

Association of Primary Care Physician Compensation Incentives and Quality of Care in the United States, 2012-2016



David S. Burstein, MD¹, David T. Liss, PhD, and Jeffrey A. Linder, MD, MPH

Feinberg School of Medicine, Division of General Internal Medicine and Geriatrics, Department of Medicine, Northwestern University, Chicago, IL, USA.

BACKGROUND: Physician compensation incentives may have positive or negative effects on clinical quality.

OBJECTIVE: To assess the association between various physician compensation incentives on technical indicators of primary care quality.

DESIGN: Cross-sectional, nationally representative retrospective analysis.

PARTICIPANTS: Visits by adults to primary care physicians in the National Ambulatory Medical Care Survey from 2012-2016. We analyzed 49,580 sampled visits, representing 1.45 billion primary care visits.

MAIN MEASURES: We assessed the association between 5 compensation incentives – quality measure performance, patient experience scores, individual productivity, practice financial performance, or practice efficiency – and 10 high-value and 7 low-value care measures as well as high-value and low-value care composites.

KEY RESULTS: Quality measure performance was an incentive in 22% of visits; patient experience scores, 17%; individual productivity, 57%; practice financial performance, 63%; and practice efficiency, 12%. In adjusted models, none of the compensation incentives were consistently associated with individual high- and low-value measures. None of the compensation incentives were associated with high- or low-value care composites. For example, quality measure performance compensation was not significantly associated with high-value care (visits with quality incentive, 47% of eligible measures met; without quality incentive, 43%; adjusted odds ratio [aOR], 1.02; 95% confidence interval [CI], 0.91 to 1.15) or low-value care (aOR, 0.99; 95% CI, 0.82-1.19). Physician compensation incentives that might be expected to increase low-value care did not: patient experience (aOR for low-value care composite, 0.83; 95% CI, 0.65-1.05), individual productivity (aOR, 1.03; 95% CI, 0.88-1.22),

and practice financial performance (aOR, 1.05; 95% CI, 0.81-1.36).

CONCLUSION: In this retrospective, cross-sectional, nationally representative analysis of care in the United States, physician compensation incentives were not generally associated with more or less high- or low-value care.

KEY WORDS: Quality of Care; Physician Compensation; Incentives; Primary Care.

J Gen Intern Med 37(2):359-66

DOI: 10.1007/s11606-021-06617-8

© Society of General Internal Medicine 2021

INTRODUCTION

Physician practice behaviors may be influenced by health systems and payers' use of physician compensation incentives.¹⁻³ Physician compensation incentives can be based on performance on quality measures, patient experience scores, productivity, practice finances, and practice profiling (also called practice efficiency). Physician compensation incentives have received attention due to increasing use of commercial value-based purchasing (VBP) strategies (i.e., pay-for-performance and accountable care organizations)^{4, 5} and an intent by Medicare to link 90% of all payments to quality.⁶

Despite interest and appeal, the relationship between physician compensation incentives and clinical quality is mixed. In ambulatory care, financial incentives may improve care processes,⁷ yet their effect on outcomes has not been clearly demonstrated.^{8, 9} Additionally, physician compensation incentives have the potential for unintended consequences: wasteful efforts to game the system;¹⁰ distraction from unincentivized activities;¹¹ adverse effects on the physician-patient relationship;¹² reduced physician autonomy;¹³ negative effects on workplace morale;^{14, 15} increased physician burnout,^{16, 17} and incentivizing low-value care, such as increased unnecessary antibiotic¹⁸⁻²⁰ or opioid prescribing.²¹

A prior study of primary care physicians in the United States found no association between physician compensation incentives and high-value care in 2007-2008.²² A more recent study found no differences in quality between primary care physicians who were practice owners and those whose compensation was based on salary, productivity, or some

Prior Presentations Burstein DS, Liss DT, Linder JA. Patient experience-based physician compensation and quality of care in the United States, 2012-2016. Abstracts from the 2020 Annual Meeting of the Society of General Internal Medicine. *J Gen Intern Med.* 2020 Jul; 35(Suppl 1):219.

Burstein DS, Liss DT, Linder JA. Patient experience-based physician compensation and quality of care in the United States, 2012-2016. Oral Presentation, Lipkin Award Nominee. Society of General Internal Medicine Annual Meeting. May 15th, 2020. (presented virtually due to COVID19).

Received November 4, 2020

Accepted January 7, 2021

Published online April 14, 2021

combination.²³ This study did not assess the association of individual compensation incentives on quality of care.

An updated assessment of physician compensation incentives and quality of care among primary care physicians in the United States can inform ongoing evaluation and adoption of value-based reimbursement strategies. We conducted a cross-sectional analysis of nationally representative primary care visits to assess the association between physician compensation incentives and quality of care, including high- and low-value care.

METHODS

Data Source

The current study was an analysis of data from the National Ambulatory Medical Care Survey (NAMCS), which is a nationally representative, weighted survey of ambulatory visits to non-federally employed office-based physician practices in the United States. The National Center for Health Statistics (NCHS), which conducts the NAMCS, uses stratified sampling to recruit physicians each year. For physicians who consent to participate, the NCHS collects practice and physician-level information and samples visits during a randomly selected 1 to 2 week period. Census field representatives and medical staff use an automated survey tool to collect visit information from medical charts. The NCHS assigns visit weights to each sampled visit to account for clustering, sampling probability, and non-response to derive national annual visit rates.²⁴ Physician participation in the NAMCS from 2012-2016 ranged from 46% to 59%. The number of unweighted sampled visits among primary care physicians ranged from 76,330 in 2012 to 13,165 in 2016. The NCHS institutional review board approved the public use of the NAMCS data, including a waiver of informed consent of participating patients.

The NAMCS captures comprehensive patient and practice information for each visit across 17 medical specialties. Data collected includes top 3 reasons for the visit; 3 diagnosis codes for years 2012-2013 and 5 diagnosis codes for years 2014-2016 (ICD-9-CM for years 2012-2015 and ICD-10-CM for year 2016); sociodemographics; limited medical history for common medical problems; comprehensive information on the patient's medication list (up to 10 medications from years 2012-2013 and up to 30 medications from years 2014-2016) and whether the medication was new or continued; select vital signs; tests ordered; services and referrals provided; and limited lab results. Physician-level information includes practice ownership, physician-employee status, electronic health record capabilities, and practice revenue structure.

One of the NAMCS survey items asked participants, "please indicate whether the practice explicitly considers the following in determining your compensation." Responses to this question include performance on quality measures ("specific measures of quality, such as rates of preventive services

for your patients"), patient experience scores ("results of satisfaction surveys from your own patients"), individual productivity ("factors that reflect your own productivity"), financial performance of the practice ("the overall financial performance of the practice"), and practice profiling ("results of practice profiling, that is, comparing your pattern of using medical resources with that of other physicians"). For this analysis, we recoded items left blank for this survey item as "No" (1.31% of weighted total).

Measures

We constructed and analyzed all variables at the visit-level. We included visits to primary care physicians by adults age greater than 18. We defined primary care physicians as internal medicine, family medicine, and general practice. We excluded visits during which the patient was not seen by a physician.

We created 17 binary outcome measures based on commonly endorsed, evidence-based quality measures.²⁵⁻²⁷ The 17 individual measures fall into two categories: High-Value Care and Low-Value Care (Medical Overuse; **Appendix 1**).

We assessed performance on each measure as the proportion of patients who fulfilled the numerator criteria divided by the number of patients who were eligible for the denominator criteria of a given measure. Each numerator criteria was binary at the visit-level (i.e., met or not met during the visit). Each denominator was the number of visits which were eligible for a given measure. For example, in the antibiotics for upper respiratory infection (URI) & bronchitis measure, the numerator was the number of visits during which a new antibiotic was prescribed, and the denominator was the number of visits for URI or bronchitis.

We tested the difference in the proportions for each outcome measure based on whether the visit was to a physician with or without a given compensation incentive. We independently tested 5 compensation incentives: Quality measures, patient experience, productivity, financial performance of the practice, and practice profiling. The referent group for each comparison was visits to physicians who did not have that incentive (e.g., visits with incentives for quality measures compared to no incentives for quality measures). Clinically relevant exclusions were applied equally to both the numerator and denominator for each measure.

We also created continuous, 0-1 composite measures of high- and low-value to summarize the visit-weighted mean proportion of eligible measures met.²⁵ To account for visits which were eligible for multiple measures, we multiplied the NAMCS survey weight by the number of eligible measures for a given visit.

We used existing variables in the NAMCS dataset as covariates: year, patient age, race/ethnicity, number of chronic diseases, patient's insurance at the visit, physician specialty, region of practice, office location, and office type. We constructed a covariate for physician owner/employee status using

variables practice ownership and physician-employee status. We constructed a covariate for electronic health record capabilities based on Stage 2 Meaningful Use criteria,²⁸ which became effective in 2014, that we used in our entire analysis. For missing data, the NCHS imputes values for age, sex, and race/ethnicity.²⁴ For number of chronic diseases, we imputed missing data (1.18% of sample) by using the number of prescribed medications as a surrogate marker for chronic illness burden. Missing patient insurance information and physician-level variables are recorded as “unknown.” We recoded missing physician-level variables as “other” or “no” as appropriate for the variable in our analysis.

Statistical Analysis

Our outcome variables of interest were performance on each of the 17 individual outcomes measures of overuse and clinical care quality and the 2 composite measures. Our independent variables were whether or not the visit was to a physician who received compensation based on each incentive: performance on measures of quality, patient experience scores, productivity, financial performance of the practice, or practice profiling.

We used sampling weights for each visit to create a multi-level model which accounts for clustering, sampling probability, and non-response bias.²⁴ Following NCHS guidelines,²⁹ we ensured that statistical tests were based on a sample of at least 30 patient records, that each estimate of the weighted data had a relative standard error (RSE) less than 30%, that the item nonresponse rate was <30%, and all records in the data files were used in each analysis to ensure correct sample variance. We excluded several measures that did not meet these criteria (**Appendix 2**). We used STATA statistical software (version 16.0; College Station, TX) for all analyses. All regressions were weighted for complex survey weighting methods using STATA’s ‘svyset’ suite of commands. We used ‘svy, subpop(x): test’ to estimate our regression models.

We used the chi-square test to assess whether visits to physicians with a certain compensation incentive is associated with performance on 17 individual quality measures. Performance on each quality measure was expressed as the proportion of patients received appropriate care who received appropriate care at the visit as defined by that measure. For each of the 17 individual quality measures, we estimated an unadjusted logistic regression model. We then estimated a multivariable logistic regression model for each individual measure which adjusted for covariates to account for patient and physician characteristics which could be responsible for differences in quality of care. We treated all variables as unordered categorical variables, except for age, which was a continuous variable. Due to the number of individual measures that were tested, we expected that 4 or 5 measures would be significant by chance alone.

For the composite measures, we calculated the weighted mean proportion of quality measures met at visits to

physicians whose compensation was or was not determined by each compensation incentive. We used fractional logistic regression to examine the independent association between each compensation incentive and performance on our composite quality measures.²⁵ Fractional logistic regression is a quasi-likelihood model used when the dependent variable – as with the proportions used in our analysis – is between 0 and 1.³⁰ For both composite measures, we estimated an unadjusted regression model and a multivariable regression model with the same covariates described above to adjust for patient and physician characteristics. All p-values were two-tailed and we considered a value of <0.05 as significant.

In a sensitivity analysis, we omitted physicians who answered the NAMCS compensation survey question with responses “item left blank,” “compensation unknown” (7.5%), and “refused to answer” (0.9%), and performed the same tests as above. In our initial analysis, we chose to code “compensation unknown” as “No” because incentives only work if physicians are aware of them.^{1, 31}

Role of the Funding Source

The Society of General Internal Medicine had no role in the design, conduct, and reporting of this study.

RESULTS

From 2012 to 2016, there were 49,850 sampled visits, which represented 1.45 billion visits to primary care physicians in the United States. Quality measure performance was used as an incentive in 22% of visits, patient experience scores for 17% of visits, productivity for 57% of visits, financial performance of the practice for 63% of visits, and practice profiling for 12% of visits (**Table 1**).

Patient sex, age, race/ethnicity, and patient insurance status were similar among visits to physicians with each compensation incentive. The various physician compensation incentives differed by region, urban versus rural, physician ownership, practice revenue, electronic health record capabilities. Incentives for quality measures were least common in the South, incentives for patient experience were most common in the West, and incentives for practice profiling were most common in the Northeast. Rural areas were more likely than urban areas to have incentives for productivity and financial performance of the practice. Physician owners were most likely to have incentives for financial performance of the practice.

In unadjusted models, physician compensation for quality measure performance was associated with better performance on our high value composite measure (43% of eligible measures met for visits to physicians without compensation for quality measure performance, versus 47% of eligible measures met for visits to physicians with compensation for quality measure performance; odds ratio (OR), 1.16; 95% 1.04 – 1.30; **Appendix 3A**).

Table 1 Patient and Physician Characteristics at Visits to Primary Care Physicians, 2012-2016 (n=1,452,525,992)

| Characteristic | Any Visit (%) | Visit Characteristics with Each Physician Compensation Incentive | | | | | |
|--|---------------|--|---|---|--|------------------------------------|------|
| | | Performance on Quality Measures (22% of visits) | Patient Experience Scores (17% of visits) | Individual Productivity (57% of visits) | Practice Financial Performance (63% of visits) | Practice Profiling (12% of visits) | |
| Sex | | | 0.76 | 0.67 | 0.06 | 0.23 | 0.79 |
| Female | 57 | 57 | 58 | 58 | 58 | 57 | 58 |
| Male | 43 | 43 | 42 | 42 | 42 | 43 | 42 |
| Age, y | | 0.18 | 0.50 | 0.75 | 0.02 | 0.51 | |
| 18-44 | 27 | 26 | 26 | 27 | 26 | 27 | |
| 45-65 | 39 | 38 | 38 | 39 | 39 | 41 | |
| ≥65 | 34 | 36 | 36 | 34 | 35 | 32 | |
| Race | | 0.43 | 0.61 | 0.49 | 0.27 | 0.79 | |
| White | 82 | 80 | 82 | 82 | 81 | 81 | |
| Black | 12 | 13 | 13 | 13 | 12 | 13 | |
| Other | 6 | 7 | 5 | 5 | 7 | 6 | |
| Ethnicity | | 0.73 | 0.83 | 0.94 | 0.46 | 0.18 | |
| Latino | 14 | 14 | 14 | 13 | 14 | 17 | |
| Non-Latino | 86 | 86 | 86 | 87 | 86 | 83 | |
| Patient's Insurance | | 0.86 | 0.90 | 0.39 | 0.01 | 0.54 | |
| Private Insurance | 48 | 48 | 47 | 49 | 48 | 48 | |
| Medicare | 31 | 32 | 33 | 31 | 33 | 29 | |
| Medicaid | 8 | 9 | 8 | 7 | 9 | 10 | |
| Other | 12 | 11 | 13 | 12 | 10 | 13 | |
| Physician Specialty | | 0.77 | 0.80 | 0.30 | 0.15 | 0.69 | |
| Family Medicine/General Practice | 61 | 59 | 59 | 59 | 58 | 63 | |
| Internal Medicine | 39 | 41 | 41 | 41 | 42 | 38 | |
| Region | | 0.04 | <0.01 | 0.09 | <0.01 | 0.02 | |
| Northeast | 18 | 20 | 15 | 16 | 21 | 29 | |
| Midwest | 21 | 28 | 28 | 24 | 20 | 20 | |
| South | 38 | 26 | 23 | 39 | 34 | 26 | |
| West | 23 | 26 | 34 | 21 | 24 | 25 | |
| Setting | | <0.01 | <0.01 | 0.11 | 0.53 | <0.01 | |
| Urban | 87 | 93 | 92 | 89 | 88 | 93 | |
| Rural | 13 | 7 | 8 | 11 | 12 | 7 | |
| Office Type | | <0.01 | <0.01 | <0.01 | <0.01 | <0.01 | |
| Solo | 34 | 17 | 12 | 21 | 39 | 16 | |
| Group | 56 | 65 | 63 | 67 | 52 | 59 | |
| Other | 10 | 18 | 25 | 12 | 9 | 25 | |
| Physician Owner/Employee Status | | <0.01 | <0.01 | <0.01 | <0.01 | <0.01 | |
| Physician Owner (Full or Part) | 54 | 34 | 23 | 44 | 64 | 29 | |
| Employee of Provider | 27 | 37 | 38 | 33 | 20 | 34 | |
| Employee of Payer | 16 | 25 | 37 | 20 | 13 | 33 | |
| Other | 4 | 4 | 2 | 3 | 4 | 5 | |
| % of Practice Revenue | | <0.01 | <0.01 | 0.30 | <0.01 | 0.10 | |
| Mixed | 57 | 41 | 43 | 55 | 52 | 47 | |
| Medicaid >25% | 15 | 20 | 15 | 15 | 18 | 21 | |
| Private >50% | 28 | 40 | 42 | 30 | 30 | 32 | |
| Electronic Health Record Capabilities | | <0.01 | <0.01 | <0.01 | 0.17 | <0.01 | |
| Meets Meaningful Use* | 66 | 80 | 84 | 72 | 63 | 79 | |
| Does not Meet Meaningful Use* | 20 | 14 | 10 | 17 | 21 | 14 | |
| No EHR | 15 | 6 | 6 | 11 | 16 | 7 | |

*Based on Stage 2 CMS Meaningful Use Requirements

Provider = Physician, medical/academic health center, community health center

Payer = Insurance company, health plan, or HMO

In adjusted models (Table 2), there were no consistent associations between any compensation incentive and individual high-value or low-value care measures. Incentives for quality measures were not associated with an increase in high-value care or a decrease in low-value care on any individual measure, though there was an association with less use of angiotensin-converting enzyme (ACE) inhibitors (47% without an incentive versus 43% with an incentive; aOR,

0.74; 95% CI, 0.59 to 0.93). Incentives for patient experience, productivity, and financial performance were not associated with an increase in low-value care for any individual measure, except for an association between productivity incentives and increased opioid prescribing (12% without an incentive versus 19% with an incentive; aOR, 1.55; 95% CI, 1.03 to 2.32). Incentives for patient experience were also associated with less unnecessary cardiac screening (9.4% without an incentive

Table 2 Adjusted Associations Between Visit-level Physician Compensation Incentives and Measures of Quality of Care

| Quality Measures from the NAMCS Dataset | Performance on Quality Measures aOR (95% CI) | Patient Experience Scores aOR (95% CI) | Individual Productivity aOR (95% CI) | Practice Financial Performance aOR (95% CI) | Practice Profiling aOR (95% CI) |
|---|---|---|---|--|------------------------------------|
| High Value Care | | | | | |
| Composite High Value Care‡ | 1.02 (0.91 - 1.15) | 0.99 (0.86 - 1.13) | 0.96 (0.90 - 1.07) | 0.98 (0.90 - 1.07) | 1.13 (0.96 - 1.34) |
| Tobacco Cessation Counselling | 0.97 (0.69-1.35) | 0.97 (0.66 - 1.44) | 0.94 (0.71 - 1.25) | 0.93 (0.70 - 1.22) | 1.47 (1.00 - 2.14) |
| Weight Loss Counselling | 1.19 (0.68 - 2.08) | 0.77 (0.47 - 1.28) | 0.85 (0.58 - 1.25) | 0.77 (0.54 - 1.08) | 0.96 (0.53 - 1.74) |
| Quality Measures in Diabetes | | | | | |
| Recent A1c Within Past Year◊ | 1.21 (0.91 - 1.60) | 1.27 (0.92 - 1.77) | 1.01 (0.80 - 1.28) | 1.05 (0.84 - 1.32) | 1.30 (0.88 - 1.91) |
| Last A1c of <8%◊ | 1.04 (0.81 - 1.33) | 1.10 (0.82 - 1.47) | 1.06 (0.85 - 1.32) | 0.99 (0.79 - 1.22) | 1.24 (0.86 - 1.78) |
| Statin Use | 0.84 (0.65 - 1.09) | 1.01 (0.76 - 1.34) | 0.97 (0.78 - 1.24) | 1.00 (0.80 - 1.23) | 1.02 (0.70 - 1.47) |
| ACE/ARB Use if HTN | 0.74 (0.59 - 0.93) | 0.86 (0.66 - 1.10) | 0.86 (0.69 - 1.06) | 0.88 (0.81 - 1.08) | 0.96 (0.66 - 1.40) |
| Eye, Foot or Urine Screening | 1.11 (0.78 - 1.57) | 0.92 (0.63 - 1.35) | 1.02 (0.72 - 1.47) | 1.24 (0.89 - 1.73) | 0.84 (0.54 - 1.30) |
| Beta-blocker Use in CAD | 0.86 (0.60 - 1.27) | 0.82 (0.54 - 1.24) | 0.89 (0.64 - 1.24) | 1.00 (0.71 - 1.43) | 0.84 (0.47 - 1.48) |
| Statin use in CVA/TIA | 1.07 (0.63 - 1.83) | 1.27 (0.71 - 2.29) | 0.93 (0.60 - 1.45) | 0.84 (0.57 - 1.24) | 1.23 (0.60 - 2.55) |
| Blood Pressure at Goal in HTN | 1.12 (0.87 - 1.43) | 0.89 (0.68 - 1.16) | 0.90 (0.77 - 1.05) | 1.07 (0.93 - 1.23) | 1.25 (0.96 - 1.62) |
| Low Value Care | | | | | |
| Composite Low Value Care‡ | 0.99 (0.82 - 1.19) | 0.83 (0.65 - 1.05) | 1.03 (0.88 - 1.22) | 1.05 (0.81 - 1.36) | 1.06 (0.83 - 1.36) |
| Antibiotics for URI or Bronchitis | 1.54 (0.96 - 2.47) | 1.11 (0.67 - 1.86) | 1.03 (0.69 - 1.55) | 1.01 (0.66 - 1.55) | 0.92 (0.45 - 1.88) |
| Opioids for Low Back Pain | 0.95 (0.55 - 1.63) | 0.81 (0.48 - 1.37) | 1.55 (1.04 - 2.32) | 0.78 (0.53 - 1.15) | 0.81 (0.42 - 1.88) |
| Inappropriate Cancer Screening | 0.99 (0.61 - 1.61) | 0.55 (0.28 - 1.06) | 0.87 (0.59 - 1.29) | 0.83 (0.55 - 1.24) | 1.15 (0.64 - 2.09) |
| Cardiac Screening without Indication at GME | 0.74 (0.49 - 1.17) | 0.63 (0.40 - 0.99) | 1.28 (0.90 - 1.82) | 0.75 (0.54 - 1.05) | - |
| Indication CBC Screening | 1.25 (0.89 - 1.76) | 1.06 (0.68 - 1.67) | 1.19 (0.86 - 1.66) | 0.90 (0.66 - 1.23) | 1.09 (0.72 - 1.64) |
| High Risk Medication Use in Elderly | 0.97 (0.64 - 1.48) | 0.80 (0.53 - 1.19) | 0.82 (0.64 - 1.06) | 0.94 (0.73 - 1.21) | 0.93 (0.61 - 1.42) |
| High Risk Sulfonylurea Use in Diabetes | 1.21 (0.76 - 1.93) | 1.02 (0.62 - 1.69) | 1.11 (0.83 - 1.50) | 0.98 (0.72 - 1.32) | 1.20 (0.71 - 2.04) |

For each individual measure, we used logistic regression to compare visits to physicians with and without a given compensation incentive, adjusted for fixed year effects, patient age, race/ethnicity, patient's insurance at the visit, region of practice, office type, physician employment/ownership status, electronic health record use, and amount of practice revenue by type of insurance. For the two composite measures, we used fractional logistic regression with adjustment for the same covariates as our logistic regression models. Bolded text indicates statistically significant associations ($p < 0.05$).

‡To create estimates which accounted for visits which were eligible for more than one measure, we multiplied the NAMCS survey weights by the number of eligible indicators for a given visit. This step was performed only for our composite measures.

◊This analysis is for years 2012-2014 and 2016. A1c data is not available for the year 2015.

URI = upper respiratory infection; GME = general medical examination; CAD = coronary artery disease; CVA = cerebrovascular accident

versus 6% with an incentive; aOR, 0.63; 95% CI 0.40 to 0.99). Incentives for practice profiling were associated with an increase in tobacco cessation counselling in smokers (19% without an incentive versus 26% with an incentive; aOR, 1.47; 95% CI, 1.00 to 2.14). None of the incentives were associated with performance on our high- or low-value composite measures.

In our sensitivity analysis, practice profiling was associated with more high value care (42% of eligible measures met per visit without incentive versus 47% with incentive; aOR, 1.19; 95% CI, 1.00 - 1.40). The association of productivity incentives and increased opioid prescribing in our initial analysis was no longer statistically significant (13% vs 19%; aOR, 1.37; 95% CI, 0.88 - 2.14). Other findings from sensitivity analysis are annotated in Appendix 3A-E.

DISCUSSION

In a nationally representative survey of visits to primary care physicians in the United States, physician compensation

incentives for quality measure performance, patient experience, productivity, financial performance of the practice, and practice profiling were not associated with quality of care. In multivariable modelling examining associations between physician compensation incentives and quality, 4 associations were significant. Only 1 significant association makes intuitive sense (productivity incentives “increasing” opioid prescriptions). With 5 physician compensation incentives evaluated across 19 quality measures (95 measures total and $p = 0.05$), we would expect 4 or 5 to be significant by chance alone. This suggests that individual financial incentives in physician compensation by themselves may not significantly influence quality of care.

Small physician compensation incentives are often ineffective at improving health care quality.^{8, 9, 32} From 2012-2016, less than 5% to 10%, on average, of primary care physicians' compensation was based on quality, patient experience, and resource utilization.³³⁻³⁵ Thus, our findings could be consistent with the ineffectiveness of small incentives.

Larger compensation incentives may improve quality when used within VBP reimbursement strategies.^{9, 32} For example, the

Fairview Health Services Accountable Care Organization (ACO), which based 40% of clinician compensation on performance incentives, resulted in improvement in quality among low performing clinicians. The United Kingdom's Quality and Outcomes Framework (QOF), which used to base 25% to 30% of physician compensation on clinical performance, may have improved performance in care processes. Overall, improvements in quality resulting from physician compensation incentives within VBP strategies may exist for processes of care, and have not yet been clearly demonstrated for health outcomes.⁷

The potential improvements in quality from highly incentivized VBP strategies must be weighed against the unintended negative consequences that can arise from performance measurement. Theoretical examples of unintended consequences of performance measurement include neglect of unmeasured outcomes, ignoring unincentivized activities, adverse selection of patients which can increase health inequities, and erosion of trust.¹⁰ The evidence for these potential harms is mixed.⁷ For physicians, performance measurement can be a source of stress and burnout due to loss of autonomy, documentation burden, unreasonable expectations, and lack of control over measured outcomes.^{13, 15, 36–38} In some cases, poorly designed incentives can result in worse medical care,^{18–21} which was not shown in our study.

Interestingly, incentives in the QOF were later revised down to 15% due to their unintended consequences,³⁹ so there appears to be an upper limit to how much of physicians' compensation should be based on clinical performance. Due to the potential negative consequences of performance measurement, some have suggested that physician compensation incentives may at times be unnecessarily relied upon for motivating clinical improvement.^{40–42} When performance on quality measures is not tied to compensation, tracking quality measures can be valuable in other ways. Examples include population management, internal quality improvement goals, transparency for consumers, and setting thresholds for ACO incentives and balancing cost-containment measures in VBP models.

Another explanation for the lack of an association between compensation incentives for quality and improvements in care is improper implementation of performance measures. Measures must be clear and simple, and have appropriate size, benchmarks for performance, and measurement intervals.¹ Measures must have capacity to change over time and should emphasize areas that most warrant performance improvement. Measures should have clinician buy-in, and would best have clinician involvement in measure selection and an evidence-based rationale. Incentives and measures should be congruent with clinician, institution, and community priorities.⁴³ Measures must be attributed to and within the control of the appropriate clinician(s) and/or other stake holders, and there must be adequate clinical and administrative support to achieve the measure. Measures which are prone to gaming should be avoided.³⁶ In our sample, 22% of visits were to physicians with incentives for quality measure performance,

17% for patient experience, and 12% for practice efficiency. We were unable to ascertain the proportion of physicians at these visits who were practicing in a setting which appropriately implemented their performance-based compensation incentives.

Our results have limitations. First, causation cannot be inferred from this retrospective, serial cross-sectional study. Strictly speaking, we have only found a lack of association between the NAMCS physician compensation incentives and the NAMCS measures of quality. Second, we were unable to specify the degree, amount, or mechanism of each incentive. Third, we may have had low power to detect differences on some measures. For example, we may have had low power on individual measures for antibiotics prescriptions for URI (unweighted visits, $n = 1,238$) and statin use in patients with history of cerebrovascular accident ($n = 1,114$). Fourth, we did not analyze how incentives were combined or used simultaneously for physicians. Sampled physicians may have been influenced by 0 to 5 different compensation incentives which may have been larger and thus more impactful than other incentives. Additionally, the mix of compensation incentives for each physician is complicated, and could bias results toward the null. Fifth, our analysis did not correct for multiple testing, but we considered the possibility of chance in interpreting our findings. Sixth, our composite measures are skewed toward measures that can be commonly met in a given visit (e.g., inappropriate cancer screening in elderly patients), rather than measures that may be more clinically significant or impacted by performance incentives (e.g., statin use in patients with history of cerebrovascular accident or yearly hemoglobin A1c for patients with diabetes). Seventh, we created our measures for quality based on their availability in the NAMCS (e.g., weight loss counseling for obesity), rather than measures which would otherwise reflect the best or most comprehensive treatment plan for a given problem (e.g., preventing obesity). Eighth, self-report of physician incentives may result in measurement error. Our sensitivity analysis showed an association between practice profiling and more high value care, which we also attribute to third party scrutiny of physician practice patterns which promotes higher quality care.

CONCLUSIONS

In a nationally representative, cross-sectional survey of primary care in the United States, we found no association between physician compensation incentives and quality of care. Physician compensation incentives have unintended negative consequences. Prior to implementation, physician compensation incentives should be correctly sized, properly implemented, and their benefits on improving quality should be clear.

Supplementary Information. The online version contains supplementary material available at <https://doi.org/10.1007/s11606-021-06617-8>.

Contributors: None.

Corresponding Author: David S. Burstein, MD; Feinberg School of Medicine, Division of General Internal Medicine and Geriatrics, Department of Medicine, Northwestern University, Chicago, IL, USA (e-mail: David.Burstein@northwestern.edu).

Funding This work was supported by the Northwestern University Division of General Internal Medicine and Geriatrics Clinical Research Fellowship and the 2019 Founder's Grant from the Society of General Internal Medicine. Dr. Linder is supported by a contract from the Agency for Healthcare Research and Quality (HHSP233201500020I) and grants from the National Institute on Aging (R33AG057383, R33AG057395, P30AG059988, R01AG069762), the Agency for Healthcare Research and Quality (R01HS026506, R01 HS028127), and the Peterson Center on Healthcare.

Declarations:

Conflicts of Interest: We have no conflicts of interest.

REFERENCES

- Emanuel EJ, Ubel PA, Kessler JB, et al. Using Behavioral Economics to Design Physician Incentives That Deliver High-Value Care. *Ann Intern Med.* 2016 Jan 19;164(2):114-9.
- Song Z, Ayanian JZ, Wallace J, et al. Unintended consequences of eliminating medicare payments for consultations. *JAMA Intern Med.* 2013;173(1):15-21.
- Hillman AL, Pauly NM, Kerstein JJ. How do financial incentives affect physicians' clinical decisions and the financial performance of health maintenance organizations? *N Engl J Med.* 1989 Jul 13;321(2):86-92.
- Ginsburg PB, Patel KK. Physician Payment Reform - Progress to Date. *N Engl J Med.* 2017 Jul 20;377(3):285-92.
- Damberg CL, Sorbero ME, Lovejoy SL, et al. Measuring Success in Health Care Value-Based Purchasing Programs: Findings from an Environmental Scan, Literature Review, and Expert Panel Discussions. *Rand Health Q.* 2014 Dec 30;4(3):9.
- Burwell SM. Setting Value-Based Payment Goals — HHS Efforts to Improve U.S. Health Care. *N Engl J Med.* 2015;372:897-99.
- Mendelson A, Kondo K, Damberg C, et al. The Effects of Pay-for-Performance Programs on Health, Health Care Use, and Processes of Care: A Systematic Review. *Ann Intern Med.* 2017 Mar 7;166(5):341-53.
- Roland M, Dudley RA. How Financial and Reputational Incentives Can Be Used to Improve Medical Care. *Health Serv Res.* 2015;50 Suppl 2(Suppl 2):2090-2115.
- Cattel D, Eijkenaar F. Value-Based Provider Payment Initiatives Combining Global Payments With Explicit Quality Incentives: A Systematic Review. *Med Care Res Rev.* 2020 Dec;77(6):511-37.
- Mannion R, Braithwaite J. Unintended consequences of performance measurement in healthcare: 20 salutary lessons from the English National Health Service. *Intern Med J.* 2012;42(5):569-74.
- Deber R, Schwartz R. What's Measured Is Not Necessarily What Matters: A Cautionary Story from Public Health. *Health Policy.* 2016;12(2):52-64.
- Powell AA, White KM, Partin MR, et al. Unintended consequences of implementing a national performance measurement system into local practice. *J Gen Intern Med.* 2012;27(4):405-12.
- Kansagara D, Tuepker A, Joos S, et al. Getting performance metrics right: a qualitative study of staff experiences implementing and measuring practice transformation. *J Gen Intern Med.* 2014;29S
- Damschroder LJ, Robinson CH, Francis J, et al. Effects of performance measure implementation on clinical manager and provider motivation. *J Gen Intern Med.* 2014;29 Suppl 4(Suppl 4):877-84.
- Hysong SJ, SoRelle R, Broussard Smitham K, et al. Reports of unintended consequences of financial incentives to improve management of hypertension. *PLoS One.* 2017;12(9):e0184856.
- Miller P. The Medical Profession's Future: A Struggle Between Caring For Patients And Bottom-Line Pressures. *Health Aff (Millwood).* 2016 Mar;35(3):407-10.
- Cassel CK, Jain SH. Assessing individual physician performance: does measurement suppress motivation? *JAMA.* 2012 Jun 27;307(24):2595-6.
- Baker DW, Gaseem A; American College of Physicians' Performance Measurement Committee. Evidence-based performance measures: preventing unintended consequences of quality measurement. *Ann Intern Med.* 2011;155(9):638-40.
- Ashworth M, White P, Jongsma H, et al. Antibiotic prescribing and patient satisfaction in primary care in England: cross-sectional analysis of national patient survey data and prescribing data. *Br J Gen Pract.* 2016 Jan;66(642):e40-6.
- Sharp AL, Shen E, Kanter MH, et al. Low-value antibiotic prescribing and clinical factors influencing patient satisfaction. *Am J Manag Care.* 2017 Oct;23(10):589-94.
- Sites BD, Harrison J, Herrick MD, et al. Prescription opioid use and satisfaction with care among adults with musculoskeletal conditions. *Ann Fam Med.* 2018 Jan;16(1):6-13.
- Bishop TF, Federman AD, Ross JS. Association between physician quality improvement incentives and ambulatory quality measures. *Am J Manag Care.* 2012 Apr 1;18(4):e126-34.
- Mosqueira AG, Rosenthal M, Barnett ML. The association between primary care physician compensation and patterns of care delivery, 2012-2015. *Inquiry.* Jan-Dec 2019;56:46958019854965
- National Center for Health Statistics. Public Use Microdata File Documentation, National Ambulatory Medical Care Survey, 2016. Hyattsville, MD: National Technical Information Service; 2018.
- Barnett ML, Linder JA, Clark CR, et al. Low-value medical services in the safety-net population. *JAMA Intern Med.* 2017 Jun 1;177(6):829-37.
- Ma J, Stafford RS. Quality of US outpatient care: temporal changes and racial/ethnic disparities. *Arch Intern Med.* 2005 Jun 27;165(12):1254-61.
- Mishuris RG, Linder JA, Bates DW. Using electronic health record clinical decision support is associated with improved quality of care. *Am J Manag Care.* 2014 Oct 1;20(10):e445-52.
- Centers for Medicare & Medicaid Services. Stage 2 Overview Tipsheet. Last updated August, 2012. Available at https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/downloads/Stage2Overview_Tipsheet.pdf. Accessed Dec 2, 2019.
- McCaig LF, Burt CW. Understanding and interpreting the National Hospital Ambulatory Medical Care Survey: Key Questions and Answers. *Ann Emerg Med.* 2012 Dec;60(6):716-21.
- Papke LE, Wooldridge JM. Econometric Methods for Fractional Response Variables with an Application to 401(k) Plan Participation Rates. National Bureau of Economic Research. 1993. https://www-nber-org.ezproxy.galter.northwestern.edu/system/files/working_papers/t0147/t0147.pdf. Accessed Dec 22, 2020.
- Khullar D, Chokshi DA, Kocher R, et al. Behavioral economics and physician compensation—promise and challenges. *N Engl J Med.* 2015 Jun 11;372(24):2281-3.
- Vlaanderen FP, Tanke MA, Bloem BR, et al. Design and effects of outcome-based payment models in healthcare: a systematic review. *Eur J Health Econ.* 2019;20(2):217-32.
- Ryan AM, Shortell SM, Ramsay PP, et al. Salary and Quality Compensation for Physician Practices Participating in Accountable Care Organizations. *Ann Fam Med.* 2015 Jul-Aug;13(4):321-4.
- Damberg CL, Silverman M, Burgette L, et al. Are value-based incentives driving behavior change to improve value? *Am J Manag Care.* 2019 Feb 1;25(2):e26-e32.
- Rosenthal M, Shortell S, Shah ND, et al. Physician practices in Accountable Care Organizations are more likely to collect and use physician performance information, yet base only a small proportion of compensation on performance data. *Health Serv Res.* 2019 Dec;54(6):1214-22.
- Sinsky CA, Privitera MR. Creating a "manageable cockpit" for clinicians: A shared responsibility. *JAMA Intern Med.* 2018 Jun 1;178(6):741-42.
- Khan N, Rudoler D, McDiarmid M, et al. A pay for performance scheme in primary care: Meta-synthesis of qualitative studies on the provider experiences of the quality and outcomes framework in the UK. *BMC Fam Pract.* 2020;21(1):142.

38. Martin B, Jones J, Miller M, et al. Health Care Professionals' Perceptions of Pay-for-Performance in Practice: A Qualitative Metasynthesis. *Inquiry*. 2020 Jan-Dec;57:46958020917491.
39. Roland M, Olesen F. Can pay for performance improve the quality of primary care? *BMJ*. 2016 Aug 4;354:i4058.
40. Kao AC. Driven to Care: Aligning External Motivators with Intrinsic Motivation. *Health Serv Res*. 2015 Dec;50 Suppl 2(Suppl 2):2216-22.
41. Berenson RA, Rice T. Beyond Measurement and Reward: Methods of Motivating Quality Improvement and Accountability. *Health Serv Res*. 2015 Dec;50 Suppl 2(Suppl 2):2155-86.
42. Himmelstein DU, Ariely D, Woolhandler S. Pay-for-performance: toxic to quality? Insights from behavioral economics. *Int J Health Serv*. 2014;44(2):203-14.
43. Kondo KK, Damberg CL, Mendelson A, et al. Implementation Processes and Pay for Performance in Healthcare: A Systematic Review. *J Gen Intern Med*. 2016 Apr;31 Suppl 1(Suppl 1):61-9.

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