



Contents lists available at ScienceDirect

Health Policy

journal homepage: www.elsevier.com/locate/healthpol



Incentivising effort in governance of public hospitals: Development of a delegation-based alternative to activity-based remuneration

Rikke Søgaard^{a,b,c,*}, Søren Rud Kristensen^d, Mickael Bech^e

^a Department of Public Health, Bartholins Alle 2, Aarhus University, 8000 Aarhus, Denmark

^b Department of Clinical Medicine, Palle Juul-Jensens Boulevard 82, Aarhus University, 8200 Aarhus N, Denmark

^c Health Economics, CFK - Public Health and Quality Improvement, Oluf Palmes Alle 15, Central Denmark Region, 8200 Aarhus N, Denmark

^d Manchester Centre for Health Economics, Institute of Population Health, University of Manchester, Oxford Road, M13 9PL Manchester, UK

^e COHERE - Centre for Health Economic Research, Department of Business and Economics, University of Southern Denmark, Campusvej 55, 5230 Odense, Denmark

ARTICLE INFO

Article history:

Received 3 October 2014

Received in revised form 11 February 2015

Accepted 9 March 2015

Keywords:

Hospital management

Governance

Remuneration

Reimbursement

Delegation

ABSTRACT

This paper is a first examination of the development of an alternative to activity-based remuneration in public hospitals, which is currently being tested at nine hospital departments in a Danish region. The objective is to examine the process of delegating the authority of designing new incentive schemes from the principal (the regional government) to the agents (the hospital departments). We adopt a theoretical framework where, when deciding about delegation, the principal should trade off an initiative effect against the potential cost of loss of control. The initiative effect is evaluated by studying the development process and the resulting incentive schemes for each of the departments. Similarly, the potential cost of loss of control is evaluated by assessing the congruence between focus of the new incentive schemes and the principal's objectives. We observe a high impact of the effort incentive in the form of innovative and ambitious selection of projects by the agents, leading to nine very different solutions across departments. However, we also observe some incongruence between the principal's stated objectives and the revealed private interests of the agents. Although this is a baseline study involving high uncertainty about the future, the findings point at some issues with the delegation approach that could lead to inefficient outcomes.

© 2015 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Activity-based reimbursement (ABR) of hospitals on the basis of case mix tariffs was introduced in many countries

in the late 1990s and early 2000s [1]. Although still being the general model of reimbursement in most of these countries, there has been a debate about its appropriateness [2,3].

In the Danish context, the model has been criticised for being a barrier to new initiatives such as telemedicine, provision of all procedures in one day and collaboration between primary and secondary health-care sectors. In addition, the lack of a direct incentive for quality and in some cases even the existence of perverse incentives such

* Corresponding author at: Department of Public Health, Bartholins Alle 2, 8000 Aarhus, Aarhus University, Denmark. Tel.: +45 28991387.

E-mail addresses: rs@ph.au.dk (R. Søgaard),

soren.kristensen@manchester.ac.uk (S.R. Kristensen), mbe@sam.sdu.dk (M. Bech).

as higher payment for patients that acquire infections during admissions, as opposed to those who do not, have been highlighted. The Danish government thus formed a 'committee on better incentives in health care' in 2012. The committee recently announced its overall conclusion, which was that an updated governance model is desirable and that broader instruments than merely financial incentives should be considered [4].

Public governance through incentive schemes is known to be a complex matter creating intended as well as unintended organisational and behavioural responses, including 'creaming', 'tunnel vision' and complex dynamics between indicators, effects and outcomes. In addition, more than one level of principal–agent relationships often coexists. In the Danish context, the coexisting actors are the central government (the Department of Health), the regional governments, the local hospital managements and the hospital department managements. Incentivising quality of care in hospital governance models thus presents a major challenge to health policy and research [5].

On this background, a trial has been launched in one of the Danish regions. The regional government, which has the responsibility for the regional health-care system, states that its new objective is inspired by The Triple Aim [6]—a simultaneous focus on cost containment, patient-experienced quality and population health—in order to allow for a stronger focus on quality than what has been implied by the ABR-based model. During the trial, Diagnosis-Related Grouping (DRG)-based productivity measurement is suspended and replaced by global payment in combination with local incentive schemes at the individual hospital department level. Uniquely, the designing of these new incentive schemes has been delegated to the hospital departments. This paper is the first to present and discuss this new approach for hospital governance.

The objective of the study is to analyse the hospital departments' (the agents') behaviour when they are delegated the authority to develop their own performance indicators and targets by the regional government (the principal). The paper proceeds with a brief background on the context of the study in Section 2, a brief presentation of the proposed theoretical framework in Section 3 and a methods description in Section 4, including the premises for this new mode of governance. In Section 5, we focus on departments' motivation for engaging in the trial, their behaviour during the development process and on the resulting incentive schemes in terms of innovativeness, usefulness as a basis for performance management and congruence with the principal's objectives. Based on this, we then discuss our expectations for delegation as a novel approach for hospital governance in the Danish context in Section 6.

2. Background and Context

The study was conducted in Central Denmark Region, which serves a population of about 1.1 million inhabitants by five somatic hospital units. The funding of the region's hospitals is based on around 77% prospective global payment and around 23% from ABR based on per case

tariffs named DK-DRG [7]. The combination of global and per case payment is similar to models used in, e.g., Sweden, Germany, United Kingdom and Switzerland, whereas the relative weight of ABR is amongst the lowest in Europe together with those of, e.g., the Netherlands and the Czech Republic.

The national and the regional governments negotiate an overall financial agreement annually and Central Denmark Region endeavours to pass on the conditions directly to its hospitals. In 2014 and 2015, this has in practice meant extrapolation of last year's budget plus productivity increases of 2.4%, which is referred to as the baseline. Performance on the baseline as well as a number of other targets that are not directly linked to payment (listed in Appendix 1) are evaluated 4–6 times per year and discussed at dialogue meetings between the regional government and its hospitals. In the case of underperformance, hospitals pay back 50% of the DRG-based production value, whereas no additional payment is made for extra activity. At the hospital level, the baseline is typically translated into internal baselines for individual departments.

3. Analytical Framework

3.1. Hospital Contracts and Delegation

As a framework for analysis, we adopt the literature on authority and delegation developed by Aghion and Tirole [8]. This is a subliterature of the broader framework of incomplete contracts [9] focusing on within-firm decentralisation decisions as opposed to vertical integration of independent firms, which is thus well suited for the study of public organisations. Aghion and Tirole's model is focussed on the relation between delegation of authority and effort incentive at the stage of project selection, as opposed to the stage of project completion, which is treated by, e.g., Bester and Kräbmer [10]. For the present study, we define the *project* as the focus of performance management, which creates the effort incentive and, which is reflected in choice of performance indicators and targets.

In the framework by Aghion and Tirole [8], a principal employs an agent to solve a task. By definition, the principal has the formal authority to decide which projects the agent should spend time on but the principal can also delegate that decision to the agent. When deciding about delegation, the principal faces a trade-off between the expected benefit from the *initiative effect* and the expected *cost of loss of control*.

The initiative effect is due to delegation of authority giving the agent an incentive to invest more effort in information collection in relation to selecting the projects (and reducing the principal's need to do so) that will most likely lead to goal achievement. A key premise thus is that the agent is better informed about which behaviour supports the objective function and, accordingly, more efficient in defining the project than the principal is.

The possible cost of loss of control is due to the principal losing the opportunity to decide which projects the agent selects, because only the right to select is contractible (and not the resulting selection). This may result in the agent selecting projects with a high private benefit,

e.g. the pursuit of narrow clinical interests, as opposed to projects that more efficiently satisfy the principal's objective function.

The agent's private benefit depends on the choice of project, but so does the benefit for the principal. The optimal choice of whether to delegate authority to the agent thus depends on the level of *preference congruence* between the principal and the agent. Below a certain level of congruence, it is better for the principal not to delegate authority to the agent, as the benefit from the initiative effect cannot be expected to outweigh the cost of loss of control.

3.2. Criteria for Good Performance Indicators

We adopt criteria for public sector management that have been proposed by a working group under the Royal Statistical Society (RSS) in the UK [11]. The choice of generic criteria for provision of public service was made due to the broadness of the principal's objective function in relation to which we have found the literature on quality indicators too narrow, although it is substantial and well established [12–14].

The RSS criteria [11] include 14 points (the phrasing of some criteria have been moderately revised to attempt making them more concise): (a) congruence with principal's objective, (b) precise definition, (c) survey-based indicators should use a shared methodology across institutions, (d) consistency over time, (e) should obviate rather than create perverse incentives, (f) should be straightforward to interpret, (g) if not collected for the entire population, they should have sufficient coverage to ensure against misleading results, (h) technical properties should be adequate (sampling scheme, response rates, precision, etc.), (i) statistical potential to exhibit change within the timescale, (j) produced with appropriate frequency and timeliness to support performance, (k) conform to international standards if these exist, (l) should not impose an undue burden on those providing the information, (m) measurement cost should commensurate with likely performance management gain and (n) should be included in a performance management protocol.

4. Method and Sample

An observational study was conducted alongside the development process for new incentive schemes at nine hospital departments in a Danish region. The process was initiated in the autumn 2013 and considered to last until final proposals for new incentive schemes, including indicator index values, were available in the autumn 2014. The new incentive schemes will replace the current model of mixed global payment and ABR (including requirements for productivity increases) in combination with 100% global payment based on the 2013 budget for a test period of 2 years.

4.1. Inclusion of Departments

All of the region's five hospital units were invited to hand in a motivated and prioritised list of five hospital departments that could engage in the trial. The hospitals

were informed that the administration would then choose one department from each hospital unit according to criteria about broad representation of different specialities, departments where the DRG-based productivity focus could present a barrier to the implementation of new initiatives, at least one department with a high degree of collaboration with the primary care sector and at least one department receiving patients from other regions. Three hospital units got their first priority, one unit got its second priority and one unit was asked by the regional government administration to participate with a department that was not suggested.

4.2. Development Process and Data Collection

A series of three initial meetings was scheduled for each of the included departments. At the first meeting, regional government representatives presented the overall premises for the department. The ABR component of their remuneration model would be suspended and replaced by delegation of authority to design new incentive schemes. The suspension of ABR would include suspension of any associated financial sanctions if activity and productivity targets were unmet. The department would thus be allocated an annual global budget for the duration of the development process and follow up from 2014–2016. It was emphasised that departments should not exceed this budget, that the regional government would follow up on the overall activity level and that the new incentives schemes should be kept simple with a total number of 3–5 indicators, preferably based on existing data. Upon this introduction, departments were delegated the responsibility for specifying their new schemes. No additional boundaries as to what ideas or projects that could be prioritised by the departments were made explicit.

During the meetings, each department was typically represented by a team of a chief physician and a chief nurse, and in some cases by a chief secretary and other key management staff. This team typically drove the process while a representative from the regional government acted as coordinator by facilitating meetings, taking minutes and coordinating the support from service departments, including region-level data offices. Hospital-level support was represented by participation of staff from management, finance, quality and/or data management offices on an ad hoc basis. Each meeting had a group size of 5–15 people with declining numbers as the focus of the process narrowed.

The research team was given access to all of these meetings and the related working documents. The entire investigation was guided by an overall set of questions, including questions about the department's motivation for participation and the rationale for, the specification of and the mechanism of action for the chosen performance indicators. Semi-structured interviews were conducted by time of the last meeting in the scheduled development process (or at a private interview session if the individual department preferred so, or if the scheduled meetings did not allow for it due to time pressure). Refinement of incentive schemes continued in a smaller forum after

the more formal development process, primarily for the purpose of developing and testing procedures for measurement, including establishment of indicator index values. The final indicators and targets were reported in to the regional government administration for presentation for politicians.

4.3. Analysis

Interviews were transcribed and the minutes of meetings and ad hoc documents used or generated during the process (statistics on historical activity, patients served, collaboration with other sectors, indicators from clinical databases, etc.) were systematically made available to the research team in order to facilitate simple thematic analysis across departments [15].

The potential benefit of delegation (the initiative effect) was assessed by asking agents directly about their motivation and by assessment of the product of their effort on project selection: the number and type of chosen indicators for performance measurement. Furthermore, the usefulness of indicators for performance management was assessed against the RSS criteria [11]. Indicators were evaluated individually and assigned a value of 0 (criterion not satisfied) or 1 (criterion satisfied) in order to report an overall % of indicators with positive evaluation scores. One investigator made a complete assessment of all indicators in a scoring sheet while making a list of any issues that could be discussed. Having discussed these issues with the co-investigators, the process was repeated until the assessment was fully consensus-based.

The potential drawback of delegation (the cost of loss of control) was evaluated in terms of incongruence between the chosen incentive schemes and the likely pursuit of The Triple Aim ambition [6]. The incentives created by individual indicators were classified according to its potential effect on the triple aim dimensions of cost, quality and outcome and/or to other dimensions (indicators were allowed to target multiple dimensions at the same time). Similar to the assessment of the RSS criteria, this assessment was undertaken by one investigator and discussed with co-investigators until fully consensus-based.

5. Results

5.1. Participating Hospital Departments

The participating departments include: a department of acute medicine, a diagnostic centre (union of radiology and all specialities of medicine to facilitate more efficient diagnostics), a department of orthopaedic surgery, a department of general medicine and a cluster of five departments organised in a large Head-Neuro centre (neurology, neurosurgery, oral and maxillofacial surgery, ear–nose–throat and ophthalmology) as presented in Table 1. These departments represent a total yearly budget of 1.2 billion DKK, which corresponds to 8% of the regional hospital budget, 23% of the number of unique patients and 10% of the number of beds.

5.2. Scope of Trial Act and Departments' Motivation

The regional government introduced the trial act as a new mode of governance that would allow the departments “to manage by quality rather than productivity”. Reference was made explicitly to the triple aim ambition but also to “more health for the same money” and “giving the right patient the right treatment at the right time”. It was thus observed that the explanation of the objective function was not entirely clear in terms of the interpretation and weight of individual dimensions and how potential conflict between them should be handled.

Almost all departments expressed their understanding of DRG-based ABR as a barrier for initiatives that could improve efficiency and quality of their activities. Accordingly, the trial act was seen as an opportunity for development or implementation of initiatives that had not previously been developed or implemented due to ABR being a barrier. Some departments explained how they had already taken up initiatives such as providing diagnostic examinations in one visit, substituting inpatient with outpatient activity and using telemedicine irrespective of the resulting decreases in productivity according to the DRG-system. A chief nurse stated “... we have worked with this new form of management focus since 2007. At that time, we started substituting inpatient with outpatient activity, which is a sick business if you think in DRG terms”.

For at least two of the surgical specialities, department managements appeared to be more split between seeing the current model as a barrier and seeing it as a “you know what you have” model under which they had learned to perform well. The principles of ABR appeared to have been adapted to an extent where these departments quantified the success of their management strategy in terms of production value. For example, a chief nurse noted that “a new incentive scheme has to come with a replacement for the DRG-system's quantification of the load of patients in our beds” and a chief doctor opposed to the premise of no longer getting performance statistics on DRG-based production value “... we who are in the business, are we no longer allowed to have any knowledge about how we perform?”

It should also be noted that some reservation towards engaging in the development of a new incentive scheme was explained by a concern that the whole trial act was just a cost saving exercise. A chief doctor noted “and when we are undertaking our activity a little bit more smart in two years' time, at least on an organisational level, it will not be remunerated in DRG terms. And where does that leave us.” This reservation also appeared to the most significant in the surgical specialities.

5.3. Project Selection

During the process, all departments had several ideas and beliefs for the path to improved performance. The resulting projects ranged from adherence to clinical guidelines, structural reorganisation (e.g. from inpatient to outpatient and from outpatient to telemedicine, etc.), reduction of process time (e.g. from referral to appointment and from admission to treatment initiation), optimisation

Table 1
Characterisation of participating hospital departments.

	Number of unique patients per year	Number of beds	% bed occupancy	% of total regional budget	% of department budget on variable costs (other than staff)	% of referrals seen within 14 days (service target)	% of productivity goal ^a
Department of acute medicine	3,982	26	115	0.68	6	29	103
Diagnostic centre	16,360	88	85	1.58	27	30	111
Department of orthopaedic surgery	14,820	33	98	0.63	19	32	NA
Department of general medicine	16,220	82	105	1.14	14	NA	98
Head-Neuro centre							
Neurology	12,804	42	82	0.64	41	NA	112
Neurosurgery	5,559	37	89	0.75	32	14	95
Oral and maxillofacial surgery	5,020	6	78	0.25	21	NA	101
Ear–nose–throat	11,300	19	84	1.02	16	35	107
Ophthalmology	17,187	8	46	1.18	48	NA	108
Total	103,252	341	87	7.87	25	NA	NA

Note: All statistics refer to the year 2012.

^a The extent to which departments have achieved the baseline activity as defined by Diagnosis-Related-Grouping-based case-mix tariffs. The extent to which departments correct baseline activity during the year varies and, accordingly, the values should be interpreted with caution.

of the service offer (e.g. that adherence is made to a standard postoperative plan and that the number of visits per treatment is minimised) and incentivising research (e.g. by monitoring the number of patients included in trials and the number of peer-reviewed publications) through improving the ultimate outcome from a societal perspective (e.g. that more patients return to work after surgery and that more patients survive after cancer treatment).

On an indicator level, an overweight of process measures was observed although, intuitively, one would expect agents to have preference over broader outcome measures, which leaves greater room for manoeuvre. Most departments expressed a wish for a generic measure of patient-reported quality experience. That idea was, however, discarded on grounds of a lack of existing data structures and it was noted that the establishment of a universal system for routine measurement was seen as the responsibility of a central office under the regional government. Instead, some departments worked with the idea of using patient panels or focus groups in order to gain insight into the user perspective and to monitor their performance in a more qualitative manner.

A focus point of the investigation was the role of risk perception and whether the overweight of process measures could be due to risk aversion, as greater uncertainty is inherently associated with the alternative of outcome measures. This was generally disconfirmed and instead explained by more pragmatic arguments such as data availability and response times that actually allow for performance management within a reasonable time frame. Also, it was observed that the long tradition for clinical quality databases, and the type of indicators typically included in clinical databases, appeared to frame the brainstorming about new indicators. It thus seemed that most departments paid little attention to the risk balance, and how the new mode of governance could affect it.

Another signal of indicator choices not being driven by risk aversion was that only limited concern was expressed in terms of discretion over performance, although the performance of most departments is interrelated to other agents' performance (e.g. primary care or other in-house specialities). For example, a common priority was to treat

'the right' patients although most departments have no influence on how many and what patients are referred to them. Nevertheless, a so-called hit rate was adopted as an indicator for the departments organised in the Head-Neuro centre. One department had good experience with affecting referral rates by providing referring doctors with benchmark statistics about the proportion of referred patients that was actually treated.

Some departments focussed on also monitoring potentially negative effects of their new schemes. One example is a department for acute medicine, which overall focussed on reducing various process times during admission, additionally chose an indicator of readmissions in order to monitor if the accelerated service compromised the overall treatment quality. Another example is a department for oral and maxillofacial surgery focussing on efficiency of treatment strategies by substituting inpatient with outpatient service and by eliminating unnecessary follow up visits, which similarly chose to monitor if this focus compromised patient- and employee satisfaction. Table 2 provides a complete list of the schemes that were ultimately chosen.

Generally, the behaviour of departments during the process appeared to be non-strategic, which was interpreted from the fact that many proposals for indicators were ambitious and involved great uncertainty in terms of the likelihood for good performance.

5.4. Evaluation of the Chosen Indicators as a Basis for Performance Management

The most critical evaluation results concerned, first, the precise definition of the indicator and the associated target (criterion b), which was in many cases lacking due to data structures not being readily available. Second, the coverage of individual indicators (criterion g), which appeared to be problematic as many indicators either focussed on a particular diagnostic group that represented only a fraction of the department's target population or, had a percentile-based target such as "90% should have an appointment within 14 days", which could represent a perverse incentive for ignoring cases with a perceived low likelihood for achieving the target. Third, technical properties of indicators such as

Table 2

Indicators and targets chosen by the hospital departments.

Department	Indicator	Target
Acute medicine	Time from admission to specialist doctor evaluation ≤ 1 h	80%
	Time from admission to treatment plan ≤ 4 h	80%
	Time from admission to discharge	10% reduction
	Patient satisfaction and security	80% positive
	Readmission due to the same diagnosis within 30 days	No increase
Diagnostic centre	Time to diagnosis ^a	No target
	Time to treatment ^a	No target
	Number of patients receiving a diagnosis after one visit ^a	>80%
	Number of examinations per patient ^a	No target
	Number of examinations per patient per day ^a	No target
	Adherence to existing quality indicators ^{a,b}	Awaits baseline
	Re-referral to department ^a	<5%
Ear–nose–throat	Patient satisfaction ^{a,b}	>95% positive/no target
	Conduction of dialogue meetings with selected patients ^a	No target
	Patient conferred at multidisciplinary conference ^a	90%
	Adherence to clinical guidelines ^a	90%
	Cancer-specific survival 1 year postoperative ^a	90%
	Complication after surgery ^a	<1%
	Adherence to postoperative plan ^a	90%
	Disease-specific quality of life score 6 months postoperative ^a	>20 points
	Disease-specific quality of life after treatment ^a	Awaits data collection
	Admissions per patient ^b	Decrease
General medicine	Number of hospital-acquired infections ^b	Decrease
	Service provision by telemedicine solutions ^a	Increase
	Service provision by home visits ^b	Increase
	Empowerment of patients ^a	Increase
	Patient-experienced quality	Increase
Neurology	Enabling new junior doctors to work independently	90% can undertake night shift after 6 weeks of employment
	Enabling new doctors under specialisation to work independently	90% can undertake night shift after 3 weeks of employment
	Enabling new doctors to work independently in thrombolysis clinic	90% can undertake first-line shift after 2 months
	Inclusion of patients into research trial ^a	25 patients per year
	Production of peer-reviewed publications weighted according to impact factor	Maintain current level
	Substitution of inpatient immunoglobulin therapy with telemedicine solutions	>80%
	Number of persons treated	No decrease
	Ad hoc access to outpatient clinic for chronic patients ^a	50% within 1 day and 90% within 14 days
Neurosurgery	Number of patients who have returned to work 1 year after spine surgery ^a	>50%
	Number of patients with pain reduction 1 year after spine surgery ^a	>50%
	Number of patients being readmitted within 30 days after spine surgery ^a	<5%
	Patient satisfaction after neck surgery	Awaits definition
	30-day survival after tumour surgery ^a	>95%
	Re-operations due to incompleteness of first tumour surgery ^a	<10%
	Re-admissions within 30 days after primary surgery due to complications ^a	<5%
	1-year survival after glioblastoma multiforme surgery ^a	>50%
	Number of persons treated ^{a,b}	No decrease
	Response to MR scanning after cancer treatment within 2 days ^a	>90%

Table 2 (Continued)

Department	Indicator	Target
Ophthalmology	Number of surgeries being cancelled	Decrease
	Number of outpatient visits per patient per year ^{a,b}	No target
	Number of unique patients treated ^a	No decrease
	Time from the patient's arrival to departure in the outpatient clinic ^{a,b}	Decrease
Orthopaedics	Number of persons treated ^a	Maintain current level
	Quality according to selected parameters in national clinical databases ^{a,b}	Maintain and increase
	Patient involvement	Increase
Oral and maxillofacial surgery	Number of patients having their surgery in the outpatient clinic (instead of inpatient) ^{a,b}	Increase
	Number of inpatient patients ^{a,b}	Decrease
	Number of outpatient visits per treatment	Decrease
	Number of persons treated	No decrease
	Patient satisfaction	No decrease
	Employee satisfaction	No decrease
Centre management Head-Neuro	Proportion of patients who have had their health problem solved by time of termination of treatment	Increase by 2% per year
	Re-admissions between 8 h and 30 days after discharge	Decrease by 5% over 2 years
	Time from referral to termination (outpatients)	Decrease by 5% over 2 years
	Time from referral to admission (elective inpatients)	Decrease by 5% over 2 years
	Time from referral to admission (acute inpatients)	Decrease by 5% over 2 years
	Time from admission to discharge (elective inpatients)	Decrease by 5% over 2 years
	Did the patient get her/his health problem solved?	Awaits baseline
	Cancellations and non-attendance	Decrease by 20% over 2 years
	Proportion of referred patients that are ultimately diagnosed with a diagnosis relevant for the referral	100%
	Proportion of referred patients that are ultimately treated in the centre	Awaits definition
	Number of outpatient visits per outpatient treatment per year	Decrease by 6% over 2 years
	Number of admissions per outpatient treatment	Decrease
	Re-operations	Decrease by 5% per year

^a Indicator restricted to proportion of production.^b Indicator monitored overall and for relevant subgroups.

Table 3

Assessment results of the usefulness of the chosen indicators for performance management.

Department	% of indicators that satisfies criterion														Average
	a	b	c	d	e	f	g	h	i	j	k	l	m	n	
Acute medicine	100	80	NA	80	80	100	60	80	100	80	NA	100	100	0	80
Diagnostic centre	100	25	100	88	100	88	38	88	88	88	NA	100	100	0	77
Ear–nose–throat	100	71	100	86	86	71	0	57	43	43	NA	86	86	0	64
General medicine	100	33	100	33	100	50	50	33	83	67	NA	100	83	0	64
Neurology	63	75	NA	88	50	88	63	50	75	75	NA	38	88	0	63
Neurosurgery	90	70	100	100	80	80	10	70	60	60	NA	90	70	0	68
Ophthalmology	75	75	NA	100	100	100	75	100	100	100	NA	75	75	0	81
Orthopaedics	67	33	100	33	100	67	100	33	33	33	NA	100	67	0	59
Oral and maxillofacial surgery	67	50	100	100	67	100	100	50	83	67	NA	100	100	0	76
Average across departments	85	57	100	79	85	83	55	62	74	68	NA	88	85	0	71
Centre management Head-Neuro	100	69	NA	100	85	77	100	69	85	100	NA	85	85	0	79

Note: the criteria were adopted from Bird et al. [11] and are described in Section 3.2. NA, not applicable.

sampling schemes and response rates (criterion h) were less optimal for indicators based on low-volume production, ad hoc patient-reported measures or ad hoc registration in the clinic. Fourth, none of the indicators were included in a formal performance management protocol (criterion n).

A couple of scoring issues in relation to application of the RSS criteria should be noted. Criteria relating to consistency over time and effort spend on monitoring (criteria d and j) will be easier to satisfy within a trial context of a limited period as opposed to a context of routine practice. The criterion about survey-based indicators having to use a shared methodology across institutions (criterion c) could be evaluated for a few indicators only, as the new incentive schemes were not intended to enable comparison across institutions and therefore (rightfully) focussed on intra-institutional development over time. In relation to the burden on those providing the information and the balance with the performance management gain (criteria l and m), most indicators received positive evaluation results because the reliance on primarily existing data was laid out as a condition by the principal. This was generally adhered to except in relation to the call for a common patient satisfaction indicator. Altogether, these issues mean that

the evaluation results should be seen as maximum scores. Table 3 shows the criteria assessment results.

5.5. Congruence between the New Incentive Schemes and the Principal's Objectives

Table 4 provides a classification of the chosen indicators according to The Triple Aim and other objectives that have not been explicitly stated by the regional government. Each indicator was allowed to have more than one focus, but almost no indicators were found to satisfy the simultaneous pursuit of all of The Triple Aims. The ratio between simple count of indicators (average across departments was 6) versus count of indicator focus (average across departments was 9) reflect the fact that most indicators were found to address 1 or 2 dimensions of The Triple Aim.

The overall congruence between the preferences of the agents, revealed in the chosen incentive schemes, and the stated preferences of the principal was affected by a main focus on quality and a more limited focus on population health and cost containment. During the development process, several arguments for not choosing health-related indicators were made. First, the lack of a system for routine

Table 4

Focus of the new incentive schemes according to the principal's objectives.

Department	Number of indicators	Number of indicator focuses ^a	% of indicator focus according to objective			
			Cost containment ^a	Patient-experienced quality	Population health	Other ^b
Acute medicine	5	7	14	71	14	0
Diagnostic centre	9	13	23	69	8	0
Ear–nose–throat	7	11	9	36	55	0
General medicine	6	9	22	56	22	0
Neurology	8	8	38	25	0	38
Neurosurgery	10	17	24	29	41	6
Ophthalmology	4	6	33	50	0	17
Orthopaedics	3	4	0	50	25	25
Oral and maxillofacial surgery	6	10	30	50	0	20
Average across departments	6	9	21	49	18	12
Centre management Head-Neuro	13	25	32	48	20	0

^a Considering the fact that some indicators aim at more than one objective simultaneously.^b Focus that is not directly related to the principal's objectives (includes, e.g., treatment volume, satisfaction of employees and research activities). NA, not applicable.

measurement of population health, second, that population health is largely outside the domain of control for the departments and, third, that the lag period from provision of service to health response is often too long to make such indicators of operational value. It should be noted, though, that most departments called for the establishment of a common system for routine reporting of patient satisfaction and that the development of such a system was initiated by the principal, however, with a time horizon exceeding that of the present trial.

6. Discussion and Conclusions

In this study, we have followed nine hospital departments and a centre management that were delegated the authority to define new incentive schemes to replace ABR. We observed some evidence of, first, that delegation is feasible, as the agents generally appeared highly motivated and engaged in the process. Second, that agents generally based project selection on innovative ambitions, which could have substantial value to the principal. Third, that the indicators chosen by agents suffered some weaknesses relating to coverage of the entire objective function of the principal and some more technical issues affecting their usefulness for performance management. Fourth and finally, that agents generally prioritised quality over other objectives – often in the form of process quality, which may or may not spill over to patient-experienced quality.

The application of the Royal Statistical Society's criteria [11] requires qualitative judgement. While it is relatively easy to determine whether an indicator definition is precise, it is more complex to assess whether the indicator obviates rather than creates perverse incentives. For the present application, this required a relatively deep understanding of rationales and mechanisms of indicator action, which we sought by observing the incentive scheme development processes and, ultimately, by asking the departments to explain any issues that pointed in more than one direction. It is important to note that we use these scorings as evidence on the overall feasibility of letting departments specify indicators. We cannot rule out misclassification bias and it would be inappropriate to use the scorings as measures of the comparative ability of individual departments.

It has long been recognised that a principal has to take into consideration the private interests of agents when choosing his strategy, as the level of effort on the part of agents will depend on their perception of the likelihood that successful performance will lead to an outcome they desire [16]. Projects that are not fully aligned with the principal's objective or projects that do not cover the entire objective function of the principal may therefore still lead to better goal achievement than if agents are forced to adopt fully aligned projects on which they will not apply high levels of effort.

Similar to the findings of Kristensen et al. [17], we found that delegation works well as a tool for identifying local needs and priorities that a central decision maker would have difficulties recognising, but also that the local capacities for the complex task of designing performance indicators are not necessarily present at clinical

departments. In this light, it seems relevant to consider whether the identified initiative effect can be maintained if a professional management office undertakes the precise definition of indicators and targets.

In addition to the adopted theoretical framework by Aghion and Tirole [8], other theoretical views would be relevant to explore in the future. Hart and Holmstrom challenge the idea that the principal can just take back the authority after projects have been selected but show that this will only happen in extreme cases, because the principal will then lose his trustworthiness [18]. Bester examines the special case of multiple decisions about projects and shows that unless strong complementarities exist between these decisions, delegation may enhance efficiency [19].

Using Donabedian's definition of health-care quality [20,21], the main focus of agents in this study was on process quality. A strength of process quality measures is that they provide quick, clear and interpretable feedback as opposed to more ultimate outcome measures, which are often outside the domain of full control and therefore less likely to be sensitive and responsive to the right behaviour. It is thus hard to assess whether the chosen projects could lead to goal achievement and such an assessment, at least partly, relies on the link between process quality and patient-experienced quality, which has not yet been fully established [22,23].

Although the new schemes are based on non-financial incentives, it should be noted that there is a large related literature on quality indicators for pay-for-performance schemes [24]. One of the potential caveats of paying for performance is the risk of getting what you pay for, not what you hope for [25,26]. This can be seen as an analogue to the issues of incomplete coverage in the present study. Some indicators were developed for specific diagnostic groups and some targets were set for percentiles only. This could create perverse incentives for investing effort in what is actually rewarded – whether financially or in terms of professional acknowledgement and pride.

In terms of the overall expectations for the delegation model as a useful device for hospital governance, the main focus of agents' signals that the likelihood of perfect agentship is the highest on the quality dimension. The key issue in relation to future perspectives is the effort incentive on project completion, as described by, e.g., Bester and Kräbmer [10]. The overall attractiveness of delegation as part of a hospital governance model should be based on both the ex ante effort spent on project selection (assessed in the present study) and on the ex post effort spent on project completion. The overall value of the delegation-based approach thus awaits assessment of the course of time in a comparative analysis.

By this study, we do not suggest that conventional performance measures relating to, e.g., activity and productivity levels should be discarded. If the delegation-based approach shows to be effective in incentivising effort, it might be used as an add-on to existing performance measures chosen by the principal. One of the main arguments for replacing the mixed payment model with 100% global payment, in this trial context, was to facilitate the agents' ability to think innovatively and free of financial risk. It

can then be discussed whether the keeping of the DRG-based production value from the departments during the trial influenced project selection – and whether it will influence the departments' longer-term performance on the chosen indicators. As long as payment is not directly linked, activity and productivity measures such as DRG-based production value could be included in future schemes without compromising innovation. Such measures could support the departments' refining of projects, the regional government's monitoring of performance and, not least, inform budget negotiations.

Conflict of interest

The authors have nothing to disclose.

Acknowledgements

The authors would like to thank the regional government administration and the hospital departments for granting the research team access to the development process and for providing relevant material. Professor Tor Iversen, University of Oslo, is acknowledged for commenting on an earlier version of this manuscript at the 35th annual meeting for the Nordic Health Economists' Study Group.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.healthpol.2015.03.005>.

References

- [1] Busse, Reinhard, Geissler, Alexander, Quentin, Wilm, et al. Diagnosis-related groups in Europe: moving towards transparency, efficiency and quality in hospitals, Vol. 2011. McGraw-Hill; 2011.
- [2] Appleby J, Harrison T, Hawkins L, Dixon A. Payment by results: how can payment systems help to deliver better care? King's Fund 2012.
- [3] O'Reilly J, Busse R, Häkkinen U, Or Z, Street A, Wiley M. Paying for hospital care: the experience with implementing activity-based funding in five European countries. *Health Economics, Policy and Law* 2012;7:73–101.
- [4] Better Incentives in Health Care [in Danish]. 2013.
- [5] Burgess S. The role of incentives in the public sector: issues and evidence. *Oxford Rev Econ Policy* 2003;19:285–300.
- [6] Berwick DM, Nolan TW, Whittington J. The triple aim: care, health, and cost. *Health Affairs (Millwood)* 2008;27:759–69.
- [7] Bech M. In: Pedersen KM, Petersen NC, editors. Is DRG also the Future Remuneration Model? [in Danish]. Munksgaard: København; 2014. p. 331–42.
- [8] Aghion P, Tirole J. Formal and real authority in organizations. *Journal of Political Economy* 1997;105:1–29.
- [9] Grossman SJ, Hart OD. The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration; 1986.
- [10] Bester H, Krähmer D. Delegation and incentives. *The RAND Journal of Economics* 2008;39:664–82.
- [11] Bird SM, Sir David C, Farewell VT, Harvey G, Tim H, Peter CS. Performance indicators: good, bad, and ugly. *Journal of the Royal Statistical Society Series A* 2005;168:1–27.
- [12] Rubin HR, Pronovost P, Diette GB. The advantages and disadvantages of process-based measures of health care quality. *International Journal for Quality in Health Care* 2001;13:469–74.
- [13] Mainz J. Developing evidence-based clinical indicators: a state of the art methods primer. *International Journal for Quality in Health Care* 2003;15(Suppl 1):i5–11.
- [14] Kotter T, Blozik E, Scherer M. Methods for the guideline-based development of quality indicators—a systematic review. *Implementation Science* 2012;7:21.
- [15] Kvale S, Brinkmann S. *InterViews. Learning the Craft of Qualitative Research Interviewing*, 2nd edition Sage; 2008.
- [16] Guth WD, Macmillan IC. Strategy implementation versus middle management self-interest. *Strategic Management Journal* 1986;7:313–27.
- [17] Kristensen SR, McDonald R, Sutton M. Should pay-for-performance schemes be locally designed? Evidence from the Commissioning for Quality and Innovation (CQUIN) Framework. *Journal of Health Services Research & Policy* 2013;18:38–49.
- [18] Hart OD, Holmstrom B. A theory of firm scope. *Quarterly Journal of Economics* 2010;125:483–513.
- [19] Bester H. Externalities, communication and the allocation of decision rights. *Economic Theory* 2008;41:269–96.
- [20] Donabedian A. The quality of care how can it be assessed? *JAMA The Journal of the American Medical Association* 1988;12:1743–8.
- [21] Donabedian A. Evaluating the quality of medical care. 1966. *Milbank Quarterly* 2005;83:691–729.
- [22] Sofaer S, Crofton C, Goldstein E, Hoy E, Crabb J. What do consumers want to know about the quality of care in hospitals? *Health Services Research* 2005;40:2018–36.
- [23] Sofaer S, Firminger K. Patient perceptions of the quality of health services. *Annual Review of Public Health* 2005;26:513–59.
- [24] Petersen LA, Woodard LD, Urech T, Daw C, Sookanan S. Does pay-for-performance improve the quality of health care? *Annals of Internal Medicine* 2006;145:265–72.
- [25] Kerr S. On the folly of rewarding A, while hoping for B. *The Academy of Management Journal* 1975;18:769–83.
- [26] Holmstrom B, Milgrom P. Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, and Organization* 1991;7:24–52.