Accepted Manuscript

Assessing physician productivity following Norwegian hospital reform: A panel and data envelopment analysis

Karl Arne Johannessen, Sverre A.C. Kittelsen, Terje P. Hagen

PII: S0277-9536(17)30008-4

DOI: 10.1016/j.socscimed.2017.01.008

Reference: SSM 10997

To appear in: *Social Science & Medicine*

Received Date: 16 June 2016

Revised Date: 2 January 2017

Accepted Date: 5 January 2017

Social Science & Medicine manuscript number SSM-D-16-01840

Assessing physician productivity following Norwegian hospital reform: a panel and data envelopment analysis

Karl Arne Johannessen[1], MD, PhD, Sverre AC Kittelsen[2], PhD and Terje P. Hagen[3], Professor, PhD.

[1]Sykehuspartner Health Enterprise, Oslo, Norway.

[2]Frisch Centre Oslo, Norway, and Institute of Health and Society, University of Oslo, Norway.

[3]Institute of Health and Society, University of Oslo, Norway.

Corresponding author:

Karl Arne Johannessen
Sykehuspartner Health Enterprise
Hoffsveien 1 D
275 Oslo
Norway
xakljo@sykehuspartner.no
Phone: +47 90 89 50 75

Assessing physician productivity following Norwegian hospital reform: a panel and data envelopment analysis

ABSTRACT

Background:

Although health care reforms may improve efficiency at the macro level, less is known regarding their effects on the utilization of health care personnel. Following the 2002 Norwegian hospital reform, we studied the productivity of the physician workforce and the effect of personnel mix on this measure in all nineteen Norwegian hospitals from 2001 to 2013.

Methods:

We used panel analysis and non-parametric data envelopment analysis (DEA) to study physician productivity defined as patient treatments per full-time equivalent (FTE) physician. Resource variables were FTE and salary costs of physicians, nurses, secretaries, and other personnel. Patient metrics were number of patients treated by hospitalization, daycare, and outpatient treatments, as well as corresponding diagnosis-related group (DRG) scores accounting for differences in patient mix. Research publications and the fraction of residents/FTE physicians were used as proxies for research and physician training.

Results:

The number of patients treated increased by 47% and the DRG scores by 35%, but there were no significant increases in any of the activity measures per FTE physician. Total DRG per FTE physician declined by 6% ($p < 0.05$). In the panel analysis, more nurses and secretaries per FTE physician correlated positively with physician productivity, whereas physician salary was neutral. In 2013, there was a 12%–80% difference between the hospitals with the highest and lowest physician productivity in the differing treatment modalities. In the DEA, cost efficiency did not change in the study period, but allocative efficiency decreased significantly. Bootstrapped estimates indicated that the use of physicians was too high and the use of auxiliary nurses and secretaries was too low.

Conclusions:

Our measures of physician productivity declined from 2001 to 2013. More support staff was a significant variable for predicting physician productivity. Personnel mix developments in the study period were unfavorable with respect to physician productivity.

Keywords: Norway; Physician productivity; Personnel mix; Health care reform; Panel analysis; Data envelopment analysis.

Introduction

The success of modern medicine may in fact become its most serious challenge. Supported by accelerating technological developments, modern medicine is pushing frontiers at increasing speeds. These rapid advancements may exceed the capacities of economic and human resources available in the future. Novel treatments for new patient groups that seemed impossible a few years ago, along with increasing complexity and specialization, have resulted in a growing demand for health personnel. With the limited workforce and labor supply confronting most developed health care systems, the continued rapid development of medicine may not be sustainable (Cooper, 2004; Simoens & Hurst, 2006; Staiger, Auerbach, & Buerhaus, 2009, 2010; Williams, Sun, Ross, & Thomas, 2010).

The need to improve efficiency is therefore urgent. To cope with economic challenges, many financial, political, and organizational investments have been made in most developed health care systems in recent decades (Busse, Schreyogg, & Smith, 2008; Magnussen, 2009; Oliver & Mossialos, 2005; Rickman & McGuire, 1999; Rumbold, 2015; Tuohy, 1999; Wiley, 2005; Wilsford, 1994). In 2002, aiming to reduce political interference, a Norwegian hospital reform transformed hospitals into enterprises owned by the government but with full autonomy. One of the major goals was to utilize personnel more efficiently by granting hospitals the power to negotiate the salaries of their own staff members and to decide on their own personnel strategies (E. Biorn, Hagen, T. P., Iversen, T., Magnussen, J., 2010; Tiemann & Schreyogg, 2012). The intention was to create solutions that would stimulate and reward personnel—physicians in particular—for increasing their competence and clinical efficiency, based on the needs of individual institutions.

Hospital productivity and efficiency have been studied extensively at the institutional level, both within individual health care systems and across different national systems. The approaches taken by these studies vary, with some using advanced techniques such as data envelopment analysis (DEA) and stochastic frontier analysis (SFA) and others relying on less advanced techniques (Castelli, Street, Verzulli, & Ward, 2015; Hollingsworth, 2008; Storfa & Wilson, 2015; Varabyova & Schreyogg, 2013). Some studies have examined efficiency within particular specialties and at the individual level (Askildsen, 2006; Bloor, Maynard, & Freemantle, 2004; Laudicella, Olsen, & Street, 2010; Romley, Goldman, & Sood, 2015; Schreyogg, 2008; Tiemann, 2008)). However, the productivity of health personnel is difficult to assess because of the multiple tasks of patient treatment, teaching, and research, and because of differences among specialties regarding diversity in patient treatments and care levels. No single measure can fully reflect this, and we are often left with macro parameters and proxies, such as billing and reimbursement. Furthermore, because productivity is only one aspect of health care systems, it has been suggested that productivity measures should be related to quality and health outcomes (Menachemi, Yeager, Welty, & Manzella, 2015; Romley et al., 2015; Sandy, Haltson, Metfessel, & Reese, 2015; Stecker & Schroeder, 2013). However, this may be challenging at the institutional level, where multiple treatment procedures and patient groups are pooled, and past work has found that the link between hospital efficiency and quality varies from a positive association to more mixed results (Heijink et al., 2015; Hussey, Wertheimer, & Mehrotra, 2013; Kittelsen et al., 2015; Menachemi et al., 2015; Romley et al., 2015; Romley, Jena, O'Leary, & Goldman, 2013; Stukel et al., 2012; Yasaitis, Fisher, Skinner, & Chandra, 2009).

Complex scientific results from DEA or SFA, based on proxies, are not everyday statistics known to health personnel and therefore may have limited impact at the bedside. Hypothetically, measures describing the number of patients to whom the personnel provide service may spark action among “the white coats” in everyday practice and have a supplemental value, despite not having the scientific basis as more advanced techniques. A report from the National Health Service (NHS) Institute revealed that patient admissions and completed consulting episodes per consultant varied by over 100% across different NHS trusts in England (Aragon, Castelli, & Gaughan, 2015; Castelli et al., 2015; Street & Castelli, 2014). If such differences are real, there would be a substantial gain if the lower-level performers could operate at the average level.

A simple description of productivity is the relation between input and output. The input of health personnel resources may be established through measures of the workforce or salary, whereas the assessment of output is more complex. Metrics such as the number of hospital admissions, daycare treatments, and outpatient consultations are not sufficient alone, but, as a group, they may cover differing pieces of a complex puzzle. However, the large degree of variations between different patient treatments and care levels are not covered. To compensate for this, researchers have used measures thought to reflect some of this variation, such as diagnosis-related groups (DRG), health care resource groups, or relative value units (E. Biorn, Hagen, T. P., Iversen, T., Magnussen, J., 2010; Castelli et al., 2015; Kentros & Barbato, 2013).

The extent of physician services available for patient treatment is the crucial issue, and the utilization of physician resources is therefore important. This, in turn, may depend on organizational perspectives as well as personnel mix (Bank & Gage, 2015; Greene, 2015; Johnson, Shah, Rechner, & King, 2008; Newhouse & Sinaiko, 2007; Rodysill, 2003; Sandbaek, Helgheim, Larsen, & Fasting, 2014; Sunshine, Hughes, Meghea, & Bhargavan, 2010). We undertook this study to examine physician productivity using panel analysis with limited Information maximum likelihood (LIML) regression and DEA analysis based on metrics of patients treated combined with health personnel indicators.

Background

In 2002, all public Norwegian hospitals were transferred from a system of county ownership to central government ownership (Hagen & Kaarboe, 2006). The aim was to increase hospital efficiency by providing greater autonomy with respect to planning, budgeting, and workforce policies. The reform aimed to define hospitals' economic responsibilities more precisely and to implement remuneration for personnel that would stimulate productivity, especially among physicians (E. Biorn, Hagen, T. P., Iversen, T., Magnussen, J., 2010; Magnussen, 2009; Verzulli, 2011). Hospitals were restructured as health enterprises comprising 1–8 of the previous hospitals and organized into five regional health authorities, which were reduced to four in 2005. During our study period (2001–2013), Norwegian hospitals consisted of five regional university hospitals (the most specialized hospitals, two of which were merged in 2010), 11 central hospitals (two with university functions), and four local hospitals. Norwegian health care is mainly funded by general taxation, and hospital care is paid through a mixture of global funding and activity-based funding (ABF), which is based mainly on the DRG system. Hospitals receive targeted compensation for teaching and research.

Aims and objectives

The current study had three aims. First, we investigated whether the utilization of the physician workforce, as assessed by indicators of patient treatment volumes in relation to the number of physicians, has improved since the 2002 hospital reform. Because we did not study the period before the reform was implemented, we had no ambition to examine causality. Second, using panel analysis with LIML estimations (Anderson, 1949) and the non-parametric DEA method for estimating a variable returns to scale cost function (Banker, Charnes, & Cooper, 1984; Charnes, Cooper, & Rhodes, 1978), we analyzed the relationship between the relative personnel mix (nurses, auxiliary nurses, and medical secretaries) and physician productivity. Third, we examined whether the new remuneration structure implemented with the reform translates into physician efficiency (Bloor et al., 2004; Devlin & Sarma, 2008). In our analyses, we used parameters reflecting patient treatment, research activity, and teaching and related these measures to workforce resources.

Methods

Data sources

The dataset covered the period from 2001, the last year before the reform was implemented, to 2013. All hospital enterprises in Norway (N = 19) were included, and we had data from each hospital each year. Hospital mergers during this period were handled by aggregating the data in the premerger period to the hospital structure in the post-merger period.

Data on workforce resources and salaries were obtained from The Employers Organization Specter and Statistics Norway and are described in Table 1. Salary data consisted of payment for regular work, casual overtime, and on-call services. Activity data were obtained from the Norwegian Patient Register and consisted of the total number of treatments, including hospitalization, daycare, and outpatient consultations and the corresponding DRG scores. The DRG system groups patients into categories with similar use of resources and reflects the total costs for patient treatment episodes. The DRG unit price is an estimated average cost of all patients at the national level and is constructed for calculation of ABF reimbursement. In addition to personnel costs, DRG scores include overhead costs, medications, blood, implants and so on, and are therefore not an exact measure of patient-related workload in relation to personnel productivity.

Table 1. Variables included in the final regressions

LIML=Limited information maximum information regression. DEA: Data envelopment analysis. DRG: Diagnosis-related groups. FTE: Full time equivalents of personnel

	Variables	Definition	Data Source
Target Dependent in LIML	Total DRG/Physician	Sum of DRG scores from hospitalization, day treatment and outpatient consultations per FTE Physician	NPR for activity, Physician FTE from Specter
Regressors in LIML Analysis	Physician Salary Physician Salary Lagged FTE Nurses/Physician FTE Secretaries/Physician Other/Physician Resident Fraction Research/Physician Scale Scale Squared	Average total salary per physician Average total salary per physician the year before Sum of FTE of Nurses per Physician Sum of FTE of Secretaries per Physician Sum of FTE of Other staff per Physician Sum of FTE Resident per Total FTE Physicians Total Research Points per Physician Number of Beds Number of Beds Squared	Specter and Statistics Norway
Input variables in DEA analysis	Labor inputs: FTE for each personnel group: Physicians, Nurses, Auxiliary nurses, Secretaries and Other staff	FTE estimates based on hours worked including overtime	Specter
	Input prices: Wage cost per FTE in each personnel group	Sum of wage costs including pension and social costs in each group for all FTE, divided by the FTE estimated above.	Statistics Norway
	Non-labor inputs	Total operating costs excluding capital costs minus total wage costs (input price normalized to 1).	Statistics Norway

DRG scores do not reflect research or education, and hospitals use a substantial amount of their resources for teaching and research. Such activities may influence the workload of the personnel. Hospital residents need considerable coaching and training, which may influence the productivity of the physician staff (Farnan, Johnson, Meltzer, Humphrey, &

Arora, 2008; Johnson et al., 2008; W. M. McDonnell, P. Carpenter, K. Jacobsen, & H. A. Kadish, 2015; Medin et al., 2011). We used the balance between residents and senior consultants (resident fraction) to examine this factor. As a proxy for research, for each year examined, we calculated the publication score (Bonastre, le Vaillant, & de Pouvourville, 2011; Linna, Hakkinen, & Linnakko, 1998; Medin et al., 2011), which is a bibliometric measure taking into account the impact-weighted number of journal articles and the number of doctoral theses completed each year. Such data were only available for 2003–2013 and were extrapolated using linear regression for 2001 and 2002 for use in the multivariate analyses.

We included the number of hospital beds (both as a linear and as a quadratic term) to account for scale effects (Aragon et al., 2015). These data were obtained from Statistics Norway. Because the hospitals differed in their scope of emergency capacity, we included fixed effects for each hospital enterprise.

Analytical approach and statistics

We used Farrell's efficiency concepts (Farrell, 1957) to define productivity as

Productivity = Output/Input,

where technical productivity is measured as the total DRG score per full-time equivalent (FTE) physician. The variables showing significance in Pearson correlations were included in the multivariate analyses, and the final regressors are listed in Table 1.

The relationship between productivity and salary raises questions of cause and effect, as increased salary may stimulate improved productivity, and improved productivity may be rewarded by increased salary. Accordingly, salary may be an endogenous variable with respect to productivity, whereas personnel mix, research, and education are not. In our final analyses, we used the LIML procedure to account for the simultaneous structure of the salary–productivity relationship with the following equations:

Productivity = a0 + b1*lagSalary + b2 * Other variables

Salary = c0 + d1*lagProductivity + d2 * Other variables

We constructed three versions of the model. The first model is a time series cross-section model that utilized all available information in the dataset (Model 1). The two other models use fixed effects for hospital, utilizing variation within each hospital over time. In Model 2, we assume that there is a 1-year lag in the effects of salary on productivity and of productivity on salary, whereas salary from the same year is used in Model 3.

To further study physician productivity in relation to the balance of resources and personnel inputs, we used the non-parametric DEA method to estimate a variable returns to scale cost function (Banker et al., 1984; Charnes et al., 1978; Farrell, 1957). We did not intend to study total factor productivity, but rather to focus on the optimal mix of various personnel groups as revealed by the cost function estimates. We included non-labor costs as described in Table 1.

The cost function is defined as the minimum cost necessary to produce a given level of output (e.g., health services) with exogenously given input prices (e.g., wages). Cost functions assume input substitution possibilities so that the use of an input increases if the wages of that group decrease. The DEA method is basically deterministic, and we used bootstrapping methods to calculate the sampling error of the estimates and assess the variance and confidence intervals (Simar & Wilson, 1998, 2000). Bootstrapping is a procedure that draws with replacement from the primary data sample, mimicking the data-generating process of the underlying true model and producing multiple pseudo-estimates that allow for the calculation of the sampling error of the estimates and estimate variance, as well as confidence intervals. The assumption is that we know how the data are generated, and we are therefore able to calculate how well our estimates reflect the true costs and efficiency levels, conditional on our data and method. The bootstrapped results are therefore robust with respect to sampling error, but the bootstrapping procedure does not account for measurement error.

Cost efficiency was decomposed into technical and allocative efficiency (Farrell, 1957). High technical efficiency implies that there is no excess input of resources to obtain a certain production level, whereas high allocative efficiency indicates that the mix of input resources is optimized. Allocative efficiency reflects the extent to which the input mix is optimal by comparing the differing marginal costs when the inputs are varied, based on the ratio of prices of the inputs.

We used SAS software version 14 for the panel analysis, the Frisch Nonparametric DEA Program (Frisch Centre, Oslo, Norway), and SPSS (IBM, version 22) for the comparison of descriptive data using ANOVA.

Results

Descriptive data

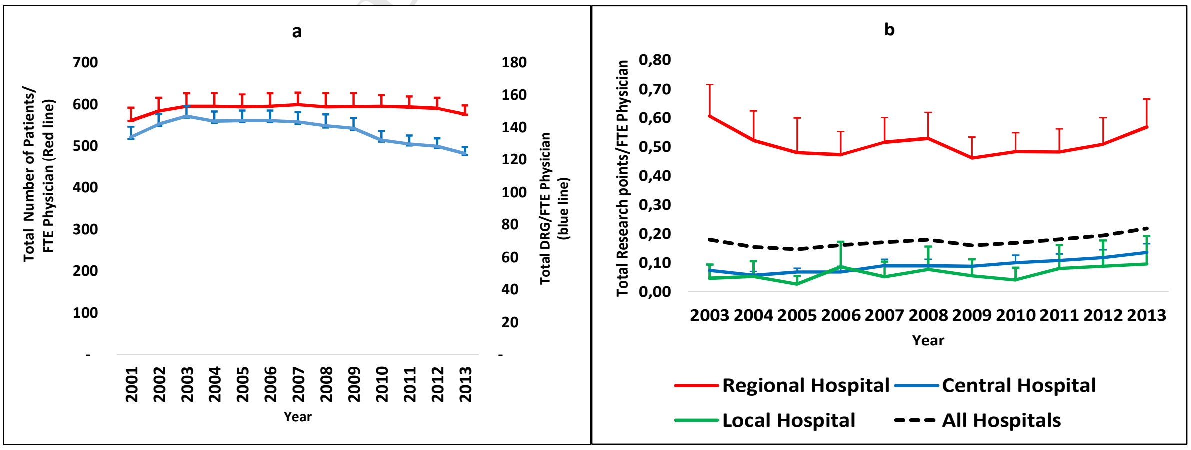
To avoid an extensive table with data from all years, we present descriptive data from 2001 and 2013 (Table 2) supplemented with graphs that illustrate developments over time in some basic variables (Figure 1).

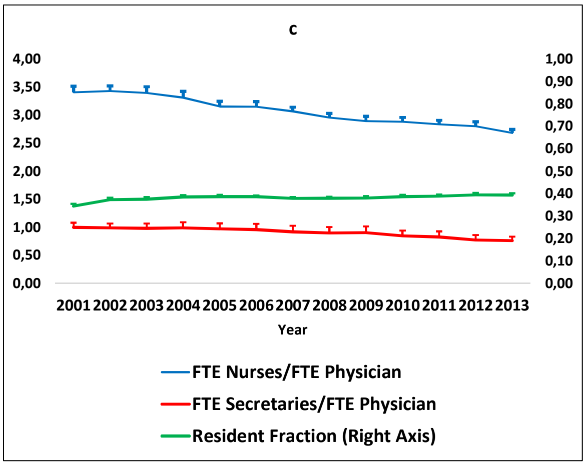
Table 2. Descriptive data on resources, activities, and productivity in 2001 and 2013

	Year 2001			Year 2013				
	Sum National Level	Mean of Hospitals	Standard Deviation	Sum National Level	Mean of Hospitals	Standard Deviation	National Change	
Hospital Output								
Hospitalstays	685 901	36 100	21 300	739 191	38 905	21 098	7.8 %	
Daycare treatments	309 112	16 269	11 478	432 376	22 757	13 130	39.9 %	
Outpatient consultations	2 859 315	150 490	98 115	4 510 978	237 420	165 063	57.8 %	
Total number of patient contacts	3 854 328	202 859	130 210	5 682 545	299 081	197 509	47.4 %	
DRG Hospitalstays	698 368	36 756	26 224	951 804	50 095	35 999	36.3 %	
DRG Daycare	98 290	5 173	3 586	92 879	4 888	3 181	-5.5 %	
DRG Outpatient consultations	110 423	5 812	4 595	192 745	10 144	7 299	74.6 %	
Research points (2003 and 2013)	2 596	122	315	3 665	193	396	41.2 %	
Physician variables								
FTE physicians	6 784	357	322	9 852	519	431	45.2 %	
Physician Salary (NOK, Deflated)		671 612	47 078		890 387	45 321	32.6 %	
Productivity	National average	Low/High	Standard Deviation	National average	Low/High	Standard Deviation	National Change	P-val
Hospitalstays / Physician	101.1	62.1 / 158.0	25.2	75.0	44.2 / 105.9	15.7	-26 %	<0.001
Daycare / Physician	45.6	31.6 / 87.2	14.1	43.9	24.6 / 67.8	11.2	-3.7 %	ns
Outpatient consultations / Physician	421.5	280.0 / 703.8	108.8	457.9	341.6 / 602.1	75.5	8.6 %	ns
Total Number of Patient Contacts /	561.1	408.5 / 831.0	112.1	576.8	404.2 / 754.1	96.3	1.4 %	ns
DRG Hospitalstays / Physician	102.9	81.8 / 175.9	21.9	96.6	72.2. / 125.2	16.8	-13 %	<0.05
DRG daycare / Physician	14.5	10.7 / 22.7	3.6	9.4	6.0 / 13.6	2.1	-34 %	<0.001
DRG outpatient consultations / Physician	16.3	11.4 / 20.8	2.4	19.6	13.5 / 26.7	3.4	20 %	<0.001
Total DRG / Physician	134	103.9 / 213.4	24.7	125	91.7 / 163.0	20.2	-6 %	<0.05
Research / Physician (2003 and 2013)	0.36	0.0 / 0.88	0.25	0.37	0.01 / 0.84	0.22	3 %	ns

Figure 1. The development of selected parameters 2001–2013

a. Total patients treated and total diagnosis-related groups (DRG) score per full time equivalent (FTE) physician. b. Research per FTE physician in regional, central, and local hospitals. c. Nurses and secretaries per FTE physician and resident fraction.





The 47% increase in the total number of patients treated was a consequence of a shift from hospitalized treatment to daycare/outpatient treatment. However, the increase varied from 12% to 92% at individual hospitals. All 19 hospitals reduced the number of hospital beds, and six reduced their volume of hospitalizations. Daycare and outpatient treatment increased in all of the studied hospitals, with a magnitude varying from 15% to 92%. For hospitalized patients, the DRG increased more than the number of patients, whereas the opposite was observed for daycare patients. This may reflect a shift of low-intensity treatment from hospitalization to daycare, leaving only the more complex cases in the hospitalized activity.

The total research scores increased by 41% at the national level, but this differed considerably across individual hospitals (Figure 1b). Regional and university hospitals accounted for 88% of the research activity.

Physician productivity

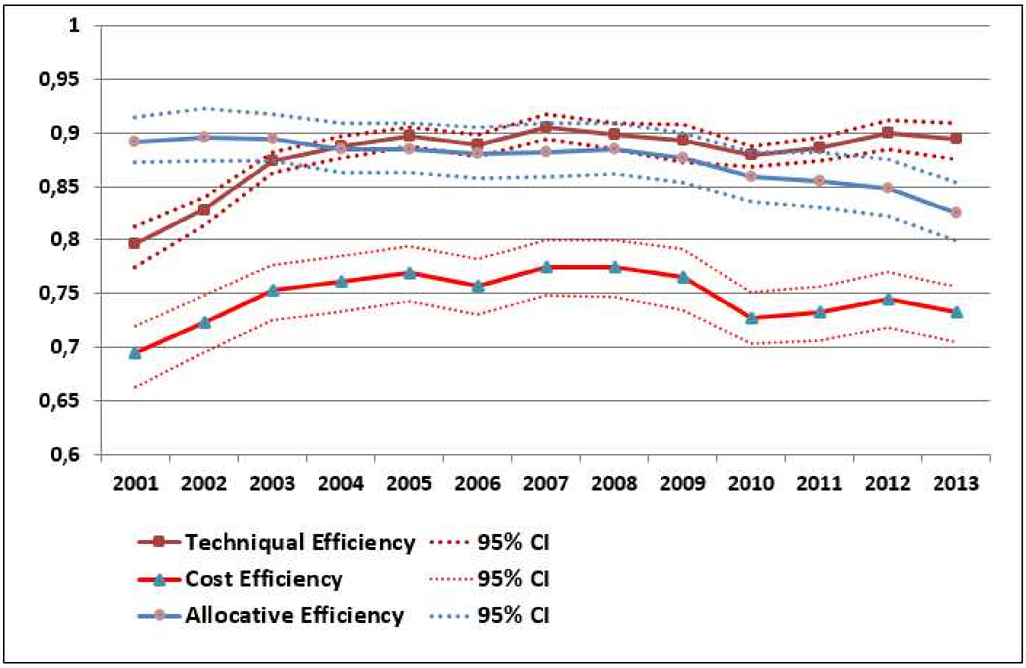
Table 2 shows that the total DRG score per FTE physician decreased by 6% ($p < 0.05$) from 2001 to 2013, whereas the total number of patients treated per physician increased by 1.4% (26 patients per physician per year, $p = 0.40$). The difference between the hospitals with the highest and lowest DRG per physician decreased from 104% (213 vs. 104) in 2001 to 77%

(163 vs. 92) in 2013, but this convergence was mainly caused by a reduction in the high scores and not by an overall increase.

The average research score per FTE physician increased from 0.08 to 0.15 for the central hospitals (p < 0.01) but was unchanged for regional and local hospitals (Figure 1b). The DEA showed that cost efficiency varied across the study years, but there was no significant upwards or downwards trend. Decomposition revealed that technical efficiency increased during the first four years but levelled off beginning in 2005. A possible interpretation for this finding is that the use of resources was excessive in relation to the patient treatment generated. Allocative efficiency, in contrast, decreased significantly throughout the study period (Figure 2). This indicates that the balance between multiple input resources deteriorated over the study period. In 2013, technical efficiency was 0.89, and allocative efficiency was 0.83.

Figure 2. Cost efficiency, technical efficiency, and allocative efficiency 2001-2013

Bootstrapped averages by year with 95% confidence intervals (CI)



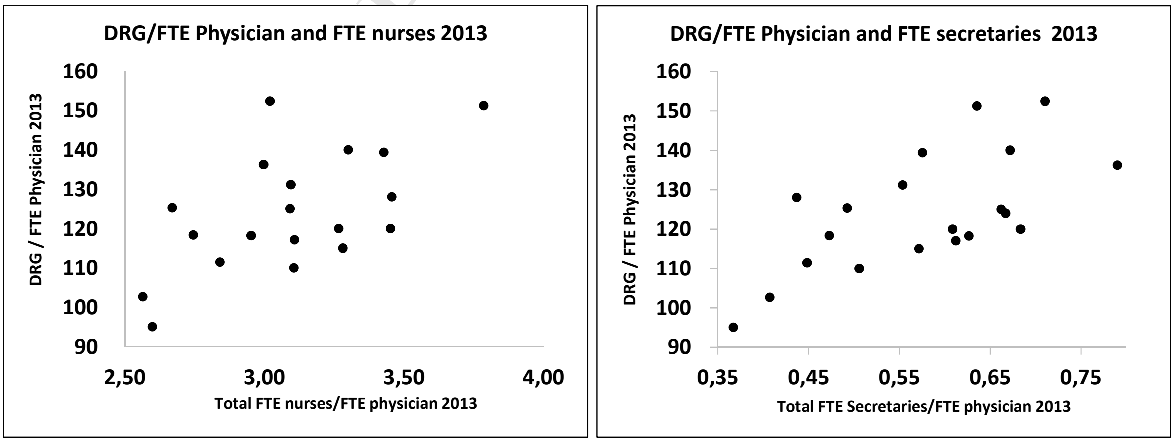
Variables influencing physician productivity

The results from the LIML regression models are presented in Table 3. The numbers of nurses and secretaries per FTE physician were the strongest correlates of productivity in all analyses, both across and within the hospitals. Figure 3 shows a simple illustration of these relations in 2013; the observations for the other years were similar.

Table 3. Results from panel analyses.

		Limited-Information Maximum Likelihood Estimation											
		Model 1: Without fixed effect				Model 2: Fixed effect for hospital, salary lagged				Model 3: Fixed effect for hospital, salary same year			
Parameter	DF	Estimate	Error	t-value	Pr > |t|	Estimate	Error	t-value	Pr > |t|	Estimate	Error	t-value	Pr > |t|
Intercept	1	51.55	42.78	1.21	ns	-70.82	34.00	-2.08	<0.05	-95.26	29.10	-3.27	<0.01
Physician Salary	1	-0.0004	0.0001	-1.09	ns					0.00004	0.00002	1.95	ns
Physician Salary Lagged	1	0.0002	0.0003	0.72	ns	0.00006	0.00002	3.02	<0.01				
FTE Nurses/Physician	1	13.51	3.81	3.54	<.0005	18.02	3.39	5.31	<0.0001	11.51	3.20	3.60	<0.001
FTE Secretaries/Physician	1	40.51	10.28	3.94	<.0001	47.63	8.89	5.35	<0.0001	41.57	9.40	4.43	<0.0001
FTE Other/Physician	1	14.57	4.69	3.11	<0.01	10.71	3.46	3.10	<0.01	15.63	3.70	4.22	<0.0001
Resident Fraction	1	-31.63	40.19	-0.79	ns	89.03	42.08	2.12	<0.05	175.31	39.19	4.47	<0.0001
Research/Physician	1	-38.48	11.51	-3.34	<0.01	6.85	16.48	0.42	ns	17.56	17.86	0.98	ns
Scale	1	0.04	0.01	3.01	<0.01	0.009	0.03	0.32	ns	0.03	0.03	1.24	ns
Scale Squared	1	-0.0001	0.0000	-2.95	<0.01	0.00005	0	-0.68	ns	-0.00001	0.000007	3.83	ns

Figure 3. The relation between diagnosis-related group (DRG) scores per fulltime equivalent (FTE) physicians and nurses and secretaries per physician in 2013



The number of other types of personnel per FTE physician also correlated significantly with productivity. This might be an effect of hospital size, but including this variable as scale and scale squared showed that the scale factor of hospital size converged, with a statistical optimum of approximately 350 beds. A negative effect of a higher resident fraction in the univariate analysis was eliminated in Model 1, and residents were shown to have a positive effect on productivity in Models 2 and 3. Also, a negative correlation between the fraction of outpatient consultations and physician productivity in the univariate analysis ($r = -0.34$, $p < 0.01$) was eliminated in the panel analysis.

The DEA confirmed the association between physician productivity and personnel mix. However, although declining allocative efficiency indicates that cost savings could be achieved by changing the input mix, this finding does not reveal which inputs are over- or under-utilized. Still, the bootstrapped estimates in Table 4 show that, when comparing the hospitals' actual 2013 cost shares to the "optimal model" based on the bootstrap, the use of some inputs was not optimal. Allocative efficiency in 2013 would be improved if the use of physicians were reduced so that their cost share was lowered from 17.1% to 14.6%, whereas the cost share of auxiliary nurses, for example, should be increased by 2.1%.

Table 4. Bootstrap estimates of optimal cost shares compared with actual observed shares for different resources in 2001 and 2013

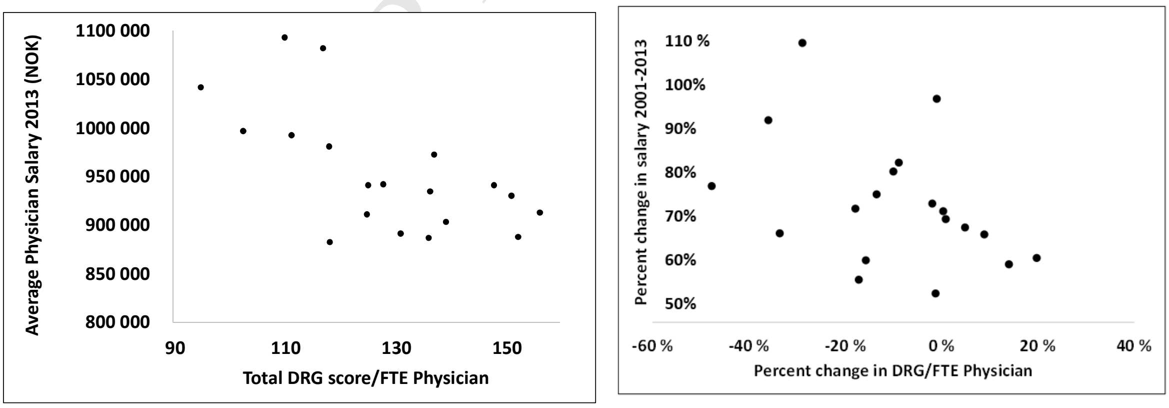
		Quantity	Price in 2013 NOK	Cost shares			
		Full time equivalents	(Personnel costs)	Actual	Optimal	Difference	(p-val)
2001	Physicians	7 108	1 266	14.3 %	15.0 %(14.7%-15.5%)	0,7 %	<0.001
	Nurses	22 032	687	24.1 %	24.2% (23.5%-25.2%)	0,0 %	ns
	Auxiliary Nurses	4 873	611	4.7 %	5.6% (5.1%-6.3%)	0,8 %	<0.001
	Secretaries	6 196	509	5.0 %	4.9% (4.8%-5.9%)	-0,2 %	<0.05
	Nonmedical staff	23 472	608	22.7 %	18.3% (17.4%-18.8%)	-4,4 %	<0.001
	Non-labor inputs	18 240	1 000	29.1 %	32.0% (30.0%-32.9%)	3,0 %	<0.001
2013	Physicians	9 852	1 330	17.1 %	14.6% (14.4%-15.2%)	-2,5 %	<0.001
	Nurses	25 695	729	24.4 %	24.3% (23.6%-25.0%)	-0,1 %	ns
	Auxiliary nurses	3 293	631	2.7 %	5.1% (4.5%-5.8%)	2,4 %	<0.001
	Secretaries	5 242	535	3.7 %	4.8% (4.8%-5.7%)	1,1 %	<0.001
	Nonmedical staff	21 653	672	18.9 %	18.1% (17.5%-19.2%)	-0,9 %	<0.05
	Non-labor inputs	25 535	1 000	33.2 %	33.2% (30.9%-34.0%)	0,0 %	ns

Physician remuneration

Physician salary correlated negatively with productivity in the univariate analysis in all years, and the hospitals with the largest increase in salary reduced their productivity. A simple illustration of this aspect is shown in Figure 4 a and b. However, the hypothesis that hospitals with higher physician salaries are characterized by lower physician productivity than are those with lower salaries did not reach significance in our Model 1, probably indicating the multifactorial aspect of this relation. The salary from the previous year correlated positively and significantly in the fixed effects model (within hospital analysis), but this was voided by the reciprocal effect of productivity on salary in the LIML analysis. We experimented with different combinations of regular salary, overtime and on-call salary, but without any significant results.

Figure 4. Univariate relation between physician salary and total diagnosis-related (DRG) scores per physician in 2013 (a) and the relation between percent change in salary and productivity 2001-2013 (b).

a b



Discussion

The results of the current study show that, although there was a significant increase in treatment activity in Norwegian hospitals from 2001 to 2013, this increase occurred primarily because of the use of more physicians and not because of an improvement in physician productivity. Furthermore, differences across Norwegian hospitals of 80% in the total number of patients treated and 64% in the DRG scores per FTE physician present a challenge with respect to overall productivity and should trigger more research. Our findings correspond well to other reports that have revealed that patient admissions and completed consulting episodes per consultant varied by over 100% between NHS trusts in England (Street & Castelli, 2014).

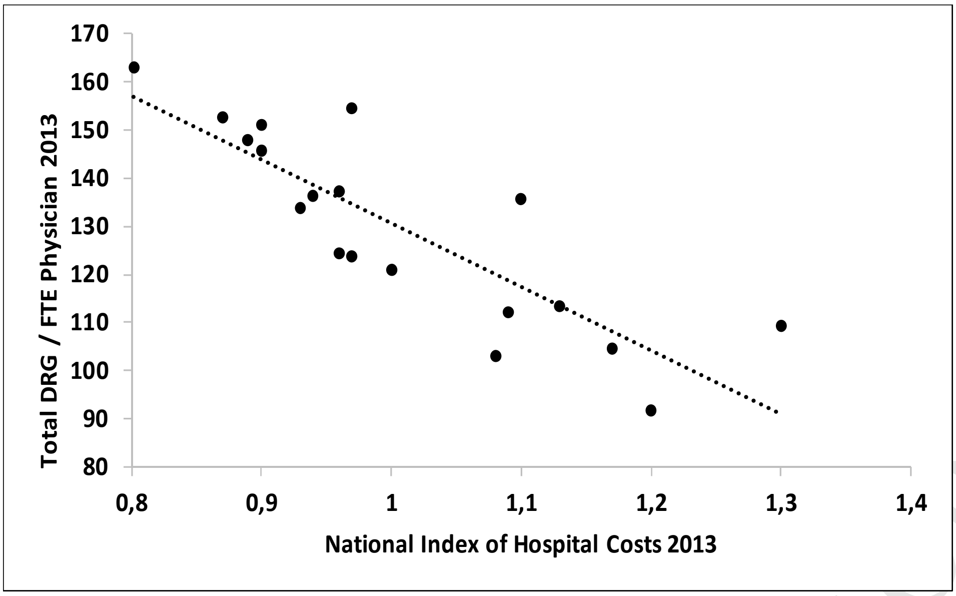
Our most striking result is the effect of personnel mix on physician productivity. The LIML analysis revealed that staffing of both nurses and secretaries correlated significantly with physician productivity, both across and within the studied hospitals. Furthermore, the DEA indicated that, with the current mix of resources, nurse staffing is close to the optimal model, but there is an overuse of physicians of approximately 15% and deficiencies in the use of auxiliary nurses and secretaries of about 89% and 30%, respectively. We interpret this finding as evidence that developments in the study period have resulted in a suboptimal personnel mix.

The substantial change from hospitalization to outpatient treatment makes it difficult to fully assess the development of a complex issue such as physician productivity. This shift in the care level is a factor that may affect our estimates of both the patient mix and the personnel mix, and it is well known that the lower weight assigned to outpatient activities by the DRG system may underestimate real efficiency (Vitikainen, Linna, & Street, 2010). However, we find it unlikely that a 26% reduction in the number of hospitalized treatments and a 3.7% reduction in day treatments per physician may be compensated by an 8.6% increase in outpatient consultations per physician. It is of note that physician productivity increased in several hospitals while it worsened in others. This large variation in the utilization of physician resources among the hospitals parallels similar differences presented in previous reports analyzing efficiency at the institutional level in Norwegian hospitals over the same period (E. Biorn, Hagen, Iversen, & Magnussen, 2003; E. Biorn, Hagen, T. P., Iversen, T., Magnussen, J., 2010). We conclude that the intention to improve personnel productivity has not yet resulted in the homogenous performance of hospitals with respect to the utilization of the physician workforce. This is also consistent with previous reports from other health care systems (Castelli et al., 2015; Hvenegaard, Street, Sorensen, & Gyrd-Hansen, 2009; Ineveld, Oostrum, Vermeulen, Steenhoek, & Klundert, 2015; Milstein & Kocher, 2014; Street & Castelli, 2014). In fact, in their study of Dutch hospitals, Ineveld et al. (Ineveld et al., 2015) found that the difference between hospitals increased over time.

Although several previous studies have reported that the overall efficiency of Norwegian hospitals improved during our study period, most of this work analyzed data only through 2004. We identified a corresponding improvement in cost efficiency until 2005, but we found no further improvement thereafter. We also found a steady reduction in physician productivity throughout the total period. If efficiency gains are mainly obtained through administrative procedures and reduced staffing in non-medical personnel categories, this may not be a sustainable strategy in the long run (Tiemann & Schreyogg, 2012).

Although DRG score per FTE physician is a rather coarse measurement, it seems to be fairly well related to the overall costs in Norwegian hospitals (Helsedirektoratet, 2013). This is illustrated in Figure 5 for 2013, and similar results were found for all of the years studied. This is an additional indication that physician productivity and, possibly, the corresponding measures for other personnel groups are important in the long-term development of hospital efficiency.

Figure 5. Univariate relation between index of national hospital and total DRG FTE physician.



Several studies have documented the effect of personnel supporting physicians on productivity (Bank & Gage, 2015; Grimshaw, 2012; McDonnell et al., 2015 ; Rumbold, 2015). The change in personnel balance observed in the present study may be caused by several factors. The shift in care level may have changed the balance between physicians and other health personnel, as daycare and outpatient treatment may require more physicians and less nursing personnel compared with hospitalized treatment. Some of the observed reduction in medical secretary resources may be related to the expected effects of technological solutions that are assumed to reduce secretary work (e.g., voice recognition and electronic patient charts). However, past work has reported that a significant increase in non-medical tasks for physicians casts some uncertainty on the effects of such technological strategies (J. Rosta & Aasland, 2014; J. A. Rosta, OG., 2015). Furthermore, the increasing specialization among physicians may not be reflected to the same extent among nurses.

The effect of resident training on productivity has been extensively studied (Farnan et al., 2008; Harvey, Al Shaar, Cave, Wallace, & Brydon, 2008; Johnson et al., 2008; Kawano, Nishiyama, & Hayashi, 2014; Zeidel et al., 2005). We found a positive correlation between productivity and the fraction of the total FTE of physicians consisting of residents. This may reflect that residents in Norwegian hospitals spend a considerable portion of their training time conducting patient treatment.

There have been several reports on the effects of incentives for physicians regarding productivity (Andreae & Freed, 2003; Conrad et al., 2002; Wilson et al., 2006). Andreae et al. (2003) found that targeted incentives and remuneration based on relative value units caused a 20% increase in clinical productivity. There was a declared ambition in the Norwegian hospital reform to improve personnel productivity by giving the hospitals freedom to implement more targeted remunerations. We cannot make any conclusions of causality based on the measures we used, but we found little evidence that the new remuneration structure for physicians has translated into improved productivity. One hypothesis may be that collective bargains still prevail despite local negotiations. Targeted incentives should definitely receive more focus in applied settings, seeking remuneration models that are related to the actual health care system as well as social and contextual factors (Wranik & Durier-Copp, 2011).

At what level do health care reforms work?

The crucial question facing health care is whether there will be enough personnel resources to meet future needs, and this question relates especially to physicians. Increasing medical specialization will call for more specialized physicians, who may restrict their medical scope for patient treatment to their own specialties. This, in turn, may increase the need for resources.

Policy makers intend to improve the efficiency of health care systems through their reforms, and an interesting question is whether we should expect an effect of political hospital reforms at the bedside (Davis & Rayburn, 2016). It is possible that a major part of the effect of political reforms is based on improvements in administrative and organizational perspectives. However, even if reforms may have effects on efficiency at the macro level, we need political initiatives that also create changes at the micro level, because improvements may not be sustained if they do not include an enhancement of the efficiency of the health care workforce (Burwell, 2015; Conrad et al., 2002; Franco, Bennett, & Kanfer, 2002; Lieberman & Allen, 2015; Marshall & Bindman, 2016; McWilliams, Chernew, Landon, & Schwartz, 2015; Milstein & Kocher, 2014; Ryskina & Bishop, 2013).

Two other political initiatives of relevance to our study were implemented in the period. In 1997, ABF was introduced with the intention to improve efficiency. The level of ABF has varied between 30-60 %. In 2001, all patients were granted free choice of hospitals combined with the removal of county border barriers. This primarily aimed to reduce long waiting times, and previous research has shown that both the introduction of ABF and the expansion of hospital budgets have been factors in reducing waiting time for elective patients (Hagen & Kaarboe, 2006; Ringard & Hagen, 2011). Although these two initiatives may also have influenced the hospitals' operational performance, we conclude that physician productivity did not improve during our study period, irrespective of these reforms.

Norway, like several other modern health care systems, will face a significant deficit of health personnel in the future (Roksvaag & Texmon, 2012). For this reason, we believe that there must be a considerably stronger focus on improving workforce productivity at the clinical level. Our data strongly indicate that staffing and personnel mix significantly influences the utilization of health personnel. Accordingly, any reform or change should also stimulate the core personnel, and managerial and organizational efforts, leadership, and economical incentives ought to focus on such goals.

The large differences in physician productivity observed across the hospitals in our study may indicate a considerable potential for improvement. Optimizing hospital staff is essential for improving efficiency, because personnel costs constitute more than 60% of the total expenses. Several factors, such as leadership, the improvement of occupational health, and the reduction of temporary staff and overtime, may contribute to this optimization. In an interview-based study of managers and clinicians in orthopedics and cardiology in acute hospitals, Bloom et al. (Bloom, Propper, Seiler, & Van Reenen, 2015) concluded that management quality was favorably correlated with indicators of hospital performance with respect to waiting times, mortality, financial performance, and staff satisfaction. Burns and Muller (Burns & Muller, 2008) also focused on such factors in their review of the literature on hospital/physician collaboration. They found that the characteristic distinguishing between high- and low-performing hospitals was “the level of both hospital executive and physician behavioral skills,” including physicians’ trust in hospital executives, mutual respect and support, communication, physicians’ involvement in clinically related decision making, and hospital executive leadership over time. This finding likely supports the idea that future reforms should promote a united process with professional medical development and system reforms.

Limitations

DRG scores are the official measurement of treatment activity for annual governmental reports of hospital productivity in Norwegian hospitals, but they have limitations as measures of productivity in our context. None of the parameters covers the activity in a complete manner individually, and the extent to which their combination may compensate for this limitation is unclear. As our data show, despite an increase in the number of patients treated per FTE physician, the DRG scores did not increase to the same extent. The DRG system has been adjusted over the study period because of economic considerations and new DRG weights have appeared due to new treatment methods and DRGs for day treatment where there was only hospitalized treatment in 2001. Although these measures may be adequate within each year, comparisons over time may be distorted. An appropriate analysis of this shift should be done for each DRG and is complex. However, when experimenting with same DRG weights at the macro level for 2001 and 2013, this did not change our main conclusion.

An increasing population of chronically ill patients may cause a shift towards more control and follow-up activities which may require full-scale personnel resources without triggering full DRG reimbursement as new patients do. In addition, the differing combinations of medical activities among the hospitals may cause unequal scores on the variables we have used. We cannot rule out the possibility that assessing more specific characteristics of hospitals could have yielded different results.

Conclusions

Despite several political reforms of the Norwegian hospital sector over the study period, physician productivity as assessed by our measures declined, and we found significant variation in productivity among Norwegian hospitals. These findings must be addressed further by future work if the coming challenges are to be solved. It is obvious that the balance between support staff and the physician workforce may have a significant effect on the utilization of physicians, and the current situation in our data indicates that future planning regarding support staff should have a factual and rational basis. Because there is a great deal of variety in the individual competence and performance of health personnel from clinical, educational, and scientific perspectives, we believe that more individual incentives and fewer collective solutions should be considered in future remuneration negotiations.

References

Anderson, T. W., & Rubin, H. (1949). Estimation of the parameters of a single equation in a complete system of stochastic equations. *The Annals of Mathematical Statistics, 20*(1), 46-63.

Andreae, M. C., & Freed, G. L. (2003). The rationale for productivity-based physician compensation at academic health centers. *J Pediatr, 143*(6), 695-696. doi:10.1016/j.jpeds.2003.09.009

Aragon Aragon, MJ., Castelli, A., Gaughan, J. (2015). *Hospital trusts productivity in the English NHS: uncovering possible drivers of productivity variations*. University of York, Centre for Health Economics.

Askildsen, J. E. H. J. (2006). Wages and work conditions as determinants for physicians' work decisions. *WORKING PAPERS NO 06/06 IN ECONOMICS. UNIVERSITY OF BERGEN*. Norway.

Bank, A. J., & Gage, R. M. (2015). Annual impact of scribes on physician productivity and revenue in a cardiology clinic. *Clinicoecon Outcomes Res, 7*, 489-495.

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management science, 30*(9), 1078-1092.

Biorn, E., Hagen, T. P., Iversen, T., & Magnussen, J. (2003). The effect of activity-based financing on hospital efficiency: a panel data analysis of DEA efficiency scores 1992-2000. *Health Care Manag Sci, 6*(4), 271-283.

Biorn, E., Hagen, T. P., Iversen, T., Magnussen, J. (2010). How different are hospitals' responses to a financial reform? The impact on efficiency of activity-based financing. *Health Care Manag Sci, 13*(1), 1-16.

Bloom, N., Propper, C., Seiler, S., & Van Reenen, J. (2015). The impact of competition on management quality: evidence from public hospitals. *The Review of Economic Studies*, 1-33. doi:doi: 10.1093/restud/rdu045

Bloor, K., Maynard, A., & Freemantle, N. (2004). Variation in activity rates of consultant surgeons and the influence of reward structures in the English NHS. *J Health Serv Res Policy, 9*(2), 76-84.

Bonastre, J., le Vaillant, M., & de Pouvourville, G. (2011). The impact of research on hospital costs of care: an empirical study. *Health Econ, 20*(1), 73-84.

Burns, L. R., & Muller, R. W. (2008). Hospital-Physician Collaboration: Landscape of Economic Integration and Impact on Clinical Integration. *Milbank Quarterly, 86*(3), 375-434.

Burwell, S. M. (2015). Setting Value-Based Payment Goals — HHS Efforts to Improve U.S. Health Care. *New England Journal of Medicine, 372*(10), 897-899.

Busse, R., Schreyogg, J., & Smith, P. C. (2008). Variability in healthcare treatment costs amongst nine EU countries - results from the HealthBASKET project. *Health Econ, 17*(1 Suppl), S1-8.

Castelli, A., Street, A., Verzulli, R., & Ward, P. (2015). Examining variations in hospital productivity in the English NHS. Eur J Health Econ 16.3, 243-254.doi:10.1007/s10198-014-0569-5

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European journal of operational research, 2*(6), 429-444.

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

Cooper, R. A. (2004). Weighing the evidence for expanding physician supply. *Ann Intern Med, 141*(9), 705-714.

Davis, D. A., & Rayburn, W. F. (2016). Integrating Continuing Professional Development With Health System Reform: Building Pillars of Support. *Acad Med, 91*(1), 26-29.

R. A., & Sarma, S. (2008). Do physician remuneration schemes matter? The case of Canadian family physicians. *J Health Econ, 27*(5), 1168-1181.

Farnan, J. M., Johnson, J. K., Meltzer, D. O., Humphrey, H. J., & Arora, V. M. (2008). Resident uncertainty in clinical decision making and impact on patient care: a qualitative study. *Qual Saf Health Care, 17*(2), 122-126.

Farrell, M. (1957). The Measurement of Productive Efficiency. *J R Stat Soc Ser A (General)*(Vol 120, NO. 3), 253-290.

Franco, L. M., Bennett, S., & Kanfer, R. (2002). Health sector reform and public sector health worker motivation: a conceptual framework. *Social Science & Medicine, 54*(8), 1255-1266.

Greene, M. (2015). Better-performing groups use NPPs to extend physician productivity. *MGMA Connex, 15*(1), 32-33.

Grimshaw, H. (2012). Physician scribes improve productivity. Oak Street Medical allows doctors to spend more face time with patients, improve job satisfaction. *MGMA Connex, 12*(2), 27-28.

Hagen, T. P., & Kaarboe, O. M. (2006). The Norwegian hospital reform of 2002: central government takes over ownership of public hospitals. *Health Policy, 76*(3), 320-333.

Harvey, M., Al Shaar, M., Cave, G., Wallace, M., & Brydon, P. (2008). Correlation of physician seniority with increased emergency department efficiency during a resident doctors' strike. *N Z Med J, 121*(1272), 59-68.

Heijink, R., Engelfriet, P., Rehnberg, C., Kittelsen, S. A., Hakkinen, U., & Euro, H. s. g. (2015). A Window on Geographic Variation in Health Care: Insights from EuroHOPE. *Health Econ, 24 Suppl 2*, 164-177.

Helsedirektoratet. (2013). SAMDATA Spesialisthelsetjenesten 2013.

Hollingsworth, B. (2008). The measurement of efficiency and productivity of health care delivery. *Health Econ, 17*(10), 1107-1128.

Hussey, P. S., Wertheimer, S., & Mehrotra, A. (2013). The association between health care quality and cost: a systematic review. *Ann Intern Med, 158*(1), 27-34.

Hvenegaard, A., Street, A., Sorensen, T. H., & Gyrd-Hansen, D. (2009). Comparing hospital costs: what is gained by accounting for more than a case-mix index? *Soc Sci Med, 69*(4), 640-647.

Ineveld, M., Oostrum, J., Vermeulen, R., Steenhoek, A., & Klundert, J. (2015). Productivity and quality of Dutch hospitals during system reform. *Health Care Management Science*, 1-12.

Johnson, T., Shah, M., Rechner, J., & King, G. (2008). Evaluating the effect of resident involvement on physician productivity in an academic general internal medicine practice. *Acad Med, 83*(7), 670-674.

Kawano, T., Nishiyama, K., & Hayashi, H. (2014). Adding more junior residents may worsen emergency department crowding. *PLoS One, 9*(11), e110801. doi:10.1371/journal.pone.0110801

Kentros, C., & Barbato, C. (2013). Using normalized RVU reporting to evaluate physician productivity. *Healthc Financ Manage, 67*(8), 98-105.

Kittelsen, S. A., Anthun, K. S., Goude, F., Huitfeldt, I. M., Hakkinen, U., Kruse, M., . . . Ratto, H. (2015). Costs and Quality at the Hospital Level in the Nordic Countries. *Health Econ, 24 Suppl 2*, 140-163. doi:10.1002/hec.3260

Laudicella, M., Olsen, K. R., & Street, A. (2010). Examining cost variation across hospital departments--a two-stage multi-level approach using patient-level data. *Soc Sci Med, 71*(10), 1872-1881.

Lieberman, D., & Allen, J. (2015). New approaches to controlling health care costs: Bending the cost curve for colonoscopy. *JAMA Internal Medicine, 175*(11), 1789-1791.

Linna, M., Hakkinen, U., & Linnakko, E. (1998). An econometric study of costs of teaching and research in Finnish hospitals. *Health Econ, 7*(4), 291-305.

Magnussen, J., Vranbaek, K., Saltman, R. (2009). Nordic Health Care Systems. Recent reforms and current policy challenges. *Open University Press: Maidenhead and New York.*

Marshall, M., & Bindman, A. B. (2016). THe role of government in health care reform in the united states and England. *JAMA Internal Medicine, 176*(1), 9-10.

McDonnell, W. M., Carpenter, P., Jacobsen, K., & Kadish, H. A. (2015). Relative productivity of nurse practitioner and resident physician care models in the pediatric emergency department. *Pediatr Emerg Care, 31*(2), 101-106. doi:10.1097/PEC.0000000000000349

McWilliams, J. M., Chernew, M. E., Landon, B. E., & Schwartz, A. L. (2015). Performance Differences in Year 1 of Pioneer Accountable Care Organizations. *New England Journal of Medicine, 372*(20), 1927-1936.

Medin, E., Anthun, K. S., Hakkinen, U., Kittelsen, S. A., Linna, M., Magnussen, J., . . . Rehnberg, C. (2011). Cost efficiency of university hospitals in the Nordic countries: a cross-country analysis. *Eur J Health Econ, 12*(6), 509-519.

Menachemi, N., Yeager, V. A., Welty, E., & Manzella, B. (2015). Are physician productivity and quality of care related? *J Healthc Qual, 37*(2), 93-101. doi:10.1111/jhq.12038

Milstein, A., & Kocher, R. (2014). Widening gaps in the wall obscuring physician performance differences. *JAMA Internal Medicine, 174*(6), 839-840.

Newhouse, J. P., & Sinaiko, A. D. (2007). Estimates of physician productivity: an evaluation. *Health Care Financ Rev, 29*(2), 33-39.

Oliver, A., & Mossialos, E. (2005). European health systems reforms: looking backward to see forward? *J Health Polit Policy Law, 30*(1-2), 7-28.

Rickman, N., & McGuire, A. (1999). Regulating Providers' Reimbursement in a Mixed Market for Health Care. *Scottish Journal of Political Economy, 46*(1), 53-71. doi:10.1111/1467-9485.00120

Ringard, A., & Hagen, T. P. (2011). Are waiting times for hospital admissions affected by patients' choices and mobility? *BMC Health Serv Res, 11*, 170. doi:10.1186/1472-6963-11-170

Rodysill, K. J. (2003). Increasing physician productivity using a physician extender: a study in an outpatient group practice at the Mayo Clinic. *J Med Pract Manage, 19*(2), 110-114.

Roksvaag, K., & Texmon, I. (2012). *Arbeidsmarkedet for helse- og sosialpersonell fram mot år 2035*. http://www.ssb.no/arbeid-og-lonn/artikler-og-publikasjoner/arbeidsmarkedet-for-helse-og-sosialpersonell-fram-mot-aar-2035.

Romley, J. A., Goldman, D. P., & Sood, N. (2015). US hospitals experienced substantial productivity growth during 2002-11. *Health Aff (Millwood), 34*(3), 511-518. doi:10.1377/hlthaff.2014.0587

Romley, J. A., Jena, A. B., O'Leary, J. F., & Goldman, D. P. (2013). Spending and mortality in US acute care hospitals. *Am J Manag Care, 19*(2), e46-54.

Rosta, J., & Aasland, O. G. (2014). Weekly working hours for Norwegian hospital doctors since 1994 with special attention to postgraduate training, work-home balance and the European working time directive: a panel study. *BMJ Open, 4*(10), e005704. doi:10.1136/bmjopen-2014-005704

Rosta, J. A., OG. (2015). Legers arbeidstid og tidsbruk. *Overlegen, 3-2015*, 28.

Rumbold, B. E. S., Judith A. Hurst, Jeremy. Charlesworth, Anita. Clarke, Aileen. (2015). Improving productive efficiency in hospitals: findings from a review of the international evidence. *Health Economics, Policy and Law, 10*, 21-43.

Ryskina, K. L., & Bishop, T. F. (2013). Physicians' lack of awareness of how they are paid: Implications for new models of reimbursement. *JAMA Internal Medicine, 173*(18), 1745-1746. doi:10.1001/jamainternmed.2013.9270

Sandbaek, B. E., Helgheim, B. I., Larsen, O. I., & Fasting, S. (2014). Impact of changed management policies on operating room efficiency. *BMC health services research, 14*(1), 1-10. doi:10.1186/1472-6963-14-224

Sandy, L. G., Haltson, H., Metfessel, B. A., & Reese, C. (2015). Measuring Physician Quality and Efficiency in an Era of Practice Transformation: PCMH as a Case Study. *Ann Fam Med, 13*(3), 264-268.

Schreyogg, J. (2008). A micro-costing approach to estimating hospital costs for appendectomy in a cross-European context. *Health Econ, 17*(1 Suppl), S59-69. doi:10.1002/hec.1323

Simar, L., & Wilson, P. W. (1998). Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. *Management science, 44*(1), 49-61.

Simar, L., & Wilson, P. W. (2000). Statistical inference in nonparametric frontier models: The state of the art. *Journal of productivity analysis, 13*(1), 49-78.

Simoens, S., & Hurst, J. (2006). The Supply of Physician Services in OECD Countries. OECD Health Working Papers, No. 21, OECD Publishing. doi:10.1787/608402211700

Staiger, D. O., Auerbach, D. I., & Buerhaus, P. I. (2009). Comparison of physician workforce estimates and supply projections. *JAMA, 302*(15), 1674-1680. doi:10.1001/jama.2009.1461

Staiger, D. O., Auerbach, D. I., & Buerhaus, P. I. (2010). Trends in the work hours of physicians in the United States. *JAMA, 303*(8), 747-753. doi:10.1001/jama.2010.168

Stecker, E. C., & Schroeder, S. A. (2013). Adding value to relative-value units. *N Engl J Med, 369*(23), 2176-2179. doi:10.1056/NEJMp1310583

Storfa, A. H., & Wilson, M. L. (2015). Physician productivity: issues and controversies. *Am J Clin Pathol, 143*(1), 6-9. doi:10.1309/AJCP5M1FAMLCYNPX

Street, A., & Castelli, A. (2014). Productivity variation persists in the NHS. *Health Serv J, 124*(6390), 16-17.

Stukel, T. A., Fisher, E. S., Alter, D. A., Guttmann, A., Ko, D. T., Fung, K., . . . Lee, D. S. (2012). Association of hospital spending intensity with mortality and readmission rates in Ontario hospitals. *JAMA, 307*(10), 1037-1045. doi:10.1001/jama.2012.265

Sunshine, J. H., Hughes, D. R., Meghea, C., & Bhargavan, M. (2010). What factors affect the productivity and efficiency of physician practices? *Med Care, 48*(2), 110-117. doi:10.1097/MLR.0b013e3181c1611b

Tiemann, O. (2008). Variations in hospitalisation costs for acute myocardial infarction - a comparison across Europe. *Health Econ, 17*(1 Suppl), S33-45. doi:10.1002/hec.1322

Tiemann, O., & Schreyogg, J. (2012). Changes in hospital efficiency after privatization. *Health Care Manag Sci, 15*(4), 310-326. doi:10.1007/s10729-012-9193-z

Tuohy, C. H. (1999). Dynamics of a changing health sphere: the United States, Britain, and Canada. *Health Aff (Millwood), 18*(3), 114-134.

Varabyova, Y., & Schreyogg, J. (2013). International comparisons of the technical efficiency of the hospital sector: panel data analysis of OECD countries using parametric and non-parametric approaches. *Health Policy, 112*(1-2), 70-79. doi:10.1016/j.healthpol.2013.03.003

Verzulli, R., Jacobs, R., Goddard, M. (2011). Do Hospitals Respond to Greater Autonomy? Evidence from the English NHS. *Centre for Health Economics, University of York, UK.*(CHE Research Paper 64), 1-32.

Vitikainen, K., Linna, M., & Street, A. (2010). Substituting inpatient for outpatient care: what is the impact on hospital costs and efficiency? *Eur J Health Econ, 11*(4), 395-404.

Wiley, M. M. (2005). The Irish health system: developments in strategy, structure, funding and delivery since 1980. *Health Econ, 14*(Suppl 1), S169-186. doi:10.1002/hec.1034

Williams, T. E., Jr., Sun, B., Ross, P., Jr., & Thomas, A. M. (2010). A formidable task: Population analysis predicts a deficit of 2000 cardiothoracic surgeons by 2030. *J Thorac Cardiovasc Surg, 139*(4), 835-840; discussion 840-831. doi:10.1016/j.jtcvs.2009.12.004

Wilsford, D. (1994). Path Dependency, or Why History Makes It Difficult But Not Impossible to Reform Health Care Systems in a Big Way. *Journal of Public Policy, 14*, 251–283.

Wilson, M. S., Joiner, K. A., Inzucchi, S. E., Mulligan, G. J., Mechem, M. F., Gross, C. P., & Coleman, D. L. (2006). Improving clinical productivity in the academic setting: a novel incentive plan based on utility theory. *Acad Med, 81*(4), 306-313.

Wranik, D., & Durier-Copp, M. (2011). Framework for the design of physician remuneration methods in primary health care. *Social work in public health, 26*(3), 231-259.

Yasaitis, L., Fisher, E. S., Skinner, J. S., & Chandra, A. (2009). Hospital quality and intensity of spending: is there an association? *Health Aff (Millwood), 28*(4), w566-572.

Zeidel, M. L., Kroboth, F., McDermot, S., Mehalic, M., Clayton, C. P., Rich, E. C., & Kinsey, M. D. (2005). Estimating the cost to departments of medicine of training residents and fellows: a collaborative analysis. *Am J Med, 118*(5), 557-564. doi:10.1016/j.amjmed.2005.02.025

Research highlights:

- Physician productivity not improved after health care reform
- Differences across hospitals in productivity reveal large potential for improvement
- Supporting staff for physicians increase physician productivity significantly