to address lack of integration in the healthcare system (McClellan et al. [2010](#)).

ACOs are defined as contractual agreements among providers who are held jointly accountable for achieving measured quality improvements and reductions in spending growth (Shortell and Casalino [2008](#); Fisher et al. [2009](#); Stukel et al. [2013](#)). McClellan et al. ([2010](#)) “[emphasize] that these quality and cost improvements must achieve overall per capita improvements in quality and cost, and that ACOs should have at least limited accountability for achieving these improvements while caring for a defined population of patients.” The ACO model can be implemented at different levels. A first step could entail the participation of a legal entity with basic health information technology and performance reporting capabilities for measuring a core “starter set” of quality, efficiency, and patient experience indicators. These measures should be linked to payment systems where shared savings are a consideration for meeting quality and spending targets. As additional steps, the ACO model could include an advanced infrastructure that is also capable of measuring outcomes, adopting a payment model with risk adjusted, partial capitation payments, and quality bonuses.

The ACO model might be easier to implement in a public health system where policy-makers have considerable responsibility and power. Yet, even in such a context, it will not succeed unless the performance evaluation system emphasizes cooperation, rather than competition, between key actors. Critical to any governance system designed to enhance quality in public health is the joint planning, development, “ownership,” and sharing of a framework – one that reflects the collaborative efforts of the various stakeholders such as health authorities, regional administrations, patients and citizens, and professionals.

Support for the health system's overall functioning, as well as its specific local components, requires a performance evaluation system (PES) capable of highlighting areas of excellence while also identifying critical or weak areas (Jones [2000](#)). Through a mutually agreed-upon set of indicators, the PES should start as a best-practice enhancement process among local healthcare institutions via a benchmarking system (McNair and Leibfried [1992](#)). An informative example of a successful performance evaluation system in a Beveridge health system with standard services (categories A and B, [Table 1](#)) is the one used in the Tuscany Region in Italy since 2004 and currently implemented in eight other Italian Regions (Nuti [2008](#), [2013](#)). In 2004, a multidimensional performance measurement system was introduced to assess and monitor Tuscany Region's 12 local health authorities (LHAs) and four teaching hospitals (THs). The PES consists of 50 measures comprised of more than 130 indicators, classified into six assessment dimensions: population health, regional health strategies, quality, patient satisfaction, staff satisfaction, and efficiency/financial performance. The performance outcomes of the two LHAs have been simplified and represented graphically as "target" diagrams ([Fig. 7](#)). Each diagram shows the six dimensions and five levels of performance (colored rings); indicators with high performance scores (dark green) are displayed closer to the center, whereas indicators with low scores (red) are displayed farther from the center.

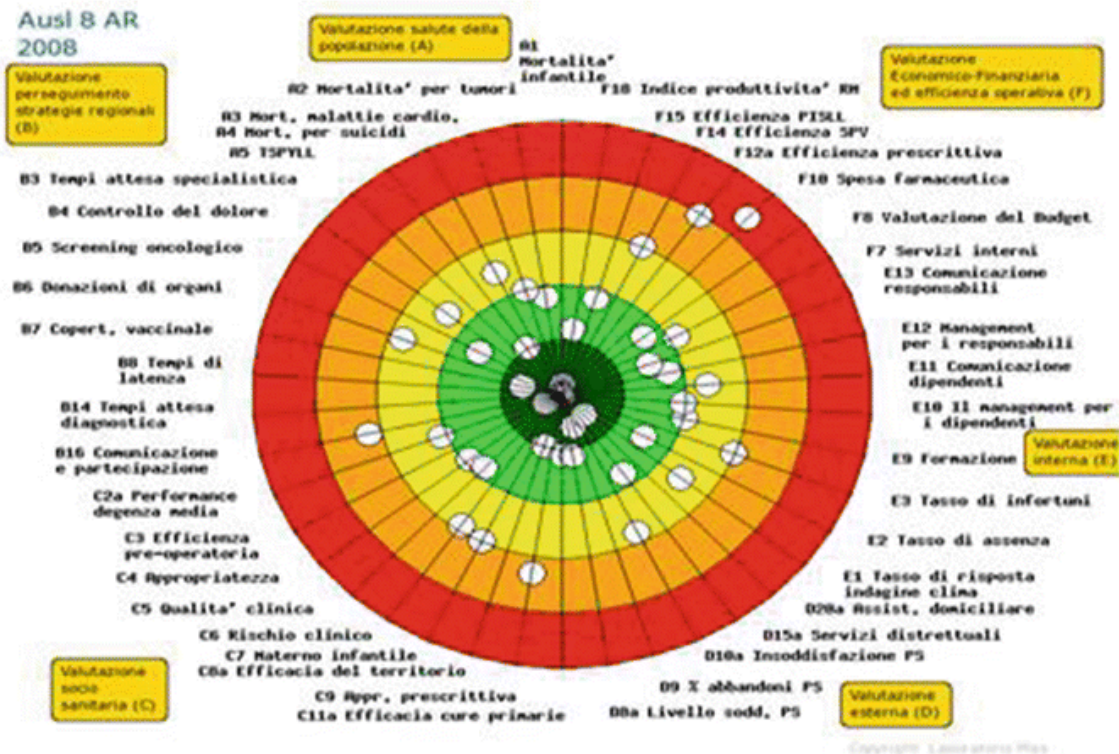
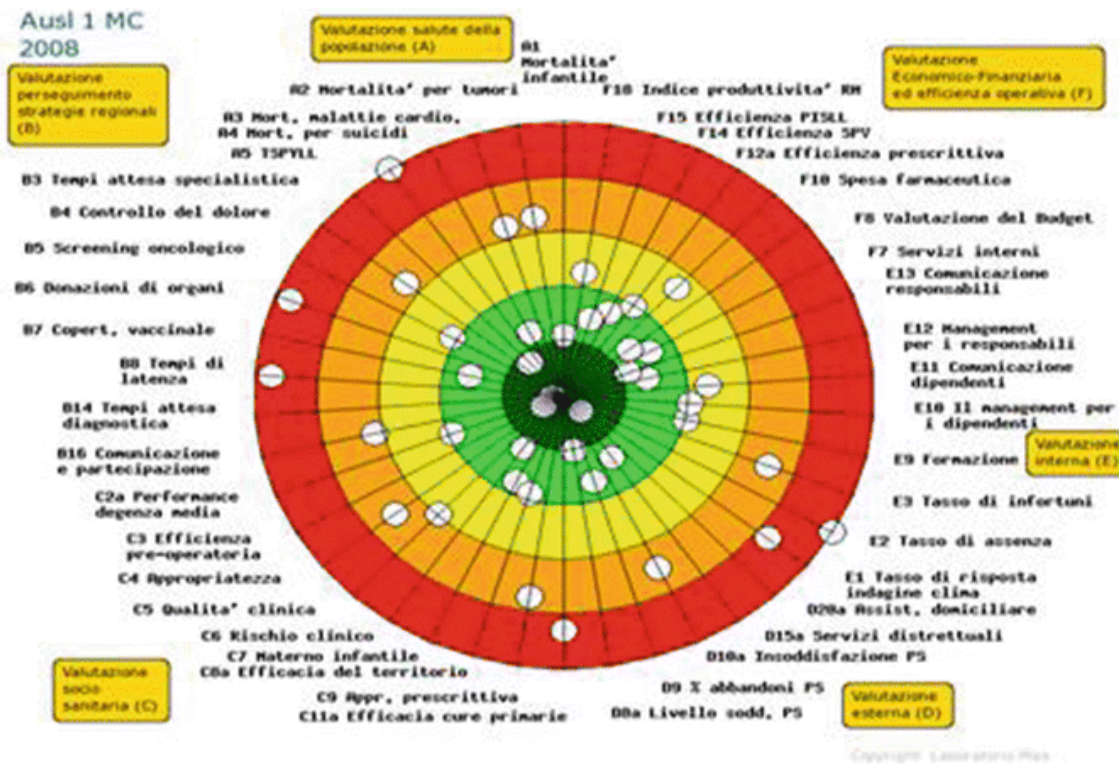


Fig. 7

Two examples of local health authorities target diagrams

From the start, the Tuscan PES was adopted as a decision support tool at both the regional and local strategic management levels. Regional performance results on selected indicators are monitored every 3 months and discussed at system-wide meetings of the LHAs' top managers and the region's

administrators. LHA managers are asked to discuss the performance of their organizations, to provide feedback on strengths and weaknesses, and to discuss future actions and goals. All of the local health authorities have incorporated PES indicators into their budgeting systems, using benchmarking across health authorities to set internal standards. Starting in December 2007, all performance indicators, presented as benchmarks, and yearly targets were made available online to top managers and professionals. Since July 2008, they have been available online to the public at www.valutazione-sanitatoscana.sssup.it. Further, the PES results have been integrated with CEO reward systems and with other regional systems such as accreditation and clinical risk monitoring.

The application of the PES has been credited with facilitating a number of improvements across the Tuscan healthcare system (Neri [2009](#)). Its success suggests that targets can be used to change the behavior of individuals and organizations if applied in the right manner (Bevan and Hood [2006](#)), particularly when there is a high level of congruence between what is bureaucratically and clinically significant. The benchmarking and advancement process associated with the PES is based on local involvement and agreement; moreover, it encourages local services to learn from the data on their own outcomes as well as those of their colleagues, enhancing their understanding of what “good” or “best” practice might look like and highlighting the factors that contribute to the framework’s success. Regular tracking of performance measurements allowed for analysis of the effects of continuous use of the PES on performance. Since 2006, on average more than 50 % of indicators registered a performance improvement; an even higher percentage of monitored indicators showed a reduction in geographical variation (see Nuti et al. [2013](#)).

Figure [8](#) shows the trends in performance and variation for the Tuscany Region. The x-axis represents the performance obtained in 2009 (ranked from 0 to 5) with respect to other Italian Regions. The y-axis represents the 2008–2009 trend grouped into five bands and indicates the percentage performance improvement (above zero) or worsening (below zero). Each plot represents an indicator. Geographical variation is represented by color: increasing variation (red), decreasing variation (green), and stability (black). As Fig. [8](#) highlights, the majority of indicators are positioned in the upper-right quadrant, which indicates both good performance with respect to the national average and an improvement during the study period. However, not all indicators positioned in this quadrant show both improved performance and reduced variation – some improved, but the variation remained stable or even increased. In a few cases (e.g., Cesarean rate and percentage of laparoscopic cholecystectomy), performance remained relatively stable, but variation within the region improved. A negative assessment was assigned to those cases – such as the percentage of medical day hospitals for diagnostic purposes – with a reduction in variation but worsened performance.

2008-2009 trend

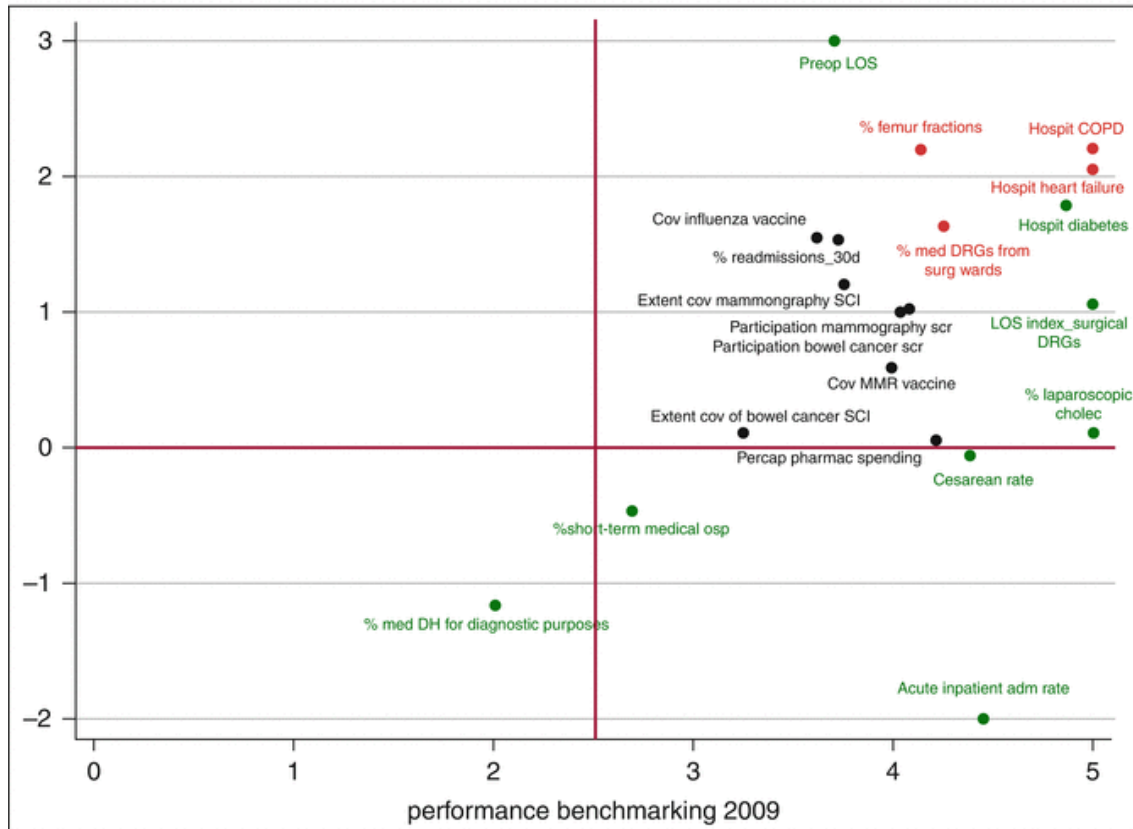


Fig. 8

Tuscany Region, trend (2008–2009) and performance benchmarking (2009). *Green*: reduced variation, *Red*: increased variation, *Black*: stable variation

The Tuscany Region’s integrated strategies of monitoring, public disclosures, discussions with professionals, and linkages with both the reward and accreditation systems have achieved a significantly greater (and faster) level of improvement compared to the performance of other regional strategies (Pinnarelli et al. [2012](#)).

Design and Implementation of Performance Evaluations Systems: Factors to Be Considered

Tuscany Region’s experience provides evidence to support the hypothesis that continuous measurement of organizational performance, together with targeted objectives, serves to orient individuals and organizations toward more responsive behavior (Bevan and Hood [2006](#)). Achieving high quality is an ever-changing process, and therefore, it is essential that regional governments support the widespread use of benchmarking and engage the entire workforce in a shared commitment to improving quality. Indicators measuring quality and appropriateness have received more attention than other indicators. Empirical studies on the Tuscan PES data explain this choice. The research findings highlight the fact that the capacity to control appropriateness and quality, rather than efficiency, is linked to financial sustainability (Nutti et al. [2011](#)).

A detailed analysis of the PES’s function and utility reveals six significant factors that contribute to success. The first factor is the **participation of clinical professionals and local managers** in the

wider process (Abernethy and Stoelwinder [1995](#); Jones and Dewing [1997](#)). The PES was designed and developed through close collaboration between healthcare professionals and managers (Nutti [2009](#)). It is important that all actors within the healthcare system, clinicians, local managers, and regional administrators, participate in the process and that they share their opinions in developing the evaluation indicators. This collaborative, inclusive approach is vital to mitigating the risk that local managers may manipulate data or that they will not use the performance evaluation system because they do not believe in the relevance and significance of the regional government's proposed indicators. Ongoing training that involves the input of HA coordinators and senior and middle management representatives has facilitated the development, diffusion, comprehension, and refinement of the PES and serves to further stimulate a sense of ownership.

The second factor is **benchmarking** – a process that helps local organizations learn from others' experiences, overcoming the limitations of self-referential evaluations and driving improvement even in the absence of a marketized form of competition. PES-produced information is represented uniformly, enabling efficient, constructive comparisons among the system's local health authorities although the comparability of budget allocations remains problematic. This crosscutting approach makes it possible to distinguish between problematic aspects related to the regional nature of HAs and those which derive from an individual authority's behavior. For instance, if all the LHAs registered a negative performance for a particular indicator, then clearly there is a more generalized problem that requires attention at a regional level. If, instead, performance varies greatly between authorities, it becomes clear that some authorities could learn from others and that collaboration (via "good practice" sharing) could help overcome problematic issues of poor performance. In this way, the means for assuring performance is closely tied to the vital role of local health authorities as learning organizations, rather than the more conflicting priorities which exist across many NHS contexts (Sheaff and Pilgrim [2006](#)).

The third factor is the strong emphasis on **reputation**. Evaluation and disclosure of hospital performance data are increasingly used to improve healthcare quality (Daneman et al. [2012](#); Fung et al. [2008](#)). The idea is that the availability of hospital performance data may improve quality of care by "empowering" accountability (Chassin et al. [1996](#); Lee et al. [2004](#)). According to some (Mannion and Davies [2002](#); Hibbard et al. [2003](#), [2005](#)), a key driver of performance change in the health sector is the threat of reputational damage. There is utility in comparing performance among individual HAs, but above all through highlighting and communicating the positive results obtained by the "best practice" examples. In order to leverage reputation as a key driver of improved performance, it is important not only to publicly report the levels of performance but also to show how HAs with pronounced changes were able to improve their performance.

The fourth factor is **periodic monitoring**. Performance indicators should be monitored systematically and periodically and presented and discussed at joint meetings of policy-makers and CEOs of each LHA. These meetings become forums for knowledge development and dissemination, where successful approaches can be discussed, adopted, and transferred to other settings. They also work as a deterrent to gaming (Nutti et al. [2008](#)). A **graphical reporting system** is the fifth hallmark of a PES. The visual approach must be easy to understand. Effective graphical representation is a way to ensure transparency so that all citizens understand how the system works. In fact, all data and results must be available on a publicly accessible website and in a report published annually. In this way, citizens become engaged in the accountability process and can access information on how public resources are used to deliver value for patients/service users. There may be disagreement over the extent to which citizens are able and willing to use this information, and citizen interest should not simply be assumed (Entwistle et al. [1998](#)). Despite this, the transparency and accessibility of such data within the public

sphere amplify the reputation-based effect (and corresponding moral obligations) on LHAs' behavior and on the system as a whole.

The sixth factor is **links with strategic planning**. It is vital that the effectiveness of the PES is not understood as a disembodied or apolitical process. A politically coherent long-term vision, personified by innovative key advocates (e.g., the Regional Health Council), is essential to assuring that "change management" that can be enacted over a number of years. Political commitment must be reflected in funding levels capable of sustaining the "organizational learning" aspects of quality development. An interesting example, reported by Veillard et al. ([2010](#)), is the alignment in Ontario, Canada, of strategies, measurement, and local health system accountability. PESs with the above characteristics focus on achieving quality and appropriateness standards and work toward reducing avoidable variation among the health system's institutions. In Spain, the La Agencia Valenciana de Salud (Gómez and Carrillo [2009](#)) also offers an example of how avoidable variation may be reduced. Each health authority sets goals based on their actual starting point. The poorest performers are asked to accomplish much more than the "best practice" performers, enabling reduction in performance gaps across HAs.

Mechanisms to Manage Variation in Preference-Sensitive and Supply-Sensitive Category Services

Admission rates for elective procedures fall within health service categories C and D (Table [1](#)). Geographical variation for such procedures has been widely documented in different countries and regions (Appleby et al. [2011](#); McPherson et al. [1982](#); Wennberg [1999](#); Right Care [2011](#); Bottle and Aylin [2005](#); Majeed et al. [2002](#); Corallo et al. [2014](#)). As demonstrated in numerous studies (Appleby et al. [2011](#); Lucas et al. [2010](#); McPherson et al. [1996](#); Grytten and Sørensen [2003](#); Nuti et al. [2010](#)), patient characteristics and preferences do not completely account for geographic variation in the provision of these types of care; rather, differences in clinicians' behavior and judgment exert a greater influence. Individual clinicians often do not realize that their treatment decisions vary greatly across geographic areas and that they have relevant impacts on healthcare costs. In such cases, performance measurement systems may help to identify variation; however, they cannot change behavior unless professionals identify a standard.

In cases when there is no performance standard, high variation can be symptomatic of under/overuse of certain procedures. The Tuscany health system's efforts to manage variation in these types of services are instructive.

The Tuscany Region, in collaboration with professionals, has fixed a standard for MRI and CT rates across districts (Nuti and Vainieri [2012](#)). Given that the Tuscany Region had the highest rates in Italy in 2009, the region's policy-makers chose the median regional rate as a reference. The evaluation considered that rates that were above (below) the median rate may indicate overtreatment (or undertreatment) (Fig. [9](#)).

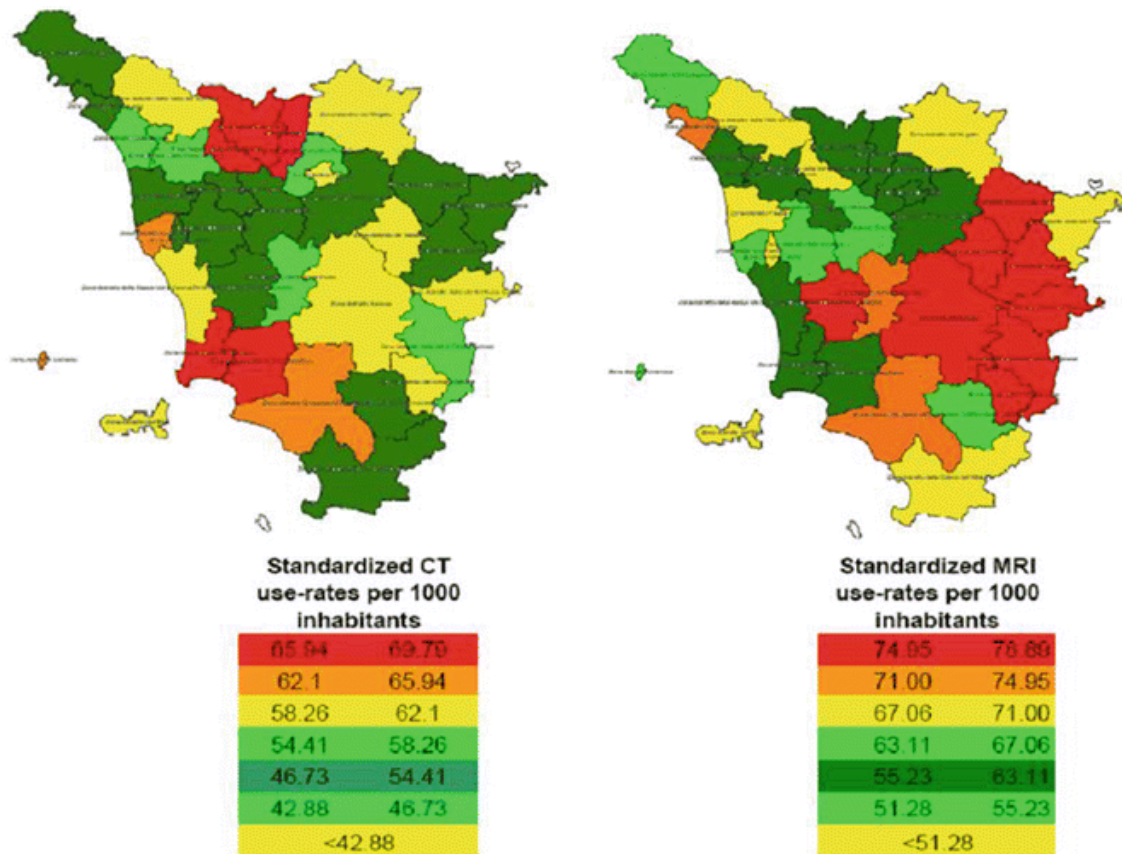


Fig. 9
CT and MRI use rates per 1000 inhabitants in the Tuscan districts in 2011

In 2012, the Tuscany Region began a series of meetings involving management representatives from the district health authorities, under the supervision of the Regional Health Council, to consider other C and D category services. The purpose of this initiative was to motivate clinicians to follow protocols and to collaborate on the creation of written guidelines on the role of patients. This intervention was designed to give patients a stronger voice to better address their needs and engage them as active participants in the variation reduction process. As reported also in other experience, an important step toward unifying practice patterns across geographic areas, patient involvement has been shown to help patients make informed choices about their care and consequently to contribute to reduction in variation (Bates [1990](#); Anthony et al. [2009](#); O'Connor et al. [2004](#); Légaré et al. [2012](#); Lomas et al. [1989](#); Brownlee et al. [2011](#)). Moreover, it called upon professionals, who were accustomed to working on their own, to work collectively as colleagues, fostering a culture of collaboration at the policy-level and helping to remove existing barriers to clinical integration. Along with establishing guidelines for patients, healthcare managers assumed responsibility for reducing geographic variation in elective surgery among their organizations and were incentivized to reduce high/low ratios within the next several years. Both these goals were then introduced into CEOs' healthcare 2012 compensation systems, in the hope that they would prioritize actions meant to reduce variation and encourage healthcare professionals and patients to collaborate on initiating this improvement process (De Jong [2008](#); Van Kemenade [2007](#)). To raise awareness about variation in healthcare, the Tuscany Region's ongoing initiative to manage variation has been proposed to a network of eight regions sharing the same PES.

Reward Systems to Orient Behaviors and Manage Variation

The reward system is one of the most widely applied managerial tools for orienting behaviors. Once performance measurement systems have been adopted, financial incentives can help to align stakeholders' differing expectations. The alignment of targets between different control mechanisms adopted by organizations is called goal congruence. Achieving goal congruence within an organization's operational systems – particularly with regard to issues of evaluation, rewards, and control – is critical to avoiding the numerous, unintended consequences of misaligned reward systems (Gomez-Mejia et al. [2010](#)).

Reward systems can serve to mitigate any negative perceptions workers may have about health sector reforms, facilitating positive changes during the health reform process (Buchan et al. [2000](#)). The financial incentive scheme, one of the most common (and studied) extrinsic rewards, is part of a wide set of remunerative strategies available to motivate employees and orient their behavior. Indeed, money appears to be a primary motivator for most people: Shortell and Kaluzny ([2006](#)) found that physicians rank money as one of the top five priorities due to the enormous investments of time and money in their training. Thus, financial incentives may have significant influence on physician behavior, encouraging quality improvements, especially if other factors also support that goal. However, the type of incentives to be used, both intrinsic (such as status acknowledgment) and extrinsic (such as financial bonus), must be carefully evaluated on the basis of a series of factors, such as the financing model adopted by the system (e.g., a Beveridge system in which the healthcare is provided by the government and financed mainly through taxes, a Bismark system providing universal coverage financed by employer/employee contributions), the type of health services to be promoted, and the kind of target to be achieved.

One way to analyze the impact of financial incentives on quality in healthcare is to consider their influence on the four health service categories outlined in Table [1](#). In this context, it may be helpful to divide the categories into two groups: (1) clinically proven and setting-sensitive care services (categories A and B of Table [1](#)) and (2) preference-sensitive and supply-sensitive care services (categories C and D of Table [1](#)).

For the first group, existing levels of health services have already suggested an appropriate reference standard, and thus, financial incentives may play an important role in encouraging health service providers to work toward meeting the standard. Several articles have reported the impact of financial incentives on quality. A variety of strategies link financial payment to results, among them payment by result, activity-based payment, lump sum, penalties, and pay for performance (P4P). Countries have adopted these incentives in different ways, both across years and services (Appleby et al. [2012](#), Street et al. [2007](#)). Appleby et al. ([2012](#)) analyzed the evidence for activity-based payments, finding that they led to efficiency but also to a number of unintended consequences, including increased variability across districts, resulting in increased hospital admissions at one extreme and undertreatment at the other.

Systematic reviews of financial incentives and pay for performance (Van Herck et al. [2010](#); Christianson et al. [2007](#)) have highlighted the heterogeneous effects on quality and performance – ranging from absent or negligible to strongly beneficial. Pearson et al. ([2008](#)) cast doubts on the effectiveness of P4P programs applied by Massachusetts health plans: the authors argued that the quality improvement achieved between 2001 and 2003 could have been reached even without those programs. Other authors have argued that P4P may lead to adverse selection, such as the exclusion of severely ill patients from care (Doran et al. [2006](#); Semuraga et al. [2011](#)).

To overcome these problems, other financial incentive schemes have been put in place. In the USA, for example, P4P has been replaced by bundled payments, also known as packaged pricing, microcapitation, risk-based contracting, and evidence-based case rates. Bundled payments programs are programs in which providers are paid a single fee for a set of evidenced-based services related to a diagnosis. The payments are typically linked to outcomes as well as other quality measures (Hackbarth et al. [2008](#); Shortell and Casalino [2008](#)). An example of a performance-based bundled payment system is ProvenCare, developed by Geisinger as a way to reimburse providers of coronary artery bypass graft surgery. Under this payment plan, physicians agreed to follow about 40 treatment guidelines, with the result that the average cost of care related to complications fell to half the historical average. To cut costs, in 2011 Medicare launched a bundled payment system for end-stage renal disease; however, there is still uncertainty about the consequences of this new system (Swaminathan et al. [2012](#)).

Consequently, there is no single financial incentive for providers that can use to improve quality of care under all circumstances. Moreover, the same incentives employed in different contexts can yield quite different results (see for instance, Town et al. [2004](#) and Glasziou et al. [2012](#)). As Glasziou et al. ([2012](#)) recently noted, financial incentives can improve the quality of clinical practice, but they can also be an expensive distraction.

For the second group of services (preference services and supply-sensitive category services), standards of performance may not yet have been established, and thus, other types of incentives should be put in place. Incentives to reduce variation must be carefully evaluated because of the high risk of unintended consequences and costs to the system in relation to uncertain results. The question raised by Glasziou et al. ([2012](#)), “Will financial incentives work and work better than other interventions, to change performance, and why?,” seems particularly appropriate for these categories. There are many nonfinancial ways to change the performance of individuals and organizations. As Flamholtz et al. ([1985](#)) have highlighted, performance outcomes are affected also by organization culture and structure. Similarly, Trisolini ([2011](#)) noted additional important factors that affect the success of reward systems: the social norms of professionalism among physicians, the range of motivational factors affecting their behavior, and the organizational settings in which they practice. With regard to motivational factors, Wynia ([2009](#)) showed that financial incentives for physicians can produce both positive and negative effects: on the one hand, they can damage physicians’ intrinsic motivation because they contribute to a perception of deprofessionalization; on the other hand, P4P agreements may be positively perceived because they provide payment for services – such as patient and family education or case management – that often go unreimbursed under fee-for-service plans. Since there is no “gold standard” for preference- and supply-sensitive services, financial incentives have a high risk of deprofessionalizing physicians. Under these circumstances, other kinds of levers may better support the goal of variation management, including strategies that lead to delivery system reforms – such as ACOs (IOM [2012](#)) – or mechanisms that influence beliefs and social norms, defined by Chenall ([2003](#)) as organic or informal mechanisms. However, formal mechanisms (such as performance measurement system, classified by Ouchi as bureaucracies) can fail when the ambiguity of performance evaluation is significantly high (Ouchi [1980](#)). In this situation organizations should seek other forms of control. Ouchi ([1980](#)) posits, “when bureaucratic failure results from excessively ambiguous performance evaluation, the sole remaining form of mediation is the clan, which relies on creating goal congruence.” Clan control is based on socialization (informal mechanisms) and peer review as the principal way to orient organization’s performance. In this sense, when uncertainty is high, physician peer groups can be an effective form of response to manage variation in health services category.

Simons ([1995](#)) identified four levers of control to stimulate employees to improve organization’s performance: belief systems, boundary systems, diagnostic control systems, and interactive control systems. While diagnostic control systems, including financial reward system and performance measurement systems previously mentioned, are not expected to have a positive impact on employee behavior and on performance because they require clear goals (which are possible in the case of clinically proven services and setting-sensitive care), the other three levers can be combined to cope with variation in the other cases (preference- and supply-sensitive care services). As [Table 2](#) shows, interactive control systems, belief systems, and boundary systems (closer to clan controls) should produce the desired reduction in variation. The interactive control systems in Simons’ paper are accomplished through the process of sharing knowledge and information across different levels of organizations. Boundary systems establish the rules of the game and identify actions and pitfalls that employees must avoid, whereas belief systems are based on the values, purpose, and direction of the organizations (Simons [1995](#)). **Table 2**
Control levers that support variation management in preference- and supply-sensitive care services

Service categories	Characteristics	Tools
Preference-sensitive care	Uncertainty	Belief systems and interactive control systems should be applied in order to identify the right approach to pursue
Supply-sensitive services	Pressure or temptation	Boundary systems and beliefs should be leveraged so that specialists will “do the right thing”

Hence, for instance, regularly held meetings to discuss performance changes may allow for a sharing of good practice and peer recognition of improvement and thus result in the reduction of inappropriate variation. Indeed socialization and the learning process, which are basic elements of the three Simons’ control levers, play an important role when uncertainty and ambiguity are high (Marengo and Pasquali [2012](#)). Moreover interactive and beliefs systems can also support and boost the results gained by formal systems adopted for service categories where standards are known.

For instance, the learning process, represented by new information, is four times more effective than profit incentives in motivating (and so orienting) professionals (Kolstad [2013](#)). In addition, interactive and belief systems can be effective ways to cope with data manipulation and gaming. In fact Bevan and Hood ([2006](#)) argued that mechanisms associated with explanations for improvements and peer review checks are much more effective at illuminating gaming practices than merely auditing results.

Conclusions

Research shows that reducing waste and increasing quality are not in opposition (Marshall and Øvretveit [2011](#); Øvretveit [2009](#)). Variation management can be a way to pursue both of these objectives. Considering the service categories from Nuti and Seghieri ([2013](#)), adapted from Wennberg et al. ([2002](#)), this chapter proposes various mechanisms to manage variation and highlights the fact that no single tool is capable of tackling unwarranted variation. As the Tuscany experience shows, system-wide performance evaluation systems can help improve performance and reduce variation in services that offer clinically proven effective care or services amenable to changes in the setting of care delivery (e.g., inpatient admissions for interventions which could be performed on a

day-surgery basis) without compromising outcomes. The key elements of success in the Tuscany program are factors linked to the design and process phase, such as:

The participation of clinical professionals and local managers in the process

Benchmarking

Emphasis on reputation

Periodic monitoring

Graphical and geographical reporting

Strong integration with other managerial mechanisms as incentive systems

Indeed diagnostic tools such as performance measurement systems may play a pivotal role in variation management and are relevant for identifying the path forward. Financial incentives may serve to reinforce the message of the performance evaluation system and can have a positive effect on changing the behavior.

For both preference- and supply-sensitive care, performance measurement systems are not enough. In these cases, variation can be managed using other mechanisms such as belief systems, which means involving professionals in a long-term process. A fundamental element is to conduct the analysis from a population-based perspective. Health professionals are, in fact, accustomed to analyzing provider-related production – indeed they focus on their own cases. Thus, a population-based approach can elicit a change in physicians' analyses, making them much more aware of the impact of their decisions on the value produced and the resources used.

In fact only a population-based analysis approach allows the extent of unwarranted variation to emerge. This requires changing the clinicians' paradigm from "medical management" to "population medicine." As Muir Gray highlighted in the July 2013 issue of the *Lancet*, "Population medicine is not a new specialty, it is a new paradigm that every clinician will sooner or later adopt, with a proportion of clinicians being allocated explicit time for working for the whole population. This is different from management of a service for the patients who present to the service. [...] In the twenty-first century, clinicians have a responsibility to the population they serve, to the patients they never see, as well as to the patients who have consulted or been referred. Individual clinicians, while still focused on the needs of the individual in front of them when in the consultation, also make decisions about the allocation and use of resources to maximise value for all the people in the population they serve" (Gray [2013](#)).

Regional strategies and governance tools may support this changing paradigm. The Tuscany Region example described in this chapter shows how a government-finance healthcare system based on the Beveridge system may manage variation to achieve not only performance improvement but also more equity. In fact, unwarranted variation reduction must be a main goal for universal coverage healthcare systems. Yet considerable work has still to be done with regard to the management of horizontal equity. Central and local governments must pay attention to geographic variation. The inclusion of indicators measuring unwarranted geographic variation and setting specific goals on this matter by including them in their PES can make a difference in the ability of Beveridge-like systems to effectively cope with horizontal equity.

References

Abernethy M, Stoelwinder J. The role of professional control in the management of complex organizations. *Acc Organ Soc.* 1995;20(1):1–17.

AHRQ Quality Indicators. Guide to prevention quality indicators: hospital admission for ambulatory care sensitive conditions. AHRQ Pub. No. 02-R0203. Rockville: Agency for Healthcare Research and Quality; 2001.

Aidemark L. The meaning of balanced scorecards in the health care organization. *Financ Account Manag.* 2001;17(1):23–40.

Anthony DL, Herndon MB, Gallagher PM, Barnato AE, Bynum JPW, Gottlieb DJ, Fisher ES, Skinner JS. How much do patients' preferences contribute to resource use? *Health Aff.* 2009;28(3):864–73.

Appleby J, Raleigh V, Frosini F, Bevan G, Gao H, Lyscom T. Variations in health care, the good, the bad and the inexplicable. London: The King's Fund; 2011.

Appleby J, Harrison T, Hawkins L, Dixon A. Payment by results how can payment systems help to deliver better care? London: The King's Fund; 2012.

Arah OA, Klazinga NS, Delnoij MJ, Ten Asbroek AHA, Custers T. Conceptual frameworks for health systems performance: a quest for effectiveness, quality, and improvement. *Int J Qual Health Care.* 2003;15(5):377–98.

[PubMed](#)

Arah OA, Westert GP, Hurst J, Klazinga NS. A conceptual framework for the OECD health care quality indicators project. *Int J Qual Health Care.* 2006;18:5–13.

[PubMed](#)

Bambas A, Casas JA. Assessing equity in health: conceptual criteria. in *Pan American Health Organization Equity and Health: Views from the Pan American Sanitary Bureau Washington, D.C: PAHO, 2001. (Occasional Publication No. 8).*

Ba-Abaad KM. Review of the literature of balanced scorecard and performance measurement: The case of healthcare organizations. *Business e-Bulletin.* 2009;1(1):33–47.

Barsanti S, Nuti S. The equity lens in the health care performance evaluation system. *Int J Health Plann Manag.* 2013. doi:10.1002/hpm.2195.

Bates T. Avoiding inappropriate surgery: discussion paper. *J R Soc Med.* 1990;83(3):176–8.

[PubMedCentral PubMed](#)

Bevan G, Hood C. What's measured is what matters: targets and gaming in the English public health care system. *Public Adm.* 2006;84(3):517–38.

Beveridge W. *Social insurance and allied services.* London: HMSO; 1942.

Blank RH, Bureau V. *Comparative health policy.* Basingstoke: Palgrave Macmillan; 2010.

Bonnefoy E, Orsi GA, Couble P, Loppin B. The essential role of drosophila HIRA for De novo assembly of paternal chromatin at fertilization. *PLoS Genet.* 2007;3(10):e182.

[PubMedCentral](#)

Bottle A, Aylin P. Variations in vaginal and abdominal hysterectomy by region and trust in England. *BJOG.* 2005;112(3):326–8.

[PubMed](#)

Bravemann P. Monitoring equity in health and healthcare: a conceptual framework. *J Health Popul Nutr.* 2003;21:181–92.

Brownlee S, Wennberg JE, Barry MJ, Fisher ES, Goodman DC, Bynum JPW. Improving Patient Decision-Making in Health Care: A 2011 Dartmouth Atlas Report Highlighting Minnesota 2011. *The Dartmouth Atlas.* The Dartmouth Institute for Health Policy and Clinical Practice.

Bubolz T, Emerson C, Skinner J. State spending on dual eligibles under age 65 shows variations, evidence of cost shifting from medicaid to medicare. *Health Affairs.* 2012;31(5):939–947.

Buchan J, Thompson M, O'May F. Health workforce incentive and remuneration strategies – a research review. Discussion paper no.4, WHO/EIP/OSD/00.14; 2000.

Cepiku D, Bonomi Savignon A. Governing cutback management: is there a global strategy for public administrations? *Int J Public Sect Manag.* 2012;25(6–7):428.

Chang L, Lin SW, Northcott DN. The NHS performance assessment framework. A balanced scorecard approach? *J Manag Med.* 2002;16:345–58.

[PubMed](#)

Chassin MR, Hannan EL, DeBuono BA. Benefits and hazards of reporting medical outcomes publicly. *N Engl J Med*. 1996;334:394e8.

Chenall RH. Management control systems design within its organizational context: findings from contingency-based research and directions for the future. *Acc Organ Soc*. 2003;28:127–68.

Christianson J, Leatherman S, Sutherland K. Financial incentives, healthcare providers and quality improvements. London: The Health Foundation; 2007.

Corallo AN, Croxford R, Goodman DC, Bryan EL, Srivastava D, Stukel TA. A systematic review of medical practice variation in OECD countries. *Health Policy*. 2014;114:5–14.

[PubMed](#)

Cylus J, Mladovsky P, McKee M. Is there a statistical relationship between economic crises and changes in government health expenditure growth? An analysis of twenty-four European countries. *Health Serv Res*. 2012;47(6):2204–24.

[PubMedCentral](#) [PubMed](#)

Dahlgren G, Whitehead M. Policies and strategies to promote social equity in health. Stockholm: Institute for Future Studies; 1991.

Dahlgren G, Whitehead M. Tackling inequalities in health: what can we learn from what has been tried? London: King's Fund; 1993.

Daneman N, Stukel TA, Ma X, Vermeulen M, Guttman A. Reduction in clostridium difficile infection rates after mandatory hospital public reporting: findings from a longitudinal cohort study in Canada. *PLoS Med*. 2012;9(7):e1001268.

[PubMedCentral](#) [PubMed](#)

De Jong J. Explaining medical practice variation. Social organization and institutional mechanisms, PhD thesis at Utrecht University; 2008.

Doran T, Fullwood C, Gravelle H, Reeves D, Kontopantelis E, Hiroeh U, Roland M. Pay-for-performance programs in family practices in the United Kingdom. *N Engl J Med*. 2006;355:375–84.

[PubMed](#)

Eccles RG. The performance measurement manifesto. *Harvard Business Review*. 1991;1:131–137.

Entwistle V, Sheldon T, Sowden A, Watt I. Evidence-informed patient choice. Practical issues of involving patients in decisions about health care technologies. *Int J Technol Assess Health Care*. 1998;14(2):212–25.

[PubMed](#)

Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder EL. The implications of regional variations in Medicare spending. Part 1: the content, quality, and accessibility of care. *Ann Intern Med*. 2003a;138:273–87.

[PubMed](#)

Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder EL. The implications of regional variations in Medicare spending. Part 2: health outcomes and satisfaction with care. *Ann Intern Med*. 2003b;138(4):288–298.

[PubMed](#)

Fisher ES, McClellan MB, Bertko J, Lieberman SM, Lee JJ, Lewis JL, Skinner JS. Fostering accountable health care: moving forward in Medicare. *Health Aff*. 2009;28(2):w219–31.

Flamholtz EG, Das TK, Tsui AS. Toward an integrative framework of organizational culture. *Acc Organ Soc*. 1985;10(1):35–50.

Fottler MD, Erickson E, Rivers PA. Bringing human resources to the table: Utilization of an HR balanced scorecard at Mayo Clinic. *HealthCare Management Review*. 2006;31(1):64–72.

Fung CH, Lim Y-W, Mattke S, Damberg C, Shekelle PG. Systematic review: the evidence that publishing patient care performance data improves quality of care. *Ann Intern Med*. 2008;148:111–23.

[PubMed](#)

Gemmill MC, Thomson S, Mossialos E. What impact do prescription drug charges have on efficiency and equity? Evidence from high-income countries. *Int J Equity Health*. 2008;7:12.

[PubMedCentral](#) [PubMed](#)

Ghobadian A, Ashworth J. Performance measurement in local government – concept and practice. *Int J Oper Product Manag.* 1994;14(5):35–51.

Glasziou PP, Buchan H, Del Mar C, Doust J, Harris M, Knight R, Scott A, Scott IA, Stckwell A. When financial incentives do more good than harm: a checklist. *BMJ.* 2012;345:e5047.

[PubMed](#)

Gómez G, Carrillo E. Evaluación del cumplimiento de objetivos en los sistemas de incentivación del desempeño. Barcelona: Agencia Valenciana de Salud. Antares Consulting; 2009.

Gomez-Mejia LR, Berrone P, Franco SM. Compensation and organizational performance. New York: ME Sharpe; 2010.

Goodman DC. Unwarranted variation in pediatric medical care. *Pediatr Clin N Am.* 2009;56:745–55.

Goodman DC, Challener GJ, et al. Tonsillectomy: a procedure in search of evidence. *J Pediatr.* 2012;160(5):814–9.

Gottlieb DJ, Zhou W, Song Y, Andrews KG, Skinner J, Sutherland JM. Prices don't drive regional Medicare spending variations. *Health Aff.* 2010;29(3):537–43.

Govender V, Mooney G. People's policies for the health of the poor globally. *Int J Health Plann Manag.* 2012;27(2):e92–103.

Gray MJA. Perspectives. *Lancet.* 2013;382:200–1.

[PubMed](#)

Grytten J, Sørensen R. Practice variation and physician-specific effects. *J Health Econ.* 2003;22:403–18.

[PubMed](#)

Guthrie J, English L. Performance information and programme evaluation in the Australian public sector. *Int J Public Sect Manag.* 1997;10(3):154–64.

Hackbarth G, Reischauer RJD, Mutti A. Collective accountability for medical care – toward bundled Medicare payments. *N Engl J Med*. 2008;359(1):3-5.

[PubMed](#)

Hibbard J, Stockard J, Tusler M. Does publicizing hospital performance stimulate quality improvement efforts? *Health Aff*. 2003;22:84–94.

Hibbard J, Stockard J, Tusler M. Hospital performance reports: impact on quality, market share, and reputation. *Health Aff*. 2005;24(4):1150–60.

Hilarion P, Sunñola R, Groene O, Vallejo P, Herrera E, Saura RM. Making performance indicators work: the experience of using consensus indicators for external assessment of health and social services at regional level in Spain. *Health Policy*. 2009;90(1):94–103.

[PubMed](#)

Hood C. The “new public management” in the 1980s: variations on a theme. *Acc Organ Soc*. 1995;20(3):93–109.

Institute of Medicine (IOM). *Best care at lower cost: the path to continuously learning in health care in America*. Washington, DC: The National Academies Press; 2012.

Jackson PM. Public service performance evaluation: A strategic perspective. *Public Money & Management*. 1993;13:9–14.

Jones C, Dewing I. The attitudes of NHS clinicians and medical managers towards changes in accounting controls. *Financ Account Manag*. 1997;13(3):261–80.

Jones G. Indicators-based systems of performance management in the national Health Service: a comparison of the perceptions of local – and national – level managers. *Health Services Management Research*. 2000;13(1):16–26.

Judge K, Judge K, Platt S, Costongs C, Jurczak K. *Health inequalities: a challenge for Europe. An independent expert report commissioned by the UK Presidency of the EU*. London; 2006.

Kelley E, Hurst J. Health care quality indicators project: conceptual framework, OECD health working papers, vol. 23. Paris: OECD; 2006.

Killoran A, Swann C, Kelly MP. Public health evidence: tackling health inequalities. Oxford /New York: Oxford University Press; 2006.

Kloot L, Martin J. Strategic performance management: a balanced approach to performance management issues in local government. *Manag Account Res.* 2000;11(2):231–51.

Kolstad J. Information and quality when motivation is intrinsic: evidence from surgeon report cards, *American Economic Review* 2013;103(7):2875–2910.

Lapsley I. Accounting and the new public management: instruments of substantive efficiency or a rationalising modernity? *Financ Account Manag.* 1999;15(3–4):201–7.

Lee TH, Meyer GS, Brennan TA. A middle ground on public accountability. *N Engl J Med.* 2004;350:2409e12.

Légaré F, Turcotte S, Stacey D, Ratté S, Kryworuchko J, Graham ID. Patients' perceptions of sharing in decisions: a systematic review of interventions to enhance shared decision making in routine clinical practice. *Patient.* 2012;5(1):1–19.

[PubMed](#)

Linard K, Basset M, Dvorsky L, Yoon J. A dynamic balanced scorecard template for public sector agencies. In: 19th international conference of the system dynamics society. Atlanta, 23–27 June 2000.

Lohr K, Brook R. Effect of cost sharing on use of medically effective and less effective care. *Med Care.* 1986;29:S31–8.

Lomas J, Anderson GM, Domnick-Pierre K, Vayda E, Enkin MW, Hannah WJ. Do practice guidelines guide practice? The effect of a consensus statement on the practice of physicians. *N Engl J Med.* 1989;321(19):1306–11.

[PubMed](#)

Lorden A, Coustasse A, Singh KP. The balanced scorecard framework – a case study of patient and employee satisfaction: what happens when it does not work as planned? *Health Care Manag Rev.* 2008;33(2):145–55.

Lucas FL, Sirovich BE, Gallagher PM, Siewers AE, Wennberg DE. Variation in cardiologists propensity to test and treat: is it associated with regional variation in utilization? *Circ Cardiovasc Qual Outcomes.* 2010;3(3):253–60.

[PubMedCentral](#) [PubMed](#)

Mackenbach JP, Stronks KA. A strategy for tackling health inequalities in the Netherlands. *BMJ.* 2002;325:1029–32.

[PubMedCentral](#) [PubMed](#)

Mackenbach JP, Meerding WJ, Kunst AE. Economic implications of socio-economic inequalities in health in the European Union. Rotterdam: Ed E.C.H.a.C.P. Directorate-General, European Commission Health and Consumer Protection, Directorate-General; 2007.

Majeed A, Eliahoo J, Bardsley M, Morgan D, Bindman AB. Variation in coronary artery bypass grafting, angioplasty, cataract surgery, and hip replacement rates among primary care groups in London: association with population and practice characteristics. *J Public Health Med.* 2002;24(1):21–6.

[PubMed](#)

Manning WG, Newhouse JP, Duan N, Keeler EB, Leibowitz A. Health insurance and the demand for medical care: evidence from a randomized experiment. *Am Econ Rev.* 1987;77:251–77.

[PubMed](#)

Mannion R, Davies H. Reporting health care performance: learning from the past, prospects for the future. *J Eval Clin Pract.* 2002;8:215–28.

[PubMed](#)

Marengo L, Pasquali C. How to get what you want when you do not know what you want: a model of incentives, organizational structure, and learning. *Organ Sci.* 2012; 23(5):1298–310, ISSN 1047–7039 (print).

Marmot M. Social determinants of health inequalities. *Lancet.* 2005;365(9464):1099–104.

[PubMed](#)

Marshall M, Øvretveit J. Can we save money by improving quality? *BMJ Qual Saf*. 2011.
doi:10.1136/bmjqs.2010.050237.

[PubMedCentral](#)

McClellan M, McKethan AN, Lewis JL, Roski J, Fisher ES. A national strategy to put accountable care into practice. *Health Aff*. 2010;29(5):982–90.

McNair C, Leibfried K. *Benchmarking: a tool for continuous improvement*. New York: Harper Business; 1992.

McPherson K, Wennberg JE, Hovind O, Clifford P. Small area variation in the use of common surgical procedures: an international comparison of new England, England and Norway. *N Engl J Med*. 1982;307:1310–14.

[PubMed](#)

McPherson K, Downing A, Buirski D. *Systematic variation in surgical procedures and hospital admission rates*. PHP publication, no 23. London: London School of Hygiene and Tropical Medicine; 1996.

Mladovsky P, Srivastava D, Cylus J, Karanikolos M, Evetovits T, Thomson S, McKee M. Health policy in the financial crisis. *Eurohealth*. 2012;18:3–6.

Mooney G. Funding aboriginal health care: the case for a new initiative. *Aust N Z J Public Health*. 1996;20(6):564–5.

[PubMed](#)

Mooney G. Vertical equity in health care resource allocation. *Health Care Anal*. 2000;8(3):203–15.

[PubMed](#)

Musgrove P. Health insurance: the influence of the beveridge report. *Bull World Health Organ*. 2000;78(6):845–6.

[PubMedCentral](#) [PubMed](#)

Neri S. The development of health care policy and the new governance of the NHS in Italy. In: The future of the welfare state: paths of social policy innovation between constraints and opportunities. Urbino, Sept 2009.

Nuti S. La valutazione della performance in Sanità. Bologna: Edizioni Il Mulino; 2008.

Nuti S, Seghieri C. Is variation management included in regional healthcare governance systems? Some proposals from Italy. Health Policy. 2013. doi:10.1016/j.healthpol.2013.08.003.

[PubMed](#)

Nuti S, Vainieri M. Managing waiting times in diagnostic medical imaging. Bmj Open. 2012;2:e001255. doi:10.1136/bmjopen-2012-001255.

[PubMedCentral](#) [PubMed](#)

Nuti S, Vainieri M. Federalismo fiscale e riqualificazione del Servizio Sanitario Nazionale. Un binomio possibile. Bologna: Edizioni Il Mulino; 2011, n.pag.238.

Nuti S, Vainieri M, Bonini, A. Managing gaming phenomena in the Tuscan performance evaluation system. In: 24thPCSI conference, Lisbon 8–11 Oct 2008.

Nuti S, Vainieri M, Bonini A. Disinvestment for reallocation: a process to identify priority in healthcare. Health Policy. 2010;95:137–43.

[PubMed](#)

Nuti S, Daraio C, Speroni C, Vainieri M. Relationship between efficiency and performance in healthcare. Int J Qual Health Care. 2011;23(3):324–30.

[PubMedCentral](#) [PubMed](#)

Nuti S, Seghieri C, Vainieri M, Zett S. Assessment and improvement of the Italian healthcare system: first evidences from a pilot national performance evaluation system. J Healthc Manag. 2012;57(3):182–98.

[PubMed](#)

Nuti S, Seghieri C, Vainieri M. Assessing the effectiveness of a performance evaluation system in the public health care sector: some novel evidence from the Tuscany region experience. J Manag Gov. 2013;17(1):59–69.

O'Connor AM, Llewellyn-Thomas HA, Flood AB. Modifying unwarranted variations in health care: shared decision making using patient decision aids. *Health Aff (Millwood)*. 2004; Suppl Variation:VAR63-72.

OECD. Strategic response to the financial and economic crisis. Contributions to the global effort. 2009. <http://www.oecd.org/economy/42061463.pdf>

OECD. Health care systems: getting more value for money. OECD economics department policy notes, no. 2; Paris: OECD Publishing; 2010.

OECD. Restoring public finances, Special issue of the OECD journal on budgeting, vol. 2. Paris: OECD Publishing; 2011.

OECD. Health at a glance 2013: OECD indicators. Paris: OECD Publishing; 2013.
doi:10.1787/health_glance-2013-en.

Oliver A, Messialo E. Equity of access to health care: outlining the foundations for action. *J Epidemiol Community Health*. 2004;58:655–8.
[PubMedCentral](#) [PubMed](#)

Ouchi W. Markets, bureaucracies and clans. *Adm Sci Q*. 1980;25:129–41.

Øvretveit J. Does improving quality save money? A review of evidence of which improvements to quality reduce costs to health service providers. London: The Health Foundation; 2009.

Parry G, Humphries G. Health and the financial crisis: a complex diagnosis, *Bulletin of the world health organization*, vol. 87. Geneva: World Health Organization; 2009.

Pearson SD, Schneider EC, Kleinman KP, Coltin KL, Singer JA. The impact of Pay-for-performance on health care quality in Massachusetts. 2001–2003. *Health Aff*. 2008;27(4):1167–76.

Pink GH, McKillop I, Schraa EG, Preyra C, Montgomery C, Baker R. Creating a balanced scorecard for hospital system. *Health Care Financ*. 2001;2(3):1–20.

Pinnarelli L, Nuti S, Sorge C, Davoli M, Fusco D, Agabiti N, Vainieri M, Perucci CA. What drives hospital performance? The impact of comparative outcome evaluation of patients admitted for hip fracture in two Italian regions. *BMJ Qual Saf.* 2012;21:127–34.

[PubMed](#)

PNE. 2013; <http://95.110.213.190/PNEed13/>.

Pollitt CJ. Looking ahead. *Public Administration* 1985;63(4):379–381.

Pollitt CJ. Beyond the managerial model: the case for broadening performance assessment in government and the public services. *Financial Accountability & Management.* 1986;2(3):155–170.

Right Care. The NHS atlas of variation in healthcare. NHS Right Care; 2011.

Semuraga B, Ross-Degnan D, Avery AJ, Elliott RA, Majumdar SR, Zhang F, Soumerai SB. Effect of pay for performance on the management and outcomes of hypertension in the United Kingdom: interrupted time series study. *BMJ.* 2011;342:d108. doi:10.1136/bmj.d108.

Sheaff R, Pilgrim D. Can learning organizations survive in the newer NHS? *Implement Sci.* 2006;1:27.

[PubMedCentral](#) [PubMed](#)

Shortell SM, Casalino LP. Health care reform requires accountable care systems. *JAMA.* 2008;300(1):95.

[PubMed](#)

Shortell SM, Kaluzny A. *Health care management: organization design and behavior.* 5th ed. Clifton Park: Thomson Delmar Learning; 2006.

Simons R. Control in an age of empowerment. Harvard: *Harvard Business Review*; 1995.

Smith PC. *Measuring up. Improving health systems performance in OECD countries.* Ottawa: OECD; 2002.

Street A, Vitikainen K, Bjorvatn A, Hvenegaard A. Introducing Activity-Based Financing: A review of experience in Australia, Norway, Denmark and Sweden. CHE Research Paper 30. York: Centre for Health Economics, University of York. 2007.

Stuckler D, Basu S, Suhrcke M, Coutts A, McKee M. The public health effect of economic crises and alternative policy responses in Europe: an empirical analysis. *Lancet*. 2009;374:315–23.

[PubMed](#)

Stukel TA, Glazier RH, Schultz SE, Guan J, Zagorski BM, Gozdyra P, Henry DA. Multispecialty physician networks in Ontario. *Open Med*. 2013;7(2):e40.

[PubMedCentral](#) [PubMed](#)

Suhrcke M, et al. *Chronic disease: an economic perspective (with annex)*. London: Oxford Health Alliance; 2006.

Swaminathan S, Mor V, Mehrotra R, Trivedi A. Medicare's payment strategy for end-stage renal disease now embraces bundled payment and Pay-for-performance to cut costs. *Health Aff*. 2012;31(9):2051–8.

Tchouaket ÉN, Lamarchel PA, Goulet L, Contandriopoulos AP. Healthcare system performance of 27 OECD countries. *Int J Health Plann Manag*. 2012;27(2):104–29.

Thomson S, Foubister T, Mossialos E. Can user charges make health care more efficient? *BMJ*. 2010;341:c3759.

[PubMed](#)

Town R, Wholey DR, Kralewski J, Dowd B. Assessing the influence of incentives on physicians and medical groups. *Med Care Res Rev*. 2004;61 Suppl 3:80S–118S.

[PubMed](#)

Trisolini MG. Theoretical perspectives on pay for performance. In: Cromwell J, Trisolini MG, Pope GC, Mitchell JB, Greenwald LM, editors. *Pay for performance in health care: methods and approaches*. Research Triangle Park: RTI Press; 2011.

Van Herck P, De Smedt D, Annemans L, Remmen R, Rosenthal MB, Sermeus W. Systematic review: effects, design choices, and context of pay-for-performance in health care. *BMC Health Serv Res*. 2010;10:247.

[PubMedCentral](#) [PubMed](#)

Van Hertem L, Van De Water HPA. Health policies on target? Review on health target setting in 18 European countries. *Eur J Pub Health*. 2000;10 Suppl 4:11–6.

Van Kemenade YW. Health care in Europe 2007: the finance and reimbursement systems of 11 European countries. Maarssen: Elsevier Gezondheidszorg; 2007.

Veillard J, Huynh T, Ardal S, Kadandale S, Klazinga NS, Brown AD. Making health system performance measurement useful to policy makers: aligning strategies, measurement and local health system accountability in Ontario. *Healthcare Policy*. 2010;5(3):49.

[PubMedCentral](#) [PubMed](#)

Wagstaff A, van Doorslaer E, Calonge S, Christiansen T, Gerfin M, Gottschalk P, Janssen R, et al. Equity in the finance of health care: some international comparisons. *J Health Econ*. 1992;11:361–87.

[PubMed](#)

Wennberg JE. Understanding geographic variations in health care delivery (editorial). *N Engl J Med*. 1999;340:52–5.

[PubMed](#)

Wennberg JE. Time to tackle unwarranted variations in practice. *BMJ*. 2011;342:d1513.

Wennberg JE, Gittelsohn A. Small area variations in health care delivery. *Science*. 1973;182(4117):1102–8.

[PubMed](#)

Wennberg JE, Fisher ES, Skinner JS. Geography and the debate over Medicare reform. *Health Aff*. 2002. doi:10.1377/hlthaff.w2.96.

WHO. The world health report 2000. Health systems: improving performance. Geneva: World Health Organization; 2000.

WHO. Health systems performance assessment: debates, methods and empiricism. Geneva: World Health Organization; 2003.

World Health Organization. Targets for health for all. The health policy for Europe. Copenhagen: WHO Regional Office for Europe; 1992.

Wynia M. The risks of rewards in health care: how pay-for performance could threaten, or bolster, medical professionalism. *J Gen Intern Med.* 2009;24(7):884–7.

[PubMedCentral](#) [PubMed](#)

Yang M, Tung Y. Using path analysis to examine causal relationships among balanced scorecard performance indicators for general hospitals: the case of a public hospital system in Taiwan. *Health Care Manag Rev.* 2006;31(4):280–8.

Zelman WN, Pink GH, Matthias CB. Use of the balanced scorecard in health care. *J Health Care Finance.* 2003;29(4):1–16.

[PubMed](#)

Zuckerman S, Waidmann T, Berenson R, Hadley J. Clarifying sources of geographic differences in Medicare spending. *N Engl J Med.* 2010;363(1):54–62.

[PubMed](#)