Value of new performance information in healthcare: evidence from Japan

Susanna Gallani, Takehisa Kajiwara & Ranjani Krishnan

International Journal of Health Economics and Management

ISSN 2199-9023
Volume 20
Number 4

Int J Health Econ Manag. (2020)
20:319-357
DOI 10.1007/s10754-020-09283-1

Springer

Your article is protected by copyright and all rights are held exclusively by Springer Science+Business Media, LLC, part of Springer Nature. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".

Springer

Author's personal copy

International Journal of Health Economics and Management (2020) 20:319–357
https://doi.org/10.1007/s10754-020-09283-1

RESEARCH ARTICLE

Value of new performance information in healthcare: evidence from Japan

Susanna Gallani[1] · **Takehisa Kajiwara**[2] · **Ranjani Krishnan**[3]

Received: 27 November 2019 / Accepted: 7 August 2020 / Published online: 18 August 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

Mandatory measurement and disclosure of outcome measures are commonly used policy tools in healthcare. The effectiveness of such disclosures relies on the extent to which the new information produced by the mandatory system is internalized by the healthcare organization and influences its operations and decision-making processes. We use panel data from the Japanese National Hospital Organization to analyze performance improvements following regulation mandating standardized measurement and peer disclosure of patient satisfaction performance. Drawing on value of information theory, we document the *absolute value* and the *benchmarking value* of new information for future performance. Controlling for ceiling effects in the opportunities for improvement, we find that the new patient satisfaction measurement system introduced positive, significant, and persistent mean shifts in performance (*absolute value* of information) with larger improvements for poorly performing hospitals (*benchmarking value* of information). Our setting allows us to explore these effects in the absence of confounding factors such as incentive compensation or demand pressures. The largest positive effects occur in the initial period, and improvements diminish over time, especially for hospitals with poorer baseline performance. Our study provides empirical evidence that disclosure of patient satisfaction performance information has value to hospital decision makers.

Keywords Value of information · Feedback · Patient satisfaction · Healthcare

JEL Classification D83 · I18 · K32 · M16 · M41

✉ Susanna Gallani
sgallani@hbs.edu

Takehisa Kajiwara
kajiwara@people.kobe-u.ac.jp

Ranjani Krishnan
krishnan@broad.msu.edu

1 Harvard Business School, 369 Morgan Hall, 15 Harvard Way, Boston, MA 02163, USA

2 Graduate School of Business Administration, Kobe University, 2-1 Rokkodai, Nada-ku, Kobe 657-8501, Japan

3 The Eli Broad College of Business, Michigan State University, N207 North Business College Complex, 632 Bogue St, East Lansing, MI 48824, USA

Author's personal copy

Introduction

Improving healthcare outcomes, quality, cost, access, and patient experience are topics of high policy interest in many countries. The effectiveness of any healthcare policy depends, in large part, on the extent to which it influences behaviors within the organizations subject to the policy. Systematic collection and disclosure of reliable, consistent, and comparable health outcomes measures is often proposed as a key policy tool. Whether the information provided by these performance measures is internalized by the organization and influences outcomes is an empirical question that we address in this study.

Economic and decision theory alike posit that new information generated by mandated measurement systems improves decisions. Performance information reduces the uncertainty about the mapping between effort and outcomes and influences subsequent allocations of effort across actions and activities to maximize desired outcomes. Performance information disclosure can therefore enhance decision making of providers and patients alike, instill participative decision making, and encourage market-based discipline and reform (Fung 2008; Lindenauer et al. 2007). On the other hand, excessive or ambiguous information could create confusion, adversely affect patient choices, and encourage healthcare providers and managers to game performance measures (Dranove et al. 2003; Weissman et al. 2005). Thus, in a healthcare setting, the effect of performance disclosures on future performance is an empirical question.

In the hospital industry, patient satisfaction is especially important and forms a central feature of patient centered medicine (Bardes 2012). As in many high-contact service environments, satisfaction is a key driver of success (Goldstein 2003). We analyze patient satisfaction panel data from 145 public Japanese hospitals over a period of 8 years (2004–2011) subsequent to the introduction of regulation mandating the periodic collection and peer disclosure of standardized patient satisfaction measurements. We develop hypotheses about the impact of mandatory performance measurement on performance improvement and its persistence. We build on the predictions of value of information (VOI) theory, which defines information value as the difference between the expected utility of an action based on the posterior probability given a new information set, and the expected utility of the action given only the prior information set (Pratt et al. 1995).[1,2] The improved posterior distribution should guide decision makers toward better choices of effort allocation to maximize desired outcomes. In particular, we posit that the information generated by the new policy has both *absolute value*—i.e., it improves the mapping between effort and performance for the individual hospital—and *benchmarking value*—i.e., it signals the hospital's performance relative to its peer group.

Our setting exhibits three ideal characteristics to explore our research question. First, the new information was not previously collected in any systematic way by the hospitals or any other entity—that is, the information is new. Second, patient satisfaction performance is not tied to incentive compensation or other pecuniary payoffs for any member of the healthcare provider organization. Third, institutional characteristics related to the implementation of the new policy make it virtually impossible for individual managers or clinicians to game the measure by improving reported but not actual performance.

---

1 Yokota and Thompson (2004) provide a review of VOI models in healthcare.

2 Although VOI is sometimes interpreted rather narrowly as the amount a decision maker would be willing to *pay* for higher quality information, the analytical models of VOI are generic and refer to "value" in a flexible sense that allows for nonfinancial interpretations (Bromwich 2007; Demski 1972; Raiffa 1968).

Author's personal copy

We leverage on a 2004 regulation which required measurement of patient satisfaction for hospitals belonging to the Japanese National Hospital Organization (NHO). Pursuant to the new policy, a neutral external agency employed by the NHO conducted an annual survey of inpatients and outpatients about their satisfaction with several aspects of their hospital experience, including medical treatments and procedures, physician and staff behavior and attitudes, and hospital infrastructure. The results of the survey containing performance information on the level as well as relative rank of individual member hospitals were disseminated to all hospitals within the NHO every year.

We analyze patient satisfaction panel data from all 145 NHO-member hospitals over a period of 8 years (2004–2011). We first conduct a factor analysis of the survey responses and identify several satisfaction constructs, as well as an overall satisfaction factor. We then examine whether patient satisfaction information leads to improvements in subsequent performance for each construct, and whether there are differences in the extent of improvement based on initial relative performance. Our analyses control for ceiling effects—that is, opportunities for further improvement are lower for hospitals that begin at greater levels of patient satisfaction and diminish as hospitals improve over time. Results indicate that new patient satisfaction information has an *absolute* performance effect in that satisfaction levels improve in the post-implementation period and do not regress to original levels, which suggests that the intervention succeeded in generating a meaningful and persistent favorable mean shift in overall performance.[3] We also find evidence of the *benchmarking* effect of new information, in that hospitals that ranked lower in the baseline year (defined as hospitals in the lowest quartile of performance in 2004) exhibit greater improvements in performance following the receipt of survey results, controlling for ceiling effects. We conclude that the new information about performance relative to a relevant peer group provides valuable feedback for directed effort.

In additional tests, we find that performance improvements are of larger magnitude in the year immediately following the introduction of the measurement, when the information is new, compared to subsequent years, especially for hospitals that were performing poorly in the baseline year. Thus, the performance effects of information updates follow a path of diminishing returns relative to new information. We further examine the relation between the first change in performance and the persistence of performance improvements over time. This analysis sheds light on the nature of the changes enacted by hospitals in response to the new information about patient satisfaction. We find that the initiatives implemented in the first year after the intervention are predictive of persistent performance improvements. Finally, we validate that patient satisfaction improvements are not driven by motives such as maximizing compensation or hospital payoffs.

Our study makes several contributions. First, we isolate the feedback role of new performance information in a healthcare setting. Prior research, such as Evans et al. (1997) finds that mandated public disclosures of hospital mortality performance is associated with subsequent improvements in mortality for hospitals that were performing poorly during the initial period. However, in Evans et al. (1997) hospitals were already collecting mortality information and the only change was to make the disclosure public. That is, the *information* was not new to the decision maker but only the *disclosure* was new. Second, we attempt to disentangle the value of information about *absolute* performance (i.e., relative to the measurement scale) and the value of information about *relative* performance (i.e., relative to a

[3] This finding also reduces the concern that regression to the mean might be an alternative explanation for our findings.

Author's personal copy

reference group). Third, a wealth of research has explored the importance of nonfinancial measures such as client satisfaction in driving future financial performance (e.g. Ittner and Larcker 1998; Nagar and Rajan 2005). These studies have stressed the importance of managerial compensation tied to nonfinancial information (Banker et al. 2000). Our evidence suggests that firms would benefit from designing planning and feedback systems where decision makers obtain performance information even if it is disassociated from compensation or other financial rewards. Finally, our study has important policy implications. The current healthcare policy discourse highlights preferences for public disclosure of quality information, under the assumption that public disclosures will promote informed choices by patients and give rise to market-based pressures on hospitals to improve quality. Our study shows that quality information can lead to quality improvements even if it is only provided to hospital medical and managerial staff.

Institutional background of the Japanese National Hospital Organization

The Japanese National Hospital Organization (NHO) is the oversight agency for Japanese government hospitals. Formally established in 2002 and headquartered in Tokyo, the NHO is the largest hospital network in Japan comprising of 145 hospitals (about 3.5% of the total number of hospitals in Japan at the time).[4] Like other Japanese Independent Administrative Institutions (IAI), the NHO is the result of the separation between political and operational responsibilities for public services. IAIs are non-profit organizations that have considerable autonomy in managing their own budgets. They are provided with adequate decision autonomy and encouraged to apply private sector management principles to their practices.

The mission of the NHO is to operate as a safety net allowing all citizens access to treatment and care "even for specialties that are difficult to access for reasons such as lack of experienced providers or system establishment or specialties that are not profitable in the private sector" (NHO Guidebook, page 8). NHO hospitals are grouped into two categories: general hospitals, which provide a wide range of services at their discretion, and sanatoriums, which, in addition to offering services similar to general hospitals, are required to supply particularly expensive and risky medical services, including care for "tuberculosis, severe motor and cognitive disabilities, and incurable neuromuscular diseases, including muscular dystrophy" (NHO Guidebook, page 8). Both types of hospitals provide inpatient as well as outpatient treatment. Funding for NHO hospitals comes from three sources: patients' copayments for service rendered, reimbursements by the National Health Insurance or Employees' Health Insurance, and public funding through government grants and subsidies.[5] Patient copayments are received directly by the hospital, while insurance reimbursements are received through a claims process. Public funding allocation

[4] Source: Guidebook of the National Hospital Organization—www.nho.hosp.go.jp.

[5] Health Insurance in Japan is compulsory for all citizens and can be obtained either through the employer (Employees' Health Insurance) or, in the case of self-employed individuals and students, through the National Health Insurance system. Special insurance programs are in place for elderly citizens (over 75 years). Patients pay about for 30% of the cost of medical services, with the remaining 70% being reimbursed to the hospital by the insurer or the government. Medical costs exceeding the equivalent of $600 in a month are fully reimbursed by the insurer or the government. Other than minor cost of living adjustments, these numbers have been steady since the year 2000.

Author's personal copy

to each hospital within the NHO is dependent upon periodic performance evaluations of medical outcomes (e.g. mortality and morbidity) and assessment of the hospital's need for resources. Prices for healthcare services are centrally determined by the Japanese government. Price is, therefore, not a driver of patients' choice of healthcare provider.

After winning the 2003 general elections, the incumbent Liberal Democratic Party enacted a series of initiatives to promote administrative reform. As part of a renewed impetus to improve healthcare, the NHO introduced in 2004 an annual patient satisfaction survey for every member hospital "to assess patients' perspectives of the services provided" (NHO Guidebook, page 12) and improve healthcare quality in NHO hospitals. Dr. Kunio Nakai, NHO board member and head of South Wakayama Medical Center, recalled, "There was a concern that national hospitals at the time were not considering patients first. […] We were transitioning into a more private business structure. Therefore, we needed to make significant changes in the way we ran the hospitals, and one of the first steps in that change was to improve patient satisfaction."[6]

Starting in 2004, patients treated in NHO hospitals were required to complete a standardized questionnaire, which assessed their satisfaction with critical aspects of their hospital experience, including medical treatment, the behavior of the staff, and the quality of the infrastructure. The survey was conducted by an independent university research agency hired by the NHO and unrelated to any of its hospitals. The university research agency compiled and analyzed the results of the annual survey, and distributed feedback reports to all member hospitals. The feedback reports included the average scores of each hospital on every patient satisfaction category and each hospital's ranking within the NHO. There were no monetary incentives tied to patient satisfaction performance.[7]

Theory and hypotheses development

Value of information

The value of information theory framework (Demski 1972; Bromwich 2007) motivates our study. Suppose a decision maker is considering an action choice from a vector of potential actions $A\{a\}$, and a potential set of uncertain states $S=\{s\}$. The utility of a particular action is $U(s, a)$, where $\emptyset(s)$ is the probability specification of the uncertain states. The decision maker optimally selects the action a^*, which maximizes expected utility, that is $E(U|a^*) = \max_{a \in A} \int_s U(s,a)\emptyset(s)$. The action choice involves the assessment of the expected utility of each action and state combination.

The subjective probability distribution of the likelihood of the states $\emptyset(s)$ is a function of the information made available to the decision maker at the time of the decision by an information system η (Feltham 1968). Suppose the decision maker obtains an additional information signal y, which allows for an improved assessment of the likelihood of the state $\emptyset(s)$. The signal y has value if it changes the decision, relative to the decision that would have been made without the signal, leading to greater levels of utility. The expected

6 The research team interviewed Dr. Kunio Nakai in October of 2017.

7 Physicians and medical staff at the NHO are compensated on a fixed wage basis and are not provided performance-contingent bonuses. Physicians and staff obtain raises based on general macro-economic conditions. Section 4 examines physician compensation at NHO hospitals in greater detail.

Author's personal copy

utility including the new information signal is therefore $E\left(U|y, \eta, a_y^*\right) = \max_{a \in A} \int_s U(s, a)\emptyset(s|y, \eta)$, where $\emptyset(s|y, \eta)$ is the revised probability distribution after the receipt of the new information signal y. The expected value of signal y is the difference between $E\left(U|y, \eta, a_y^*\right)$ and $E(U|a^*)$.[8] If this expected value is positive, then the new information has value to the decision maker.

Performance effect of patient satisfaction information

Theory predicts that that healthcare providers derive utility from non-financial outcomes such as patient satisfaction, independent of monetary compensation (Arrow 1963; Kolstad 2013). At the same time, however, analytical research on multitasking (Holmstrom and Milgrom 1991) predicts that managers faced with competing priorities and demands on their time tend to allocate more effort to those dimensions of performance that are more directly linked to monetary rewards. Since the inception of the IAI, hospital administrators in our setting are evaluated annually based on medical and managerial aspects of performance that do not include patient satisfaction. Therefore, it is plausible that new information on patient satisfaction might not be salient enough to generate a change in behaviors.

The introduction of the patient satisfaction survey provided hospitals with an additional information system η that contained two new signals—*absolute* patient satisfaction level (y_1) and patient satisfaction *relative* to peer hospitals (y_2). We posit that the healthcare provider's utility from patient satisfaction is a function of both the hospital's individual performance level relative to the measurement scale (*absolute* value of information) and its performance relative to peers (*benchmarking* value of information). We therefore test whether the new information provided by each signal has value, in that it affects the choice of actions leading to improvements in subsequent performance.

Absolute value of patient satisfaction information

NHO sources indicated that before 2004, neither the NHO nor individual member hospitals collected systematic information on patient satisfaction, as it was generally not a priority for hospitals in the system. Consequently, even those hospitals that were interested in assessing patient satisfaction only had, at best, noisy priors about their own performance. Individuals tend to be overconfident about their ability and overestimate their effort levels (Camerer and Lovallo 1999; Benoit and Dubra 2011; Kruger and Dunning 1999, 2002); thus, individual hospitals likely concluded that their patient satisfaction performance was (a) on the higher end of an absolute scale, and (b) above average relative to their peers.

In line with a standard assumption in decision theory, we posit that decision makers in hospitals are Bayesian, i.e., they use new information to update their prior beliefs (Pratt et al. 1995). The new information about absolute performance would therefore influence strategy and effort allocations required to move from utility function $E(U|a^*)$ to $E\left(U|y, \eta, a_y^*\right)$. The information on individual performance *level* (y_1) provided a more precise signal of *absolute* performance, reduced uncertainty in the mapping between actions

[8] Although VOI is sometimes interpreted rather narrowly as the amount a decision maker would be willing to *pay* for higher quality information, analytical VOI models are generic and refer to "value" in a flexible sense that allows for non-pecuniary interpretations (Bromwich 2007; Demski 1972; Raiffa 1968).

Author's personal copy

and outcomes, permitted belief revision, and increased employees' knowledge (Sprinkle 2003). The new performance signal enhanced employees' ability to make guided effort choices that were better suited to the circumstances (Morris and Shin 2002; Bandura and Jourden 1991; Ederer 2010) and devise corresponding interventions to improve (Anderson and Kimball 2019). This leads to the following prediction about the *absolute performance effect* of nonfinancial performance information:

H1 New information about patient satisfaction performance improves subsequent performance.

Benchmarking value of patient satisfaction information

The second signal contained in the patient information system was patient satisfaction *relative* to peer hospitals (y_2). The relative signal y_2 has two benchmarking uses, one related to its effect on healthcare providers' utility function, and the other related to its feedback value. Economic theory recognizes referent performance as an important driver of utility (Kolstad 2013; Sugden 2003; Lant and Hewlin 2002). In a reference-dependent utility model, decisions are influenced both by the expected outcome of the decision and by a reference point, which could be performance of a peer or competitor (Sugden 2003). Public policy intervention, such as the NHO's introduction of patient satisfaction measurement and reporting, typically aim to improve or sustain performance across all organizations subject to the stimulus, but, in particular, to improve performance of those that operate significantly sub-standard.

If a relative performance signal indicating poorer performance relative to a referent group is valuable, it then prompts decision makers to increase effort and to search for new strategies that can improve the rankings (Bandura and Jourden 1991), especially if decision makers have flexibility to respond to the new information (Abernethy and Bouwens 2005). Organizational theory posits that decision makers pay more attention to activities that fail to meet targets compared to those that succeed (Levitt and March 1988). In our setting, the noise reduction value of relative feedback is higher for poorly performing hospitals because they likely expected to be above average in the pre-information period, and therefore the relative information represents an unpleasant surprise.[9] Therefore, poor relative performance can be a higher motivator of performance than good performance. For example, Ramanarayanan and Snyder (2012) find that information disclosure in the dialysis industry is associated with reduction in mortality for poorly performing firms, but do not find comparable effects for highly performing firms. On the other hand, poor relative performance could also be demotivating if poor performers perceive the gap between themselves and the top performers to be too wide to bridge. Casas-Arce and Martínez-Jerez (2009) find that the introduction of a relative performance evaluation system (i.e., a contest) in a retail setting introduces complacency among the high performers, who reduce effort, and leads poor performers to give up. To be an effective motivator for poor performers, relative performance information needs to offer information about the feasibility of performance improvements required to close the gap with better performers.

[9] Prior literature finds that in the absence of information, individuals and firms tend to hold optimistic beliefs about their ability and therefore are overconfident about their performance relative to competitors (Kahnemann et al. 1982).

Author's personal copy

The relative performance signal y_2 has feedback value for the decision maker to the extent that it provides an assessment of the range of possible outcomes that are achievable (i.e., if peer organizations are performing better, then those higher levels of performance must be achievable), as well as opportunities for learning. y_2 reduces idiosyncratic uncertainty creating a level field to assess performance. y_2 also increases the accuracy of the posterior belief function about the mapping between effort and output relative to the organization's peer group. This serves as motivation for poor performers to increase effort to improve performance (Ederer 2010). Based on the above, we test the following hypothesis:

H2 Lower baseline performance on patient satisfaction measures is associated with higher magnitude of subsequent improvements.

Diminishing returns from patient satisfaction information

Because the patient satisfaction survey is administered every year, the Bayesian belief revision of decision makers is likely more significant in the first year, when the information is completely new. Subsequent iterations of the survey likely offer decision makers opportunities to revise and fine tune their performance strategies. Therefore, we explore whether *new* performance information has greater value relative to performance information *updates*.

Theory posits that value of information depends on the prior distribution that decision makers use to represent the current situation (Raiffa 1968). Before the availability of any standardized information about own and peer performance, decision makers are likely to (a) systematically over-estimate their own performance, (b) exhibit overconfidence about their own performance, and (c) estimate probability distributions that are tighter than the actual distributions (Alpert and Raiffa 1982; Lichtenstein, Fischhoff, and Phillips 1982; Morgan and Henrion 1990; Hammitt and Shlyakhter 1999). This implies that release of the first information signal is likely to cause a greater Bayesian adjustment of decision makers' beliefs (Kolstad 2013) compared to subsequent updates. The idea is that subsequent probability distribution adjustments based on further releases of informative signals would be smaller in magnitude relative to their initial release, when decision makers went from a no-information regime to a quality information regime. In sum, the initial performance signal captures the effect of new information, while subsequent signals capture the effect of information updates. This leads to the following hypothesis:

H3 Patient satisfaction information drives greater improvements relative to the baseline during the initial period compared to subsequent periods.

Methodology

Sample characteristics

Our dataset includes satisfaction information about the entire population of 145 NHO hospitals for the period 2004–2011. The standardized patient satisfaction survey was administered every year during the months of June and July in all NHO hospitals. There are two types of NHO hospitals—general hospitals and sanatoriums. General hospitals are similar to private hospitals and are allowed greater discretion in the choice of healthcare services they offer. Sanatoriums are expected to provide not only the services that are provided by

Author's personal copy

general hospitals, but also special services "that cannot be dealt with properly by Non-National Hospital Organizations due to historical and social reasons."[10] These include treatment of expensive (thus often unprofitable), long-term, risky, or communicable ailments such as tuberculosis, AIDS, Alzheimer's, ALS, complex mental illnesses, and invasive or terminal cancer.

Separate surveys were administered to inpatients and outpatients ("Appendix 1"). Each survey contained ten common questions related to the patient's overall satisfaction ("Appendix 1", Panel A). Inpatient satisfaction was assessed with 19 questions ("Appendix 1", Panel B), and 15 questions related to satisfaction of outpatients ("Appendix 1", Panel C). The survey items related to aspects that were considered to be critical for the delivery of high-quality healthcare services,[11] such as quality of medical treatments, behavior of the staff, quality of infrastructure and facilities, waiting periods, etc. All the questions used a 5-point Likert-type scale, where 1 indicated strong dissatisfaction and 5 indicated strong satisfaction. Data were collected and processed by a university research center, which was unrelated to any of the hospitals in the NHO. Feedback reports were subsequently distributed to each member hospital. These reports contained the average score for each of the questions included in the questionnaires and the ranking of the hospital within the NHO on each item included in the survey. Our dataset consists of the hospital/year level average score for each question included in the questionnaire as well as the associated rankings for each year between 2004 and 2011.[12]

Descriptive statistics for each survey item are reported in "Appendix 2".[13] In general, we note that the distribution of average hospital scores exhibits slightly higher means in the inpatient subsample compared to the outpatients one. However, inpatients are also more likely than outpatients to assign lower scores, as evidenced from the greater variation in the distribution of reported scores.

Variables reduction

Patient satisfaction is a multidimensional construct (Chen 2009). To obtain a measure of the underlying dimensions, we conduct a factor analysis using the patient satisfaction survey data. We use a principal component factor approach with oblique rotation to allow for the possibility of significant correlation among factors. We retain factors with eigenvalues greater than 1. Consistent with the design and administration process of the survey, we conduct separate factor analyses for inpatients and outpatients and, within each patient type, for detailed aspects of satisfaction and for the general satisfaction block of the survey questions. With respect to the general satisfaction construct, we confirm that all items load onto a single factor for each type of patient. General satisfaction items load with weight greater than 0.80 (0.85) for inpatients (outpatients) with a Cronbach's alpha of 0.98 (0.96). Exploratory factor analysis of the detailed satisfaction questionnaires reveals two factors for inpatients and three for outpatients. Based on the items loading onto each factor, we label

[10] National Hospital Organization (Independent Administrative Institution) page 1; http://www.mof.go.jp/english/filp/filp_report/zaito2004e-exv/24.pdf.

[11] Source: Guidebook of the National Hospital Organization—www.nho.hosp.go.jp.

[12] We do not have access to individual patient-level responses.

[13] The surveys include sub-items for each of the 15 (19) questions. After validating that each group of sub-questions loaded on individual factors corresponding to the "header" question, we decided to focus on the 15 (19) header questions in order to ensure we would have sufficient statistical power for our analyses.

Author's personal copy

inpatient factors as *Staff/Treatment* (factor loadings > 0.63; Cronbach's alpha = 0.97) and *Logistics/Infrastructure* (factor loadings > 0.72; Cronbach's alpha = 0.91). For outpatients, we label our factors as *Staff/Treatment* (factor loadings > 0.62; Cronbach's alpha = 0.93), *Administrative Procedures* (factor loadings > 0.67; Cronbach's alpha = 0.81), and *Logistics/Infrastructure* (this factor is composed of one single item corresponding to the direct question about inconvenience of the hospital).[14]

Descriptive statistics and correlations

Table 1 provides descriptive statistics for all hospitals in our population. Because patient satisfaction factors are normalized by construction, the corresponding variables exhibit zero mean and a standard deviation of one. Higher values of the factor scores correspond to higher levels of satisfaction. General hospitals represent 40% of the hospitals included in our population, while sanatoriums constitute the remaining 60%. On average, Japanese prefectures[15] are served by seven hospitals for each 100 thousand inhabitants, and in less than 25% of the cases that number is lower than 2.6. Among expenses, salary and bonuses represent an average of 31% of total costs. Medical revenues represent the predominant source of funding.

Table 2 provides the pairwise Pearson correlation coefficients for the variables included in our dataset. Larger hospitals and general hospitals appear to have greater overall satisfaction among the outpatient population, and lower among the inpatient ones. However, inpatients seem to be generally more satisfied about the quality of the staff and treatment compared to outpatients. Compensation expenses appear to be positively correlated with outpatient overall satisfaction but have no significant correlation with inpatients' satisfaction. Grant revenues are positively correlated with satisfaction with logistics and infrastructure aspects of the outpatients' experience (likely due to greater resources that can be dedicated to renovations of older buildings—a nation-wide initiative promoted by the NHO during the years in our sample period),[16] negatively correlated with satisfaction on administrative procedures, and not significantly associated with any other aspect of satisfaction.

Exploratory analysis of satisfaction components for inpatients and outpatients

Inpatients and outpatients in different organizational settings may differentially weigh the importance of each factor in formulating their assessments of overall satisfaction with the hospital.[17] Therefore, to further validate the soundness of our factor analysis, we examine

[14] Items that cross-loaded on multiple factors were dropped (Ho 2013).

[15] A prefecture is a geographical subdivision of the Japanese territory, conceptually equivalent to a county in the US.

[16] Source: Guidebook of the National Hospital Organization—www.nho.hosp.go.jp.

[17] A survey conducted by the Japanese Ministry of Health, Labor and Welfare during the period of the study explored the major drivers of hospital choice for inpatients and outpatients. The sample consisted of more than 150,000 respondents, randomly selected from the patient population of all Japanese Hospitals. Overall, outpatients (inpatients) identified the following drivers of hospital choice: 38% (34.9%) prior experience at the same hospital, 37.6% (29.9%) physical closeness to their residence, school or place of work, 33.2% (49%) recommendation by doctors, 31.4% (34.7%) kindness of doctors and nurses, and 28.7% (25.5%) size/technology of the hospital. Source: Japanese Ministry of Health, Labour and Welfare. (2011). Patients Behavior Survey, from http://www.mhlw.go.jp/english/new-info/2012.html.

Author's personal copy

the extent to which each of the component factors influences overall inpatient and outpatient satisfaction. Using OLS regressions with heteroskedasticity-robust standard errors clustered by hospital, we estimate the following model, first for the pooled sample of all hospitals and then separately for the general hospitals and sanatoriums:

$$OverallSatisfaction_{i,t} = \alpha + \sum_{j=1}^{k} \beta_j Factor_{j_{i,t}} + \delta_1 Hospital_i + \delta_2 Size_i + \delta_3 Concentration_i + \sum_{t=1}^{T} \gamma_t Year_t + \varepsilon \quad (1)$$

Variable descriptions are provided in "Appendix 3". Results reported in Table 3 indicate that both *Staff/Treatment* and *Logistics/Infrastructure* are significant drivers of overall inpatient satisfaction. Wald tests comparing the magnitude of coefficients associated with each factor indicated that, in our setting, inpatients assign similar weights to these aspects of healthcare services when considering their overall satisfaction with the hospital. However, inpatients of sanatoriums—who likely deal with lengthier and riskier conditions—value their experience with staff and treatment more than the quality of the infrastructure of the provider organization (i.e., the Wald test rejects the null hypothesis that the two coefficients are not different; $p > 0.10$). Similarly, outpatients tend to value equally the three satisfaction factors we identified in our variable reduction effort, with the exception of outpatients in sanatoriums, who assign a significant lower weight to administrative procedures compared to their interactions with the staff and their treatment.

Univariate test of H1: average patient satisfaction performance improvements

H1 predicts a positive effect of patient satisfaction information. Table 4 provides information on the mean inpatient and outpatient satisfaction for each year subsequent to the information release. Statistical tests (*t* test) comparing the mean performance in each year with the baseline (patient satisfaction in 2004) indicate a positive mean shift in satisfaction with staff and treatment for both inpatient and outpatient respondents, as evidenced by a positive difference with the baseline recorded in each year subsequent to the introduction of the survey. In other words, we document a persistent improvement in overall satisfaction. Interestingly, satisfaction with logistics and infrastructure improves in the eyes of inpatients, while it deteriorates for outpatients. While this is likely to be a consequence of prioritization in the renovation programs managed by the NHO, we do not have sufficient information to further explore this finding.[18] Administrative procedures do not appear to change as a result of the intervention. Taken together, our univariate tests suggest that hospitals acted on aspects of healthcare delivery that were more directly impactable by clinicians (e.g. staff behaviors, communication with the patients, empathy), whereas effects of the intervention on other more structural aspects of the patient experience (e.g. infrastructure, administrative procedures) were more difficult to affect.

[18] Patient satisfaction with hospitals' infrastructure is likely negatively impacted by aging buildings that had not been properly maintained during the pre-NHO era. Since 2004, the NHO has invested significant sums, mostly in the form of grants, to remodel and renovate its hospitals with a view to improve patient experience. However, because it is the policy of the NHO to balance their budget each year, and each hospital is responsible for breaking even, actual investments were slow to produce visible results. The disruption caused by renovation activities is likely to have caused the deterioration of patient satisfaction in some cases. Source: Guidebook of the National Hospital Organization—www.nho.hosp.go.jp.

Author's personal copy

Table 1 Descriptive statistics

	N	Mean	SD	p25	p50	p75
Satisfaction factors						
Staff and treatment—outpatient	1148	0.000	1.000	−0.529	0.077	0.653
Administrative procedures—outpatient	1148	0.000	1.000	−0.697	−0.086	0.670
Logistics and infrastructure—outpatient	1148	0.000	1.000	−0.530	0.110	0.644
Staff and treatment—inpatient	1050	0.000	1.000	−0.189	0.228	0.552
Logistics and infrastructure—inpatient	1050	0.000	1.000	−0.575	0.086	0.747
Overall satisfaction—outpatient	1150	0.000	1.000	−0.609	0.007	0.626
Overall satisfaction—inpatient	1076	0.000	1.000	−0.192	0.214	0.517
Hospital characteristics						
Size	1160	4.024	1.384	3.000	3.800	4.810
Concentration	1160	7.003	2.620	5.100	6.300	8.300
Hospital	1160	0.400	0.490	0.000	0.000	1.000
Salary expenses (¥B)	1150	1.799	0.958	1.172	1.547	2.124
Bonus expenses (¥B)	1150	0.478	0.227	0.333	0.425	0.572
Grant revenues (¥B)	1149	0.031	0.046	0.000	0.016	0.043
Medical revenues (¥B)	1160	5.076	3.491	2.665	4.129	6.231
Education revenues (¥B)	1150	0.050	0.077	0.000	0.001	0.086
R&D revenues (¥B)	1150	0.066	0.099	0.005	0.032	0.076
Other costs (¥B)	1150	5.065	3.364	2.696	4.115	6.203

Multivariate test of H1: absolute value of patient satisfaction information

We test H1 using the following multivariate model, which estimates the level of patient satisfaction in each year subsequent to the introduction of the satisfaction survey:

$$Satisfaction_{i,t} = \alpha + \sum_{t=1}^{T} \gamma_t Year_t + \beta_1 Hospital_i + \beta_2 Size_i + \beta_3 Concentration_i + \varepsilon \quad (2)$$

We control for hospital type (time invariant indicator variable *Hospital* coded as 1 for general hospitals and zero for sanatoriums), hospital *Size* (measured as the number of staffed beds divided by 100) and *Concentration* (which is the number of private and public hospitals per 100,000 people in the geographic area (*prefecture*) where each NHO hospital is located). We perform seven separate estimations, one for the overall satisfaction score for inpatients and outpatients respectively, one for each of the satisfaction factors i.e., *Staff/Treatment* and *Logistics/Infrastructure* for inpatients, and *Staff/Treatment*, *Logistics/Infrastructure*, and *Administrative Procedures* for outpatients. Satisfaction results for 2004 represent the baseline and therefore 2004 is the omitted year dummy. All estimations are performed using OLS with heteroskedasticity-robust standard errors clustered by hospital.[19]

[19] While the distribution of the dependent variable is bounded above (below) by the value of the corresponding factor calculated for a hypothetical hospital that scores 5 (1) on all indicators relative to the factor, the construction of the factor variable is normalized by construction. Therefore, OLS is an appropriate estimator for this model.

Author's personal copy

Table 2 Correlations

		(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1)	Staff and treatment—outpatient	1.0000						
(2)	Administrative procedures—outpatient	0.6177***	1.0000					
(3)	Logistics and infrastructure—outpatient	0.3464***	0.1418***	1.0000				
(4)	Staff and treatment—inpatient	0.2857***	−0.0186	0.3042***	1.0000			
(5)	Logistics and infrastructure—inpatient	0.3434***	0.2257***	0.2664***	0.6411***	1.0000		
(6)	Overall satisfaction—inpatient	0.8478***	0.7360***	0.3032***	0.1806***	0.3042***	1.0000	
(7)	Overall satisfaction—outpatient	0.3300***	0.0523	0.3344***	0.8474***	0.7220***	0.2474***	1.0000
(8)	Size	−0.0142	−0.2705***	0.1902***	0.1562***	−0.0066	−0.0646*	0.1062***
(9)	Concentration	0.0953***	0.0619*	0.1162***	0.1731***	0.1608***	0.1434***	0.2102***
(10)	Hospital	−0.1364***	−0.4431***	0.3402***	0.2546***	0.1115***	−0.1970***	0.2225***
(11)	Salary expenses (¥B)	0.0489	−0.3353***	0.3347***	0.2471***	0.1521***	−0.0254	0.2320***
(12)	Bonus expenses (¥B)	0.0323	−0.3332***	0.3123***	0.2312***	0.1096***	−0.0278	0.2048***
(13)	Grant revenues (¥B)	−0.0256	−0.1230***	0.0744**	0.0168	−0.0114	−0.0324	0.0178
(14)	Medical revenues (¥B)	0.0614*	−0.3315***	0.3325***	0.2646***	0.1692***	−0.0028	0.2575***
(15)	Education revenues (¥B)	−0.0346	−0.2886***	0.3099***	0.1703***	0.1119***	−0.0524	0.1804***
(16)	R&D revenues (¥B)	0.1015***	−0.2091***	0.2144***	0.1564***	0.0366	0.0444	0.1338***
(17)	Other costs (¥B)	0.0553*	−0.3358***	0.3472***	0.2640***	0.1737***	−0.0094	0.2573***

		(8)	(9)	(10)	(11)	(12)	(13)
(8)	Size	1.0000					
(9)	Concentration	−0.0902***	1.0000				
(10)	Hospital	0.1941***	0.0714**	1.0000			
(11)	Salary expenses (¥B)	0.7686***	−0.0494	0.5419***	1.0000		
(12)	Bonus expenses (¥B)	0.8089***	−0.0306	0.4971***	0.9773***	1.0000	
(13)	Grant revenues (¥B)	0.4325***	−0.0463	0.0855***	0.4221***	0.4219***	1.0000
(14)	Medical revenues (¥B)	0.7285***	−0.0250	0.5677***	0.9759***	0.9641***	0.3921***
(15)	Education revenues (¥B)	0.5509***	−0.0280	0.5108***	0.7405***	0.7274***	0.3443***

Author's personal copy

Table 2 (continued)

		(8)	(9)	(10)	(11)	(12)	(13)
(16)	R&D revenues (¥B)	0.6006***	−0.1106***	0.3264***	0.7569***	0.7568***	0.2917***
(17)	Other costs (¥B)	0.7304***	−0.0393	0.5838***	0.9851***	0.9664***	0.3970***

		(14)	(15)	(16)	(17)
(14)	Medical revenues (¥B)	1.0000			
(15)	Education revenues (¥B)	0.7444***	1.0000		
(16)	R&D revenues (¥B)	0.7565***	0.4993***	1.0000	
(17)	Other costs (¥B)	0.9885***	0.7607***	0.7713***	1.0000

The pairwise Pearson correlation coefficients

Statistical significance is indicated as follows: ***$p < 0.01$; **$p < 0.05$; *$p < 0.10$

Author's personal copy

Table 3 Regression results: relationship between factors and overall satisfaction

DV = Overall satisfaction levels	Inpatients			Outpatients		
	(1)	(2)	(3)	(4)	(5)	(6)
	Pooled	General hospitals	Sanatoriums	Pooled	General hospitals	Sanatoriums
Staff and treatment—inpatient	0.439***	0.430***	0.441***	0.616***	0.560***	0.640***
	(13.42)	(12.64)	(11.73)	(12.12)	(7.46)	(12.04)
Logistics and infrastructure—inpatient	0.233***	0.200***	0.248***	0.067*	0.124***	0.052
	(10.76)	(8.56)	(8.04)	(1.80)	(2.96)	(1.19)
Administrative procedures—outpatient				0.350***	0.311***	0.339***
				(8.01)	(7.60)	(6.38)
Hospital	0.036			0.030		
	(1.51)			(0.50)		
Size	0.010	0.026**	−0.011	0.007	0.053**	−0.038
	(0.67)	(2.51)	(−0.43)	(0.29)	(2.03)	(−1.06)
Concentration	0.046***	0.038**	0.057***	0.041**	0.064*	0.035
	(3.54)	(2.12)	(3.28)	(2.04)	(1.83)	(1.34)
Intercept	−0.107***	−0.158***	−0.057	0.038	−0.100	0.079
	(−2.76)	(−3.31)	(−1.13)	(0.54)	(−1.16)	(0.74)
Year FE	YES	YES	YES	YES	YES	YES
Adj. R^2	0.792	0.788	0.784	0.772	0.811	0.753
N	1000	433	567	1148	463	685

Estimation results of the following panel OLS regression, including year fixed effects:

$$OverallSatisfaction_{i,t} = \alpha + \sum_{j=1}^{k} \beta_j Factor_{j_{i,t}} + \delta_1 Hospital_i + \delta_2 Size_i + \delta_3 Concentration_i + \sum_{t=1}^{T} \gamma_t Year_t + \varepsilon$$

Columns (1) and (4) report the estimation of the above regression using the entire population of hospitals for which factors can be calculated using our dataset. Columns (2) and (5) refer to the estimation of the above regression on a sample containing only the general hospitals, while columns (3) and (6) relate to the sample including only sanatoriums. In all estimations, standard errors (reported in parentheses underneath the corresponding coefficients) are clustered by hospital to account for correlation of satisfaction performance across time for the same hospital. Statistical significance is indicated as follows: ***$p<0.01$; **$p<0.05$; *$p<0.10$. An alternative specification of the model including hospital fixed effects instead of hospital characteristics as controls yields equivalent results

Author's personal copy

Table 4 Univariate test of H1: patient satisfaction levels in each year compared to the baseline

	Year	Inpatients			Outpatients		
		N	Mean difference	*t*	N	Mean difference	*t*
Overall satisfaction	2005	228	0.330***	2.93	289	0.401***	2.98
	2006	229	0.461***	3.71	290	0.468***	3.53
	2007	230	0.559***	4.66	290	0.541***	4.18
	2008	229	0.685***	6.23	289	0.631***	4.93
	2009	227	0.581***	4.18	287	0.593***	4.74
	2010	228	0.516***	2.83	287	0.649***	5.16
	2011	227	0.645***	5.14	288	0.665***	5.12
Staff and treatment	2005	271	0.229	1.62	287	0.612***	5.53
	2006	265	0.478***	3.54	288	1.087***	10.67
	2007	264	0.604***	4.68	288	1.125***	10.42
	2008	266	0.689***	5.10	287	1.261***	12.51
	2009	261	0.681***	4.93	285	1.272***	11.65
	2010	261	0.665***	4.72	285	1.316***	12.98
	2011	260	0.751***	6.14	286	1.353***	12.40
Logistics and infrastructure	2005	271	0.347***	2.83	287	– 1.547***	– 16.07
	2006	265	0.492***	4.06	288	– 1.425***	– 14.50
	2007	264	0.488***	4.01	288	– 1.494***	– 14.29
	2008	266	0.678***	5.40	287	– 1.418***	– 14.13
	2009	261	0.849***	6.85	285	– 1.405***	– 14.32
	2010	261	0.849***	6.39	285	– 1.300***	– 13.27
	2011	260	0.957***	8.33	286	– 1.436***	– 14.13
Administrative procedures	2005				287	0.014	0.10
	2006				288	0.050	0.37
	2007				288	0.039	0.30
	2008				287	0.177	1.41
	2009				285	0.101	0.82
	2010				285	0.044	0.35
	2011				286	0.071	0.57

Results of univariate tests, in which we compare the level of each satisfaction factor in each year subsequent to the intervention to the baseline level (i.e., satisfaction levels in 2004)

Statistical significance is indicated as follows: $^{***}p<0.01$; $^{**}p<0.05$; $^{*}p<0.10$

Table 5 reports the results, which indicate that the introduction of the patient satisfaction has, on average, a positive effect on hospital *Overall Satisfaction* for inpatients and outpatients alike for each of the subsequent years following the first year of the patient satisfaction information release. The results for the two satisfaction factors for inpatients and the *Staff/Treatment* factor for outpatients are similar and indicate that relative to 2004, satisfaction performance was higher in each subsequent year. General hospitals appear to score greater inpatient satisfaction compared to sanatoriums, whereas the relation is reversed for the outpatient population (with the exception of satisfaction with logistics and infrastructure). Greater concentration of hospitals in the prefecture is associated with

Author's personal copy

Table 5 Multivariate test of H1: patient satisfaction levels over time

DV = Satisfaction level	Inpatient			Outpatients			
	Overall satisfaction	Staff and treatment	Logistics and infrastructure	Overall satisfaction	Staff and treatment	Administrative procedures	Logistics and infrastructure
2005	0.259***	0.245***	0.334***	0.408***	0.614***	0.011	−1.537***
	(2.84)	(2.82)	(3.84)	(4.98)	(9.07)	(0.17)	(−14.98)
2006	0.420***	0.425***	0.476***	0.468***	1.081***	0.042	−1.416***
	(5.71)	(5.63)	(5.90)	(4.73)	(15.94)	(0.55)	(−13.39)
2007	0.509***	0.503***	0.433***	0.541***	1.118***	0.031	−1.484***
	(6.61)	(6.25)	(5.16)	(5.63)	(14.60)	(0.41)	(−12.74)
2008	0.633***	0.612***	0.629***	0.631***	1.255***	0.174**	−1.413***
	(8.32)	(7.16)	(7.50)	(6.27)	(18.07)	(2.10)	(−12.74)
2009	0.535***	0.594***	0.806***	0.593***	1.272***	0.104	−1.393***
	(5.85)	(6.82)	(8.64)	(5.89)	(16.00)	(1.31)	(−13.24)
2010	0.468***	0.596***	0.802***	0.657***	1.311***	0.063	−1.296***
	(3.26)	(6.34)	(6.78)	(6.57)	(20.03)	(0.77)	(−11.89)
2011	0.586***	0.654***	0.903***	0.670***	1.354***	0.080	−1.427***
	(5.57)	(8.12)	(9.53)	(5.73)	(15.26)	(0.86)	(−12.23)
Hospital	0.331***	0.539***	0.213*	−0.384***	−0.244**	−0.833***	0.413***
	(3.45)	(4.38)	(1.76)	(−3.54)	(−2.39)	(−7.74)	(4.13)
Size	0.113*	0.224**	0.027	−0.046	0.024	−0.211***	0.095
	(1.85)	(2.35)	(0.37)	(−0.73)	(0.40)	(−3.98)	(1.49)
Concentration	0.139***	0.166**	0.155***	0.126**	0.099*	0.058	0.084
	(3.02)	(2.53)	(2.62)	(2.23)	(1.84)	(1.18)	(1.55)
Intercept	−0.593***	−0.808***	−0.673***	−0.339***	−0.899***	0.279***	1.079***
	(−5.50)	(−5.65)	(−5.80)	(−2.93)	(−9.77)	(2.84)	(12.44)
Adj. R^2	0.099	0.149	0.118	0.095	0.211	0.253	0.297
N	1076	1050	1050	1150	1148	1148	1148

Author's personal copy

Table 5 (continued)

Estimation results of the following OLS regression model:

$$Satisfaction_{i,t} = \alpha + \sum_{t=1}^{T} \gamma_t Year_t + \beta_1 Hospital_i + \beta_2 Size_i + \beta_3 Concentration_i + \varepsilon$$

Separate estimations are performed for each measure of patient satisfaction reported in the header of each column. The estimation is performed using panel OLS with heteroskedasticity robust standard errors clustered by hospital to account for behavior correlation through time. In all cases, the reference year is the baseline (2004). Statistical significance is indicated as follows: ***$p<0.01$; **$p<0.05$; *$p<0.10$. An alternative specification of the model including hospital fixed effects instead of hospital characteristics as controls yields equivalent results

Author's personal copy

higher satisfaction, likely the effect of a combination of greater competitive pressures and opportunities for learning and knowledge exchange.

Taken together, our univariate and multivariate results are consistent with the *absolute performance effect* predicted by *H1*—that is, patient satisfaction information improves performance, even after controlling for time invariant characteristics and unobservable hospital level factors. In other words, new patient satisfaction information has decision facilitating value for the decision makers.

Test of H2: benchmarking effect of patient satisfaction information

H2 predicts that a relatively lower initial performance on patient satisfaction is associated with higher magnitude of subsequent nonfinancial performance improvements, arising from the value of the *relative* performance signal. To test H2, we estimate the following model:

$$SatisfactionChange_{i,t} = \alpha + \beta_1 PoorInitialPerformer_i + \beta_2 Size_i + \beta_2 Concentration_i + \beta_3 Hospital_i + \sum_{t=1}^{T} \gamma_t Year_t + \varepsilon \quad (3)$$

where *SatisfactionChange* is defined as the annual change in patient satisfaction ($Year_t - Year_{t-1}$). To account for the bounded nature of patient satisfaction and the influence that variation in the initial level of satisfaction might have on the opportunities for improvement of different hospitals, we scale the change in satisfaction by a measure of the "available change", as follows:

$$SatisfactionChange_{i,t} = \frac{Satisfaction_{i,t} - Satisfaction_{i,(t-1)}}{Max\_Satisfaction - Satisfaction_{i,2004}} \quad (4)$$

Therefore, our dependent variable captures the portion of the initial available improvement that is realized by the hospital in each year subsequent to the intervention.[20] *PoorInitialPerformer* is an indicator variable that takes value 1 if the hospital was in the lowest quartile of performance in the baseline year, and zero otherwise. The control variables *Competition*, *Size*, and *Hospital* were defined previously. Estimations are performed using OLS with heteroskedasticity-robust standard errors clustered by hospital and include year fixed effects.

Results presented in Table 6 indicate a positive coefficient for *PoorInitialPerformer* for overall satisfaction of outpatients and inpatients alike and for all satisfaction factors, with the exception of satisfaction with logistics and infrastructure. These results indicate that hospitals that had lower baseline satisfaction scores (2004) realized greater portions of their opportunities for satisfaction improvement (i.e., performance improvement scaled by opportunity to improve) thereafter, compared to hospitals that were already performing at a higher level. These results collectively support our prediction relative to the referent performance value of information (*H2*).

[20] To calculate the values corresponding to the maximum performance for each satisfaction factor, we computed each factor score for a hypothetical hospital scoring 5 on each question in the inpatient and outpatient questionnaires.

Author's personal copy

Table 6 Test of H2: comparison of changes in patient satisfaction between poor initial performers and the rest of the population

DV = Annual performance change scaled by initial opportunity for improvement	Inpatients			Outpatients			
	Overall satisfaction	Staff and treatment	Logistics and infrastructure	Overall satisfaction	Staff and treatment	Administrative procedures	Logistics and infrastructure
Poor initial performer	0.068*	0.037***	−1.965	0.022***	0.014**	0.030***	0.059
	(1.72)	(2.59)	(−0.96)	(3.74)	(2.24)	(6.46)	(0.52)
Size	0.027	0.004	−0.603	0.009*	0.007	0.006	−0.037
	(1.29)	(1.18)	(−0.97)	(1.91)	(1.59)	(1.54)	(−0.56)
Concentration	0.004	0.006*	−0.678	0.005	0.002	0.002	0.016
	(0.35)	(1.95)	(−0.98)	(1.45)	(0.70)	(0.55)	(0.68)
Hospital	0.047	0.013	−1.872	0.013***	0.005	0.006	0.015
	(1.29)	(1.39)	(−1.00)	(2.68)	(1.18)	(1.08)	(0.10)
Intercept	0.004	0.019	2.594	0.051***	0.132***	−0.030*	−0.926*
	(0.07)	(0.61)	(1.02)	(3.53)	(8.03)	(−1.90)	(−1.86)
Year FE	YES	YES	YES	YES	YES	YES	YES
Adj. R^2	0.022	0.011	0.002	0.024	0.065	0.011	0.015
N	587	856	856	1002	988	988	988

Estimation results of the following OLS regression model:

$$SatisfactionChange_{i,t} = \alpha + \beta_1 PoorInitialPerformer_i + \beta_2 Size_i + \beta_2 Concentration_i + \beta_3 Hospital_i + \sum_{t=1}^{T} \gamma_t Year_t + \varepsilon$$

To account for the differences in the opportunities to improve depending on the performance starting point, for each hospital we scaled the change in satisfaction in each year by the original opportunity for improvement calculated as the maximum performance level and the baseline, as follows:

$$Satisfaction\_Change_{i,t} = \frac{Satisfaction_{i,t} - Satisfaction_{i,(t-1)}}{Max\_Satisfaction - Satisfaction_{i,2004}}$$

Separate estimations are performed for each measure of patient satisfaction reported in the header of each column. The estimation is performed using panel OLS with heteroskedasticity robust standard errors clustered by hospital to account for behavior correlation through time. In all cases, the reference year is the baseline (2004). Statistical significance is indicated as follows: ***$p<0.01$; **$p<0.05$; *$p<0.10$

Author's personal copy

Test of H3: value of first information versus subsequent information

Next, we examine if information has a higher value when it is new, compared to subsequent updates of the same information. First, we focus on the absolute value of information and estimate the following model:

$$SatisfactionChange_{i,t} = \alpha + \sum_{t=1}^{T} \gamma_t * Year_t + \beta_1 Size_i + \beta_2 Concentration_i + \beta_3 Hospital_i + \varepsilon \tag{5}$$

where the dependent variable is, as before, calculated as the ratio between the change in satisfaction between year $(t-1)$ and year (t) scaled by the initial opportunities for improvement (see Eq. (4) above). All other variables are defined as previously described. The base (omitted) case in all estimations is the first change—i.e., the change in satisfaction performance between 2004 and 2005. In order to fully support H3, we would need to estimate negative coefficients associated with each of the years subsequent to 2005. We estimate Eq. (5) separately for each satisfaction factor using OLS with heteroskedasticity robust standard errors clustered by hospital.

Our estimation results are reported in Table 7, Panel A, and show partial support for H3. In particular, we find that the prediction formalized with H3—that is, the improvement in absolute satisfaction is greater in the first year compared to subsequent adjustments—is supported when the dependent variable refers to overall outpatient satisfaction. We also find consistent results, for the most part, for changes in outpatient satisfaction with staff and treatment. With the exception of changes in outpatient satisfaction with respect to logistics and infrastructure, our results weakly support H3, namely, performance changes in subsequent years are smaller or not statistically different than the change in the first year. We note that all our tests described so far are robust to the inclusion of hospital fixed effects.

Considering that initial poor performers had greater opportunities to improve, we explore whether subsequent information about performance might have different value for this group. For example, hospitals that performed really poorly before the intervention might face greater difficulties to identify solutions and implement changes in the culture of their organizations. Therefore, it is possible that performance improvements might be smaller in the initial periods and increase subsequently as organizations learn from the feedback received about what solutions are more effective to improve patient satisfaction. As Dr. Nakai recalled, “there were many [members of the] staff who struggled to get used to the new way of treating patients. We told them to take their time and adjust at the speed at which they were comfortable.” On the other hand, the benchmarking information effect for initial poor performers might prompt decision makers in these organizations to devote considerable effort and act quickly to gain the favor of their patients, which could lead to greater improvements in the first year compared to subsequent ones. We estimate the following model:

Author's personal copy

Table 7 Test of H3: performance changes year over year

DV = Performance change scaled by initial potential for improvement	Inpatients			Outpatients			
	Overall satisfaction	Staff and treatment	Logistics and infrastructure	Overall satisfaction	Staff and treatment	Administrative procedures	Logistics and infrastructure
Panel A: Pooled sample							
2006–2005	0.043	0.050	0.009	−0.068**	−0.036	0.013	1.549***
	(0.58)	(1.06)	(0.23)	(−2.34)	(−1.03)	(0.37)	(3.82)
2007–2006	−0.039	−0.010	−0.053*	−0.051***	−0.134***	0.012	1.316***
	(−0.60)	(−0.26)	(−1.86)	(−2.70)	(−5.79)	(0.62)	(6.26)
2008–2007	0.035	0.007	0.002	−0.053***	−0.109***	0.035	0.295
	(0.40)	(0.19)	(0.07)	(−2.77)	(−5.20)	(1.56)	(0.35)
2009–2008	−0.080	−0.046	−0.026	−0.075***	−0.131***	0.006	1.115***
	(−1.50)	(−1.02)	(−0.81)	(−4.03)	(−5.85)	(0.26)	(3.43)
2010–2009	−0.195	−0.040	−0.069**	−0.050***	−0.126***	0.010	0.949**
	(−1.24)	(−1.09)	(−2.04)	(−2.62)	(−5.43)	(0.53)	(2.22)
2011–2010	−0.053	−0.016	−0.015	−0.054***	−0.128***	0.027	0.662
	(−0.70)	(−0.30)	(−0.41)	(−3.12)	(−5.98)	(1.30)	(1.20)
Size	0.025	0.002	−0.524	0.009*	0.007	0.007*	−0.039
	(1.24)	(0.43)	(−0.97)	(1.89)	(1.50)	(1.66)	(−0.63)
Concentration	−0.003	0.004	−0.571	0.003	0.001	0.000	0.012
	(−0.22)	(1.11)	(−0.98)	(1.07)	(0.44)	(0.01)	(0.51)
Hospital	0.034	0.006	−1.790	0.016***	0.006	0.016***	0.032
	(1.07)	(0.70)	(−1.00)	(3.16)	(1.40)	(2.77)	(0.18)
Intercept	0.028	0.031	2.068	0.055***	0.135***	−0.027*	−0.918*
	(0.44)	(1.02)	(1.03)	(3.82)	(8.21)	(−1.68)	(−1.89)
Adj. R^2	0.020	0.009	0.002	0.022	0.064	0.007	0.015
N	587	856	856	1002	988	988	988

Author's personal copy

Table 7 (continued)

DV = Performance change scaled by initial potential for improvement	Inpatient			Outpatients			
	Overall satisfaction	Staff and treatment	Logistics and infrastructure	Overall satisfaction	Staff and treatment	Administrative procedures	Logistics and infrastructure
Panel B: Comparison between poor performers and the rest of the population							
Poor initial performer * 06-05	−0.322**	−0.207*	−0.141*	−0.068^	−0.088^	−0.107**	−1.708***
	(−2.11)	(−1.93)	(−1.63)	(−1.39)	(−1.50)	(−1.99)	(−3.14)
Poor initial performer * 07-06	−0.301***	−0.347***	−0.164**	−0.080**	−0.085**	−0.111***	−1.468***
	(−2.71)	(−3.71)	(−2.26)	(−2.30)	(−2.10)	(−3.20)	(−5.25)
Poor initial performer * 08-07	−0.210^	−0.296***	−0.115^	−0.047	−0.071^	−0.077**	−0.036
	(−1.52)	(−3.69)	(−1.59)	(−1.16)	(−1.55)	(−2.21)	(−0.03)
Poor initial performer * 09-08	−0.309***	−0.271***	−0.080	−0.089***	−0.127***	−0.114***	−1.179***
	(−2.76)	(−3.08)	(−0.86)	(−2.67)	(−3.25)	(−2.98)	(−2.70)
Poor initial performer * 10-09	−0.029	−0.132^	−0.043	−0.087**	−0.077*	−0.127***	−0.820
	(−0.12)	(−1.51)	(−0.53)	(−2.39)	(−1.67)	(−3.57)	(−1.42)
Poor initial performer * 11-10	−0.262^	−0.304***	−0.194**	−0.088***	−0.056	−0.101***	−0.603
	(−1.58)	(−3.32)	(−2.41)	(−2.63)	(−1.30)	(−2.77)	(−0.81)
Poor initial performer	0.273***	0.254***	−1.825	0.087***	0.086***	0.121***	0.889*
	(2.61)	(4.19)	(−0.89)	(3.15)	(2.89)	(4.46)	(1.70)
Size	0.027	0.004	−0.592	0.009*	0.007	0.006	−0.037
	(1.28)	(1.27)	(−0.97)	(1.90)	(1.58)	(1.53)	(−0.56)
Concentration	0.005	0.006**	−0.665	0.005	0.002	0.002	0.016
	(0.37)	(2.07)	(−0.98)	(1.45)	(0.69)	(0.54)	(0.69)
Hospital	0.047	0.014	−1.832	0.013***	0.005	0.006	0.014
	(1.28)	(1.61)	(−0.99)	(2.66)	(1.15)	(1.06)	(0.10)

Author's personal copy

Table 7 (continued)

DV = Performance change scaled by initial potential for improvement	Inpatient			Outpatients			
	Overall satisfaction	Staff and treatment	Logistics and infrastructure	Overall satisfaction	Staff and treatment	Administrative procedures	Logistics and infrastructure
Intercept	−0.046	−0.034	2.541	0.034**	0.114***	−0.053***	−1.136*
	(−0.56)	(−0.99)	(1.00)	(2.05)	(5.60)	(−2.74)	(−1.88)
Year FE	YES	YES	YES	YES	YES	YES	YES
Adj. R^2	0.030	0.033	0.002	0.030	0.071	0.019	0.019
N	587	856	856	1002	988	988	988

Results of our tests of H3. All estimations are performed via separate panel OLS regressions with heteroskedasticity-robust standard errors clustered by hospital to account for correlations across time. To account for the differences in the opportunities to improve depending on the performance starting point, for each hospital we scaled the change in satisfaction in each year by the original opportunity for improvement calculated as the maximum performance level and the baseline, as follows:

$$Satisfaction\_Change_{i,t} = \frac{Satisfaction_{i,t} - Satisfaction_{i,(t-1)}}{Max\_Satisfaction - Satisfaction_{i,2004}}$$

In all cases, statistical significance is indicated as follows: ***$p<0.01$; **$p<0.05$; *$p<0.10$. Standard errors are reported in parentheses underneath the corresponding coefficient

Panel A, reports the estimation results of the following regression model:

$$SatisfactionChange_{i,t} = \alpha + \sum_{t=1}^{T} \gamma_t * Year_t + \beta_1 Size_i + \beta_2 Concentration_i + \beta_3 Hospital_i + \varepsilon$$

An alternative specification of the model including hospital fixed effects instead of hospital characteristics as controls yields equivalent results

Panel B reports the estimation results of the following regression model, in which we compare annual changes in patient satisfaction between hospitals that ranked in the bottom quartile in 2004 (i.e. indicator variable *PoorInitialPerformer* = 1) and the rest of the population:

$$SatisfactionChange_{i,t} = \alpha + \sum_{t=1}^{T} \gamma_t (PoorInitialPerformer_i * Year_t) + \beta_1 PoorInitialPerformer_i + \beta_2 Size_i + \beta_3 Concentration_i + \beta_4 Hospital_i + \sum_{t=1}^{T} \delta_t Year_t + \varepsilon$$

^The significance of the coefficient is barely outside of canonical ranges (two-tail $p<0.13$), thus indicating one-tail significance. An alternative specification of the model including hospital fixed effects instead of hospital characteristics as controls yields equivalent results

Author's personal copy

$$SatisfactionChange_{i,t} = \alpha + \sum_{t=1}^{T} \gamma_t (PoorInitialPerformer_i * Year_t) + \beta_1 PoorInitialPerformer_i + \beta_2 Size_i + \beta_3 Concentration_i + \beta_4 Hospital_i + \sum_{t=1}^{T} \delta_t Year_t + \varepsilon \quad (6)$$

where all the variables are defined as previously described. The coefficients associated with the interaction terms between *PoorInitialPerformer* and each year inform us on the trend in performance improvement subsequent to the first change (which is the omitted base case). As before, we estimate the equation using OLS with heteroskedasticity robust standard errors clustered by hospital and we include year fixed effects. Results are reported in Table 7, Panel B, and indicate that, while the main effect of *PoorInitialPerformer* remains consistent with our tests of H2, the diminishing returns of patient satisfaction performance information are more pronounced for initial poor performers than better performing hospitals. In fact, for the majority of the years after the first change, hospitals initially performing poorly exhibit lower relative improvements than in the first year (recall that our measure of performance changes is scaled by the available opportunities for improvement).[21] Taken together, our results support H3, especially for hospitals for which the initial benchmarking effect was the strongest.

Additional analyses

Relation between initial changes and improvement persistence

Our analyses provide evidence that the intervention of the NHO did indeed generate effects consistent with the intended purposes. On average, hospitals in the NHO system improved their patient satisfaction performance after the introduction of the survey, and performance did not regress. Improvements were greater for poorly performing hospitals and in the early stages of the intervention, where the Bayesian revisions were likely bigger. We next explore whether the changes enacted in the early stages of the intervention significantly predict the persistence of the performance improvements at the hospital level. On the one hand, hospitals might have responded to the initial signal with cosmetic changes that might gain improvements in the short term but, without affecting the organizational culture in a stable manner, were less likely to yield to persistent results (Cavalluzzo and Ittner 2004). On the other hand, it is possible that changes enacted by individual hospitals were structural in nature and, therefore, led to persistent subsequent performance. We estimate the following model:

$$ImprovementPersistenceYear\_t_i = \alpha + \beta_1 FirstYearChange_i + \beta_2 Size_i + \beta_3 Concentration_i + \beta_4 Hospital_i + \varepsilon \quad (7)$$

The hospital-level dependent variable *ImprovementPersistenceYear_t* is defined as the difference between the performance in year t and the baseline performance (i.e., performance in year 2004). Our tests of H2 showed that changes subsequent to the first year

[21] We would not be able to perform the test of H2 reported in Table 6 with hospital FE, since *InitialPoorPerformer* is a time-invariant characteristic that would be absorbed by the fixed effects. However, untabulated tests on two subsamples (respectively, poor initial performers and the rest of the population) yield consistent results.

Author's personal copy

Table 8 Additional analysis: persistence of initial performance change

DV = Persistence in year t	Inpatients						Outpatients					
	2006	2007	2008	2009	2010	2011	2006	2007	2008	2009	2010	2011
Overall satisfaction	0.346***	0.385***	0.432**	0.336***	0.720**	0.657**	0.781***	0.716***	0.722***	0.737***	0.662***	0.777***
	(3.24)	(3.27)	(2.62)	(2.99)	(2.47)	(2.48)	(5.15)	(3.36)	(4.14)	(4.05)	(3.58)	(3.71)
Staff and treatment	0.551***	0.630***	0.525***	0.530**	0.678***	0.718***	0.272*	0.252	0.317*	0.367*	0.388***	0.306
	(4.06)	(3.74)	(2.78)	(2.56)	(4.01)	(4.62)	(1.68)	(1.10)	(1.92)	(1.71)	(2.93)	(1.38)
Logistics and infrastructure	0.683***	0.540***	0.459***	0.566***	0.797***	0.468*	0.887***	0.926***	0.929***	0.806***	0.867***	0.413***
	(10.26)	(3.74)	(3.14)	(4.72)	(4.63)	(1.74)	(22.19)	(19.81)	(20.15)	(18.60)	(21.74)	(4.39)
Administrative procedures							0.566***	0.585***	0.753***	0.627***	0.588***	0.301**
							(3.87)	(4.77)	(7.58)	(6.30)	(5.56)	(2.01)

Estimations of the β_1 coefficients for the following regression model:

$$ImprovementPersistenceYear\_t_i = \alpha + \beta_1 FirstYearChange_i + Size_i + Concentration_i + Hospital_i + \varepsilon$$

We perform the estimations of the above equation using OLS with heteroskedasticity robust standard errors. Separate cross-sectional estimations are performed for each factor/year combination (e.g. we regress the persistence in overall satisfaction performance improvements observable in 2006 on the initial change in overall satisfaction between 2005 and 2004). In the interest of parsimonious exposition, we report only the coefficients associated with the first change (*FirstYearChange*). Standard errors are reported in parentheses underneath the corresponding coefficient. Statistical significance is indicated as follows: ***$p < 0.01$; **$p < 0.05$; *$p < 0.10$

Author's personal copy

were, on average, smaller or non-distinguishable from the first change. To the extent that the performance level in subsequent years remains higher than the baseline, we conjecture that the effect of the intervention is persistent. The variable of interest is *FirstYearChange*, defined for each hospital as the difference between satisfaction performance in year 2005 and the baseline performance. All other variables are defined as previously described.

Table 8 reports the results of our estimations. We estimate Eq. (7) separately for each satisfaction factor and for each year, using OLS with heteroskedasticity robust standard errors clustered by hospitals. That is, we perform 42 separate regressions. For clarity of exposition and ease of interpretation, we report only the coefficients associated with the predictor of interest. That is, each cell of Table 8 reports the coefficient of *FirstYearChange* for a *different* regression. The results indicate that the changes implemented in the first year of the program reliably predict the persistence of the results. Therefore, we conclude that the changes enacted by hospitals in response to the availability of new information about patient satisfaction performance were structural in nature. The systematic relation between the first change in performance and the persistence of improvements also reduces the concern that our documented effects are due to mean reversion.

Alternative explanations: compensation practices and government grants

Our setting allows us to examine the value of information of patient satisfaction performance feedback in the absence of incentive compensation tied to performance improvement. To validate that this is really the case, we examine physician compensation practices at NHO hospitals using field and archival data to determine if there is any link between patient satisfaction performance and compensation. Further, the NHO provides member hospitals with grants to support a broad range of initiatives, including renovations of older buildings, training programs, clinical trials, medical research, etc.[22] While improvements in patient satisfaction are not clearly stated as criteria for the allocation of grant resources, we explore whether a relation might exist, which would point to potential discretion among the NHO leadership with respect to grant assignments to promote the success of the patient satisfaction initiative.

Field evidence of compensation practices at the NHO

We conducted interviews with hospital administrators at the NHO headquarters to glean information about physician compensation practices. These interviews revealed that there is no explicit link between physician or administrative officer compensation and patient satisfaction performance. Essentially, NHO clinicians and administrators are government employees. Each hospital worker is classified into a particular grade based on a pre-defined hierarchy and paid based on a government salary schedule. The typical compensation package includes: a monthly salary, allowances for cost of living, overtime and travel. There is no performance-related bonus. "Appendix 4" contains information on employment, compensation, and promotion systems at NHO at the time of the study. For example, a senior trauma surgeon would earn a monthly salary of 572,900 Yen (Employment Grade 10, Step 21). Adding a 20% allowance for cost of living and housing, the total compensation package is 687,480 Yen (about $6027 per month in current U.S. Dollars). Dr. Nakai confirmed

[22] Source: Guidebook of the National Hospital Organization—www.nho.hosp.go.jp.

Author's personal copy

Table 9 Additional analysis—patient satisfaction performance, compensation, and grant revenues

DV = Satisfaction level	Inpatient			Outpatients			
	Overall satisfaction	Staff and treatment	Logistics and infrastructure	Overall satisfaction	Staff and treatment	Administrative procedures	Logistics and infrastructure
Salary expenses (¥B)	0.010	0.014	−0.266	−0.197	−0.146	0.099	0.311
	(0.05)	(0.08)	(−0.87)	(−0.84)	(−0.78)	(0.43)	(1.39)
Allowance expenses (¥B)	−1.516***	−0.217	−0.971	−0.044	−0.029	−0.829	−0.387
	(−2.81)	(−0.46)	(−1.36)	(−0.06)	(−0.06)	(−1.44)	(−0.64)
Grant revenues (¥B)	−2.007***	−0.012	−1.187	0.209	−0.242	0.061	−0.559
	(−2.58)	(−0.02)	(−1.33)	(0.35)	(−0.55)	(0.14)	(−0.99)
Medical revenues (¥B)	0.016	−0.007	−0.047	−0.046*	−0.057***	−0.068**	−0.062**
	(0.67)	(−0.35)	(−1.19)	(−1.90)	(−3.32)	(−2.53)	(−2.36)
Education revenues (¥B)	−0.590	−0.170	−0.364	−0.125	−0.526	0.143	0.755
	(−1.24)	(−0.47)	(−0.46)	(−0.20)	(−1.00)	(0.20)	(1.50)
R&D revenues (¥B)	−0.837**	−0.295	−2.336***	0.240	0.563	−0.100	0.730
	(−2.08)	(−0.75)	(−3.75)	(0.40)	(1.44)	(−0.15)	(1.26)
Other costs (¥B)	0.196***	0.042	0.421***	0.266***	0.212***	0.181***	0.070
	(3.19)	(0.74)	(4.26)	(3.95)	(4.10)	(2.64)	(1.11)
Size	0.016	0.196	−0.268***	−0.385***	−0.209**	−0.374***	−0.112
	(0.17)	(1.56)	(−2.71)	(−3.62)	(−2.47)	(−4.61)	(−1.27)
Concentration	0.144***	0.167**	0.157***	0.141***	0.112**	0.072	0.099*
	(3.65)	(2.54)	(2.96)	(2.76)	(2.28)	(1.47)	(1.91)
Hospital	−0.015	0.476***	−0.445***	−0.888***	−0.584***	−1.122***	0.137
	(−0.13)	(3.42)	(−2.91)	(−5.36)	(−4.88)	(−7.24)	(1.25)
Intercept	−0.652***	−0.817***	−1.073***	−0.832***	−1.226***	0.090	0.763***
	(−3.98)	(−4.57)	(−6.05)	(−3.05)	(−7.34)	(0.49)	(4.70)
Year FE	YES	YES	YES	YES	YES	YES	YES
Adj. R^2	0.172	0.152	0.197	0.136	0.252	0.256	0.325

Author's personal copy

Table 9 (continued)

DV = Satisfaction level	Inpatient			Outpatients			
	Overall satisfaction	Staff and treatment	Logistics and infrastructure	Overall satisfaction	Staff and treatment	Administrative procedures	Logistics and infrastructure
N	1070	1043	1043	1143	1141	1141	1141

Estimation results of the following equation:

$$Satisfaction_{i,t} = \alpha + \beta_1 SalaryExpenses_{i,t} + \beta_2 AllowanceExpenses_{i,t} + \beta_3 GrantRevenues_{i,t} + \beta_4 MedicalRevenues_{i,t} + \beta_5 EducationRevenues_{i,t} + \beta_6 R\&DRevenues_{i,t} + \beta_7 OtherCosts_{i,t} + \beta_8 Size_i + \beta_9 Concentration_i + \beta_{10} Hospital_i + \sum_{t=1}^{T} \gamma_t Year_t + \varepsilon$$

Separate estimations are performed for each of the dependent variables indicated in the header of each column. All estimations are performed via panel OLS regressions, with heteroskedasticity-robust standard errors, clustered by hospital. Standard errors are reported in parentheses underneath the corresponding coefficient. All estimations include year fixed effects. Statistical significance is indicated as follows: ***$p<0.01$; **$p<0.05$; *$p<0.10$. An alternative specification of the model including hospital fixed effects instead of hospital characteristics as controls yields equivalent results

Author's personal copy

during our interview that, even when the patient satisfaction intervention was launched, there was no plan to attach any monetary reward to performance improvements.

Archival evidence of compensation practices at the NHO

To empirically examine whether there is any link between patient satisfaction and physician compensation, we estimate the following model:

$$\begin{aligned} Satisfaction_{i,t} = {} & \alpha + \beta_1 SalaryExpenses_{i,t} + \beta_2 AllowanceExpenses_{i,t} + \beta_3 GrantRevenues_{i,t} \\ & + \beta_4 MedicalRevenues_{i,t} + \beta_5 EducationRevenues_{i,t} + \beta_6 R\,DRevenues_{i,t} \\ & + \beta_7 OtherCosts_{i,t} + \beta_8 Size_i + \beta_9 Concentration_i + \beta_{10} Hospital_i \\ & + \sum_{t=1}^{T} \gamma_t Year_t + \varepsilon \end{aligned} \tag{8}$$

where *SalaryExpenses* is total compensation received by hospital workers including benefits. *AllowanceExpenses* is the total annual incremental payout by the hospital to clinicians which is predominantly based on increases in the cost of living, *GrantRevenues* is the total amount of NHO grants allocated to the individual hospital in each year. We include control variables for size, competition, hospital type, as well as Medical Revenue, Educational Revenue, R&D Revenue, and Other Costs. Descriptive statistics for the variables used in these analyses are included in Table 1.

If patient satisfaction were taken into consideration in salary or allowance payouts, the coefficient on β_1 and β_2 would be positive. If patient satisfaction was associated with the amount of grants made available to the hospital, we would estimate a positive value for β_3. Estimation results reported in Table 9 indicate generally no significant association between salary or allowances and either inpatient or outpatient satisfaction, with the exception of inpatient overall satisfaction, which exhibits a negative relation with allowances and grants. These results suggest that patient satisfaction does not determine of clinician or hospital payoffs. Thus, we conclude that improvements in patient satisfaction are not driven by the motive to increase monetary payoffs.

A plausible concern could be raised with respect to the effect of variation in the compensation of top physician executives, to the extent that their compensation might be dependent, in part, on the overall performance of the hospital. It is feasible that patient satisfaction could increase more prominently in those hospitals where the compensation of top executives is most elastic to performance factors. Unfortunately, the NHO does not disclose any detail about executive compensation or any breakdown of compensation data by hierarchical level within the hospital. Consequently, we cannot explore this source of variation.[23] We encourage future research to explore the moderating effect of top management pay sensitivity to hospital performance on the relation between patient satisfaction information and subsequent improvements.

[23] To address the potential impact of variation in levels of pay on patient satisfaction improvements, we performed an additional analysis, in which we restricted the estimation of Eq. (8) to a subsample of hospital/year observations, constructed by identifying the hospitals in the highest quartile of average salary per staffed bed in each year and each type of hospital. The results (untabulated) of our estimation are consistent with those reported in Table 9.

Author's personal copy

Alternative explanations: regression to the mean

We consider the extent to which results may be a manifestation of regression towards the mean rather than actual improvements in patient satisfaction. That is, poorly performing hospitals may improve in performance simply because of the nature of the behavior of extreme values in a statistical distribution rather than an actual improvement.[24] We are able to reduce the concerns with respect to the alternative explanation of regression to the mean based on the following. First, Table 4 shows that the overall mean inpatient and outpatient satisfaction measure increases over time. This rules out aggregate mean stability, a necessary condition for regression to the mean (Cook and Campbell 1979; Zhang and Tomblin 2003). Second, the overall satisfaction never decreases significantly below the initial (2004) levels in aggregate or for the hospitals in the highest quartile of performance (untabulated results), which is further evidence against the aggregate mean stability condition. Third, we estimate the correlation between subsequent satisfaction measures after controlling for hospital fixed effects. Untabulated results show non-significant correlations coefficients, which is inconsistent with the conjecture that the improvement over time may be purely a result of regression towards the mean (Cook and Campbell 1979; Zhang and Tomblin 2003). Finally, as described above, results documented in Table 8 provide evidence of a systematic relation between hospital response to the new patient satisfaction information and the persistence of performance improvements in subsequent years. Taken together, these findings lead us to conclude that regression towards the mean does not fully explain the results of this study.

Conclusions

Performance information generally has decision value, regardless of whether it is private or public. However, most studies analyzing the value of information are confounded with decision makers' responses to pecuniary incentives stemming from performance pressures or public disclosure pressures. In this study we examine the effect of mandatory patient satisfaction performance measurement on subsequent satisfaction performance in a setting where there is minimal confound from public disclosures or incentive compensation. However, from a practical perspective, as Narayanan and Davila (1998, page 272) state: "Most firms collect a plethora of information for belief revision, even though only a few signals are directly linked to incentives." Further, they note that when a signal that is useful for belief-revision is also used for performance evaluation, the manager has incentives to manipulate the signal. Managerial manipulation lowers the value of the performance signal in its belief revision and learning role.

Our empirical setting allows an examination of the value of information with minimal confound from public disclosures or incentive compensation. We use patient satisfaction data for all of the 145 hospitals members of the Japanese NHO for a period of 8 years (2004–2011) to explore the value of information. In our setting, hospitals obtained two

[24] Note that regression to the mean is primarily an issue when the analysis consists of only two observations, such as two variables measured on one occasion (e.g. control and treatment group in an experiment) or one variable measured on two occasions (e.g. pre-test post-test comparison after an experimental intervention). Regression to the mean is not a phenomenon that is relevant to multiple observations over time (Nesselroade et al. 1980).

Author's personal copy

new signals: first, a signal about the absolute level of patient satisfaction performance, and, second, a signal about performance relative to a referent group. We find evidence of an *absolute performance effect*—the new information signal about the level of performance induces a more precise posterior distribution of beliefs and facilitates decision making. We also find a *benchmarking effect* arising from the value of the new information signal about performance relative to a peer group. Results further indicate that information has higher decision support value when it is new. New information results in the greatest belief revision, compared to subsequent updates. Relatedly, we find that changes implemented as a result of belief revisions when the information is new are predictive of the persistence in performance improvements.

Performance measurement systems are generally developed to facilitate better decisions. Prior research suggests that when performance measurement systems are mandated through legislative provisions, subordinate organizations are likely to comply with the regulatory requirement but make little use of such information for internal decisional processes (Cavalluzzo and Ittner 2004). In our setting, the subordinate organizations (i.e., NHO hospitals) appear to effectively use the new information in a quick fashion. Our results support the notion that information generated by new performance measurement systems can have persistent decision value for firms.

We contribute to both research, practice, and policy by showing that in healthcare settings, new information about patient satisfaction is valuable to improve subsequent performance. Our findings are particularly relevant to healthcare organizations, in light of the fact that compensation contracts in this industry are rarely materially linked to quality measures while, at the same time, policy provisions aiming at improving the quality and value of healthcare services for patients continue to be developed and introduced in many countries. Our results provide evidence that systematic collection of information about absolute and relative performance leads to the internalization of the goals of the policy and subsequent performance improvements, even in absence of compensation incentives and/or pressures from public disclosure.

Acknowledgements We thank for their valuable comments and suggestions Jeff Biddle, Clara Chen, Leslie Eldenburg, Regina Herzlinger, Bob Kaplan, Matthias Mahlendorf, Melissa Martin, Pam Murphy, Steve Salterio, Greg Sabin, Daniel Thornton, Stephanie Tsui, Jeff Wooldridge, workshop participants at the "Patient-Centric Healthcare Management in the Age of Analytics" conference, University of Arizona, Erasmus University, Michigan State University, Queen's School of Business, and Wilfrid Laurier University. We appreciate the help we received in gathering and interpreting information about the Japanese healthcare industry from Nobuo Sato and Mayuka Yamazaki at the Harvard Business School Research Center in Tokyo, and for the precious insights on the institutional settings shared with us by Dr. Kunio Nakai, and by Kanoko Oishi. We thank Kenji Yasukata, Yoshinobu Shima and Chiyuki Kurisu for their support in the collection of the data used in this study, and to Sho Nihei for his translation services. All errors are our own.

Funding This research did not receive any specific grant funding from agencies in the public, commercial, or not-for-profit sector.

Data availability Data used in this study can be obtained from the Japanese National Hospital Organization.

Appendix 1: Survey instrument

Panel A: Overall satisfaction (Same questions asked separately for outpatients and inpatients; Scale 1 = Strongly Dissatisfied; 2 = Somewhat Dissatisfied; 3 = Neutral; 4 = Somewhat Satisfied; 5 = Strongly Satisfied

Author's personal copy

I am generally satisfied with this hospital
I am satisfied with the results of the treatment
I am satisfied with the period of the treatment
I am satisfied with treatment I have been taking
I am satisfied with the hospital
I think this hospital provides safe medical services
I think the explanations provided by the medical staff were very clear
I think the treatment I have received was acceptable
I generally trust this hospital
I would like to recommend this hospital to family members and friends

Panel B: Individualized questions for inpatients (Scale 1 = Strongly Agree; 2 = Somewhat Agree; 3 = Neither Agree nor Disagree; 4 = Somewhat Disagree; 5 = Strongly Disagree)
I am not satisfied with the explanation by doctors when I was hospitalized
I was unhappy with the procedure of medical admission
I was unhappy with hospital's explanation about my life during the hospital stay
I think that the doctors behave badly and use bad language in this hospital
I was worried about some doctors' skills and knowledge
I think that the nurses behave badly and use bad language in this hospital
I was unhappy with the assistance received for daily life activities
I think that medical staff such as doctors, nurses and other medical staff lacked teamwork
I did not like today's medical tests (For patients who accepted medical tests)
I did not like today's medical surgeries (For patients who accepted medical surgeries)
I did not like today's medical treatment (For patients who accepted medical treatment)
I did not like today's drip, injection, medicine, or prescription (For patients who had a drip, injections, medicine, or prescription)
I did not like today's rehabilitation (For patients who had rehabilitation)
I am unhappy with the toilets and bathrooms in this hospital
I think that passageways, stairs and elevators are inconvenient
I am unhappy with my room
I am unhappy with the food in this hospital
I am unhappy with the other environment such as stores, and interiors
I am unhappy with the hospital's explanation of my discharge

Panel C: Individualized questions for outpatients (Scale 1 = Strongly Agree; 2 = Somewhat Agree; 3 = Neither Agree nor Disagree; 4 = Somewhat Disagree; 5 = Strongly Disagree)
I felt uneasy when I came to the hospital at the initial visit
I think that this hospital is very inconvenient
I have a bad impression about this hospital
I am unhappy with waiting time
I am unhappy with the waiting room
I think that doctors behave badly and use bad language in this hospital
I was worried about some doctors' skills and knowledge
I think that nurses behave badly and use bad language in this hospital
I did not like today's medical tests (For patients who accepted medical tests)
I did not like today's medical treatment (For patients who accepted medical treatment)

Author's personal copy

I did not like today's drip, injection, medicine, or prescription (For patients who had a drip, injections, medicine, or prescription)
I did not like today's rehabilitation (For patients who had rehabilitation)
I am unhappy with the treatment room
I am unhappy with the other environment such as shops, ATM, and interiors
I am unhappy with the procedures for payment

This appendix lists the questions used in the survey administered to Japanese National Hospital Organization (NHO) general hospitals and sanatoriums. The translation from Japanese to English aimed at maintaining the original meaning as close as possible

Appendix 2: Descriptive statistics for each question in the survey

Panel A: Inpatients

Inpatients	N	Mean	SD	p25	p50	p75	Min	Max
Doctors explanations when hospitalized	1126	4.332	0.411	4.285	4.438	4.543	1.000	5.000
Admission procedures	1128	4.301	0.392	4.248	4.393	4.500	1.000	5.000
Explanation about life during hospital stay	1128	4.137	0.373	4.014	4.196	4.333	2.000	5.000
Doctors' behavior	1128	4.522	0.306	4.473	4.585	4.678	2.000	5.000
Doctors' skills	1127	4.478	0.324	4.419	4.548	4.645	2.000	5.000
Nurses' behavior	1126	4.381	0.388	4.333	4.481	4.580	2.000	5.000
Assistance for daily life	1126	4.453	0.347	4.382	4.536	4.632	2.000	5.000
Clinician teamwork	1126	4.399	0.353	4.336	4.474	4.580	1.714	5.000
Medical tests	1120	4.518	0.323	4.468	4.578	4.681	2.000	5.000
Medical surgeries	1051	4.571	0.430	4.543	4.683	4.776	1.000	5.000
Medical treatment	1114	4.575	0.351	4.546	4.657	4.745	2.000	5.000
Drip, injection, medicine, prescription	1118	4.493	0.401	4.440	4.578	4.676	1.000	5.000
Rehabilitation	949	4.379	0.417	4.250	4.441	4.593	1.000	5.000
Toilets and bathrooms	1125	4.019	0.476	3.746	4.049	4.357	1.000	5.000
Passageways, stairs, elevators	1123	4.272	0.376	4.098	4.329	4.509	1.000	5.000
My room	1125	4.072	0.448	3.818	4.108	4.398	1.500	5.000
Food	1124	4.040	0.399	3.885	4.092	4.268	2.000	5.000
Stores and interiors	1123	4.025	0.424	3.859	4.084	4.287	1.000	5.000
Explanations at discharge	1126	4.316	0.311	4.200	4.360	4.479	2.556	5.000

Author's personal copy

Overall satisfaction—inpatients	N	Mean	SD	p25	P50	P75	Min	Max
Generally satisfied	1130	4.317	0.360	4.223	4.388	4.520	1.000	5.000
Results of the treatment	1130	4.335	0.360	4.226	4.419	4.528	1.000	5.000
Length of the treatment	1129	4.233	0.384	4.152	4.318	4.433	1.000	5.000
Treatment	1130	4.381	0.356	4.326	4.460	4.559	1.000	5.000
Hospital	1129	4.233	0.353	4.155	4.298	4.406	1.000	5.000
Safety of medical services	1130	4.453	0.345	4.360	4.538	4.637	1.000	5.000
Clear explanations	991	4.492	0.314	4.440	4.549	4.635	1.000	5.000
Treatment was acceptable	991	4.467	0.356	4.416	4.555	4.644	1.000	5.000
Trust	991	4.534	0.321	4.486	4.600	4.689	1.000	5.000
Recommend to family and friends	1081	4.359	0.439	4.305	4.459	4.577	1.000	5.000

Panel B: Outpatients

Outpatients	N	Mean	SD	P25	P50	P75	Min	Max
Felt uneasy	1149	3.695	0.255	3.582	3.716	3.839	2.667	5.000
Inconvenient	1149	3.683	0.386	3.454	3.746	3.955	2.043	4.857
Bad impression	1149	4.051	0.333	3.938	4.114	4.250	1.667	5.000
Waiting time	1150	3.116	0.374	2.841	3.055	3.337	2.279	5.000
Waiting room	1150	3.789	0.318	3.583	3.818	4.000	2.688	5.000
Doctors' behavior	1150	4.165	0.221	4.018	4.155	4.318	3.333	5.000
Doctors' skills	1150	4.075	0.243	3.920	4.068	4.230	2.667	5.000
Nurses' behavior	1150	4.101	0.235	3.963	4.100	4.242	2.563	5.000
Medical tests	1149	4.108	0.248	3.964	4.119	4.273	2.667	5.000
Medical treatment	1149	4.303	0.235	4.182	4.325	4.455	3.000	5.000
Drip, injection, medicine, prescription	1149	4.299	0.253	4.156	4.319	4.467	2.750	5.000
Rehabilitation	1002	4.093	0.350	3.898	4.086	4.290	1.000	5.000
Treatment room	1150	4.141	0.269	3.982	4.157	4.326	1.000	5.000
Shops, ATM, interiors	1149	3.833	0.312	3.629	3.848	4.042	1.000	5.000
Procedures for payment	1150	3.859	0.327	3.670	3.865	4.070	1.000	5.000

Overall satisfaction—outpatients	N	Mean	SD	P25	P50	P75	Min	Max
Generally satisfied	1150	4.071	0.213	3.930	4.080	4.208	2.667	5.000
Results of the treatment	1150	4.026	0.217	3.891	4.025	4.161	2.667	5.000
Length of the treatment	1150	3.921	0.222	3.789	3.911	4.057	2.333	5.000
Treatment	1150	4.032	0.223	3.895	4.016	4.169	2.333	5.000
Hospital	1150	3.952	0.217	3.817	3.938	4.088	2.333	5.000
Safety of medical services	1150	4.156	0.204	4.026	4.157	4.289	2.667	5.000
Clear explanations	1150	4.174	0.216	4.052	4.172	4.311	2.000	5.000
Treatment was acceptable	1005	4.157	0.211	4.038	4.163	4.294	3.158	5.000
Trust	1005	4.266	0.197	4.150	4.276	4.396	3.000	5.000
Recommend to family and friends	1150	4.077	0.252	3.938	4.086	4.236	2.000	5.000

Author's personal copy

Appendix 3: Variables definition

Hospital characteristics	
Size	Number of beds available in the hospital, expressed in hundreds (i.e., number of beds/100)
Concentration	Number of hospitals (NHO and not) per 100 thousand inhabitants in the prefecture
Hospital	Indicator variable coded as 1 if the hospital is a general hospital and coded as 0 if the hospital is a sanatorium
Salary expenses (¥B)	Expenses due to salary compensation (billion Yen)
Bonus expenses (¥B)	Expenses due to bonus compensation (billion Yen)
Grant revenues (¥B)	Grant revenues received by the hospital (billion Yen)
Medical revenues (¥B)	Expenses related to medical services provided by the hospital (billion Yen)
Education revenues (¥B)	Expenses related to teaching (medical school, nursing school) (billion Yen)
R&D revenues (¥B)	Expenses related to clinical and academic research (billion Yen)
Other costs (¥B)	Total medical costs other than the categories identified above (billion Yen)

Appendix 4: Physician compensation at NHO

Salary schedule Each NHO post is classified into a certain grade in a salary schedule. The classification of the employee into a post is based on two factors: educational classification and experience. Most Japanese government agencies have ten *grades*. Within each grade employees receive raises in *steps*, which are based on time in grade. A sample of the pay scale for a Japanese government agency is provided below.

Author's personal copy

Salary per month (Yen)											
Grade											
		1	2	3	4	5	6	7	8	9	10
Steps	1	135,600	185,800	222,900	261,900	289,200	320,600	366,200	413,000	466,700	532,000
	5	140,100	192,800	230,200	270,200	298,200	329,800	376,300	422,800	479,000	544,700
	9	144,500	200,000	237,500	278,600	307,300	338,600	386,400	432,300	491,300	554,700
	13	149,800	207,000	244,900	287,000	316,400	347,200	397,100	441,300	503,600	562,100
	17	155,700	214,600	252,200	295,400	325,200	355,500	406,400	449,300	513,300	568,100
	21	161,600	222,000	260,100	303,800	333,500	363,500	414,800	456,500	519,000	572,900
	25	172,200	229,300	267,700	312,100	341,500	371,500	422,900	462,500	524,800	
	29	178,800	236,100	275,300	320,400	349,400	379,500	429,400	467,800	529,600	
	33	185,800	242,100	282,700	328,400	357,000	386,900	434,600	471,000	533,100	
	37	191,600	248,000	290,100	336,500	364,200	393,700	439,700	474,200	536,700	
	41	196,900	254,200	297,400	344,400	370,100	398,400	443,200	477,400	540,300	
	45	202,000	259,700	304,200	352,000	374,700	403,000	446,400	480,500		
	49	207,100	265,200	310,600	358,500	378,400	405,900	449,400			
	53	211,600	270,100	317,100	363,000	381,700	408,800	452,400			
	57	215,400	275,200	323,400	367,100	384,500	411,600	455,400			
	61	219,200	279,700	328,100	369,800	387,000	414,300	458,400			
	65	223,000	283,500	331,900	372,400	389,600	416,900				
	69	226,900	287,200	335,200	375,000	392,200	419,400				
	73	230,100	290,400	337,800	377,600	394,800	422,000				
	77	233,000	292,300	340,000	380,200	397,300	424,600				
	81	236,100	293,800	342,000	382,700	399,900					
	85	239,000	295,300	344,000	385,100	402,500					
	89	241,900	296,800	345,900	387,600						
	93	243,700	298,200	347,700	390,100						
	97	299,600	349,500								

Composition of salary In addition to the monthly salary, government employees also get *allowances* averaging at about 20% of base salary. The allowances include: living expenses (cost of living adjustment), housing allowance, commuter allowance, overtime allowance, cold weather allowance, and diligence allowance (typically based on the number of months of consecutive work in the previous 6-month period). There are some compensation adjustments related to macroeconomic conditions. Individual performance-based bonuses are not commonly found.

References

Abernethy, M. A., & Bouwens, J. (2005). Determinants of accounting innovation implementation. *ABACUS, 41*(3), 217–240.

Alpert, M., & Raiffa, H. (1982). A progress report on the training of probability assessors. In D. Kahneman, P. Slavic, & L. A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 294–305). Cambridge: Cambridge University Press.

Anderson, S. W., & Kimball, A. (2019). Evidence for the feedback role of performance measurement systems. *Management Science, 65,* 4385–4406.

Arrow, K. J. (1963). Uncertainty and the welfare economics of medical care. *The American Economic Review, 53*(5), 941–973.

Author's personal copy

Bandura, A., & Jourden, F. J. (1991). Self-Regulatory mechanisms governing the impact of social comparison on complex decision making. *Journal of Personality and Social Psychology, 60*(6), 941–951.
Banker, R. D., Potter, G., & Srinivasan, D. (2000). An empirical investigation of an incentive plan that includes nonfinancial performance measures. *The Accounting Review, 75*(1), 65–92.
Bardes, C. L. (2012). Defining "patient-centered medicine". *New England Journal of Medcine, 366*(9), 782–783.
Benoit, J.-P., & Dubra, J. (2011). Apparent overconfidence. *Econometrica, 79*(5), 1591–1625.
Bromwich, M. (2007). Economics in management accounting. In C. S. Chapman, A. G. Hopwood, & M. D. Shields (Eds.), *Handbook of management accounting* (pp. 137–162). Amsterdam: Elsevier.
Camerer, C., & Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *American Economic Review, 89*(1), 306–318.
Casas-Arce, P., & Martínez-Jerez, F. A. (2009). Relative performance compensation, contests, and dynamic incentives. *Management Science, 55*(8), 1306–1320.
Cavalluzzo, K. S., & Ittner, C. D. (2004). Implementing performance measurement innovations: Evidence from government. *Accounting, Organizations and Society, 29*(3–4), 243–267.
Chen, C. X. (2009). Who really matters? Revenue implications of stakeholder satisfaction in a health insurance company. *The Accounting Review, 84*(6), 1781–1804.
Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston: Houghton Mifflin.
Demski, J. S. (1972). Information improvement bounds. *Journal of Accounting Research, 10*(1), 58–76.
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*(2), 555–588.
Ederer, F. (2010). Feedback and motivation in dynamic tournaments. *Journal of Economics and Management Strategy, 19*(3), 733–769.
Evans, J. H. I., Hwang, Y., Nagarajan, N., & Shastri, K. (1997). Involuntary benchmarking and quality improvement: The effect of mandated public disclosure on hospitals. *Journal of Accounting, Auditing and Finance, 12*(3), 315–346.
Feltham, G. A. (1968). The value of information. *The Accounting Review, 43*(4), 684–696.
Fung, C. (2008). Systematic review: The evidence that publishing patient care performance data improves quality of care. *Annals of Internal Medicine, 148*(2), 111.
Goldstein, S. M. (2003). Employee development: An examination of service strategy in a high-contact service environment. *Production and Operations Management, 12*(2), 186–203.
Hammitt, J., & Shlyakhter, A. (1999). The Expected value of information and the probability of surprise. *Risk Analysis, 19*(1), 135–152.
Ho, R. (2013). *Handbook of univariate and multivariate data analysis with IBM SPSS* (2nd ed.). Boca Raton: CRC Press.
Holmstrom, B., & Milgrom, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law Economics and Organization, 7,* 24–52.
Ittner, C. D., & Larcker, D. F. (1998). Are nonfinancial measures leading indicators of financial performance? An analysis of customer satisfaction. *Journal of Accounting Research, 36,* 1–35.
Kahnemann, D., Slovic, P., & Tverski, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, MA: Cambridge University Press.
Kolstad, J. T. (2013). Information and quality when motivation is intrinsic: Evidence from surgeon report cards. In *NBER*. Retrieved August 18, 2016, from http://www.nber.org/papers/w18804.
Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*(6), 1121–1134.
Kruger, J., & Dunning, D. (2002). Unskilled and unaware–but why? A reply to Krueger and Mueller (2002). *Journal of Personality and Social Psychology, 82*(2), 189–192.
Lant, T. K., & Hewlin, P. F. (2002). Information cues and decision making: The effects of learning, momentum, and social comparison in competing teams. *Group and Organization Management, 27*(3), 374–407.
Levitt, B., & March, J. G. (1988). Organizational learning. *Annual Review of Sociology, 14,* 319–340.
Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1982). Calibration of probabilities: The state of the art to 1980. In D. Kahneman, P. Slavic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 305–334). Cambridge: Cambridge University Press.
Lindenauer, P., Remus, D., Roman, S., Rothberg, M., Benjamin, E., Ma, A., et al. (2007). Public reporting and pay for performance in hospital quality improvement. *The New England Journal of Medicine, 356*(5), 486–496.

Author's personal copy

Morgan, M. G., & Henrion, M. (1990). *Uncertainty: A guide to dealing with uncertainty in quantitative risk and policy analysis*. Cambridge: Cambridge University Press.

Morris, S., & Shin, H. S. (2002). Social value of public information. *The American Economic Review, 92*(5), 1521–1534.

Nagar, V., & Rajan, M. V. (2005). Measuring customer relationships: The case of the retail banking industry. *Management Science, 51*(6), 904–919.

Narayanan, V. G., & Davila, A. (1998). Using delegation and control systems to mitigate the trade-off between the performance-evaluation and belief-revision uses of accounting signals. *Journal of Accounting and Economics, 25*(3), 255–282.

Nesselroade, J. R., Stigler, S. M., & Baltes, P. B. (1980). Regression toward the mean and the study of change. *Psychological Bulletin, 88*(3), 622–637.

Pratt, J. W., Raiffa, H., & Schlaifer, R. O. (1995). *Introduction to Statistical Decision Theory*. Cambridge, MA: MIT Press.

Raiffa, H. (1968). *Decision analysis; introductory lectures on choices under uncertainty*. Boston: Addison-Wesley.

Ramanarayanan, S., & Snyder, J. (2012). Information disclosure and firm performance: Evidence from the dialysis industry. In *SSRN*. Working Paper # 2152796.

Sprinkle, G. (2003). Perspectives on experimental research in managerial accounting. *Accounting, Organizations and Society, 28,* 287–318.

Sugden, R. (2003). Reference-dependent subjective expected utility. *Journal of Economic Theory, 111*(2), 172–191.

Weissman, J. S., Annas, C. L., Epstein, A. M., Schneider, E. C., Clarridge, B., Kirle, L., et al. (2005). Error reporting and disclosure systems: Views from hospital leaders. *JAMA, 293*(11), 1359–1366.

Yokota, F., & Thompson, K. M. (2004). Value of information analysis in environmental health risk management decisions: Past, present, and future. *Risk Analysis, 24*(3), 635–650.

Zhang, X., & Tomblin, J. B. (2003). Explaining and controlling regression to the mean in longitudinal research design. *Journal of Speech, Language and Hearing Research, 46*(6), 1340–1351.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.