Association Between HEDIS Performance and Primary Care Physician Age, Group Affiliation, Training, and Participation in ACA Exchanges

Jill R. Glassman, PhD, MSW¹ , David S. P. Hopkins, PhD¹,
M. Kate Bundorf, PhD, MBA, MPH², Robert M. Kaplan, PhD¹,
Meera V. Ragavan, MD, MPH¹, Alan Glaseroff, MD, ABFM¹, and Arnold Milstein, MD, MPH¹

¹Clinical Excellence Research Center, School of Medicine, Stanford University, Stanford, CA, USA; ²Health Research and Policy, School of Medicine, Stanford University, Stanford, CA, USA.

BACKGROUND: There are a limited number of studies investigating the relationship between primary care physician (PCP) characteristics and the quality of care they deliver.

OBJECTIVE: To examine the association between PCP performance and physician age, solo versus group affiliation, training, and participation in California's Affordable Care Act (ACA) exchange.

DESIGN: Observational study of 2013-2014 data from Healthcare Effectiveness Data and Information Set (HEDIS) measures and select physician characteristics.

PARTICIPANTS: PCPs in California HMO and PPO practices ($n = 5053$) with part of their patient panel covered by a large commercial health insurance company.

MAIN MEASURES: Hemoglobin A1c testing; medical attention nephropathy; appropriate treatment hypertension (ACE/ARB); breast cancer screening; proportion days covered by statins; monitoring ACE/ARBs; monitoring diuretics. A composite performance measure also was constructed.

KEY RESULTS: For the average 35- versus 75-year-old PCP, regression-adjusted mean composite relative performance scores were at the 60th versus 47th percentile (89% vs. 86% composite absolute HEDIS scores; $p < .001$). For group versus solo PCPs, scores were at the 55th versus 50th percentiles (88% vs. 87% composite absolute HEDIS scores; $p < .001$). The effect of age on performance was greater for group versus solo PCPs. There was no association between scores and participation in ACA exchanges.

CONCLUSIONS: The associations between population-based care performance measures and PCP age, solo versus group affiliation, training, and participation in ACA exchanges, while statistically significant in some cases, were small. Understanding how to help older PCPs excel equally well in group practice compared with younger PCPs may be a fruitful avenue of future research.

KEY WORDS: primary care; health care quality; HEDIS performance measures; physician age; physician characteristics.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11606-020-05642-3>) contains supplementary material, which is available to authorized users.

Received August 4, 2019

Accepted January 3, 2020

Published online January 23, 2020

J Gen Intern Med 35(6):1730–5

DOI: 10.1007/s11606-020-05642-3

© Society of General Internal Medicine 2020

INTRODUCTION

Individual physician characteristics, such as age, affiliation with a group versus solo practice, training in family versus internal medicine, and participation in narrow networks, may impact performance. Yet only a limited number of studies have investigated the relationship between provider characteristics and the quality of care they deliver.

Studies on the effects of physician age and affiliation with a group versus solo practice have produced inconsistent results. One study found that Medicare patients treated by solo physicians were less likely to receive timely and appropriate treatment for acute myocardial infarction (AMI) once they were admitted to a hospital.¹ Another study showed 30-day hospital readmissions and cost were worse for physicians affiliated with large group practices compared with those for physicians practicing in small group or solo settings.² Likewise, some studies have found performance drops off for older physicians,³ while others have not.⁴

The introduction of narrow network plans in the Affordable Care Act (ACA) insurance markets introduced the concern that physicians agreeing to participate in these plans at lower payment rates might provide lower quality care; however, there is little evidence to date that this is the case given the complexities involved in determining which providers are included in a narrow network.⁵ Finally, quality-of-care comparisons of internists versus family medicine physicians differ depending on the quality and nature of the metrics examined.^{6, 7}

Even less is known about the relationship between physician characteristics and quality for primary care physicians (PCPs). Care process measures in the Healthcare Effectiveness Data and Information Set (HEDIS) have been widely collected for a long time and may provide a necessary, though not sufficient, indicator of quality of care provided by PCPs. However, HEDIS measures are rarely reported at the

individual physician level,⁸ and clinical significance of differences in scores at any level, much less at the physician level, are not often discussed.

In this study, we used a subset of physician-level HEDIS performance measures to examine their relationship with PCP age, affiliation with solo versus group practice, certification in internal versus family medicine, participation in an ACA exchange, and geographic area.

METHODS

Study Design

This was a cross-sectional, observational study in which we analyzed associations between 2013 and 2014 data on HEDIS measures compiled by the California Healthcare Performance Information System (CHPI), and select physician characteristics, for 5053 primary care physicians (PCPs) affiliated with one of the largest commercial health insurance companies in California. The study was approved by Stanford's Institutional Review Board. Analyses were conducted using Stata version 14.

Data

Physician-level HEDIS performance scores were obtained from CHPI, which operated the only multi-payer claims database in California. In 2016, they released their second cycle of publicly available data on physician and practice site-level performance measures calculated from claims and encounter data for 2013–2014. These represented all claims and encounters for roughly 10 million members continuously enrolled for that period with one of the three largest commercial health insurance plans in California or with Medicare's fee-for-service program. CHPI aggregated these data to calculate 14 select HEDIS clinical quality process measures.⁹ The measures were selected as being appropriate at the individual physician level by an expert Physician Advisory Group convened by CHPI. The present study used data from the 7 measures that were applicable to adult PCPs. Data on physician characteristics were obtained from one of the three commercial insurers and represented 88% of all PCPs with available data in the CHPI database.

Measures

Performance Measures. The following HEDIS performance measures were included: comprehensive hemoglobin A1c testing (diabetes patients); medical attention for nephropathy (diabetes patients); appropriate treatment of hypertension (ACE/ARB, diabetes patients); breast cancer screening (women age 50–69); proportion of days covered: statins; annual monitoring: ACE inhibitors/ARBs; annual, monitoring: diuretics.

For each measure, the score was calculated as a ratio of the number of patients for whom care met the evidence-based criterion over the number of patients eligible for that evidence-based recommended care within the time period January 1, 2013–December 31, 2014, reported in percentage units. CHPI risk-adjusted scores using patient-level age and gender; zip code–level education, poverty, and race/ethnicity; payer; and product (e.g., HMO vs. PPO). CHPI reported each of these performance scores for a given physician only if the following criteria were met: Spearman-Brown prophecy reliability was at least .70, and a minimum of 11 patients were eligible for the measure to protect patient confidentiality. Further, CHPI only reported on measures for which at least 100 PCPs had reportable data. Finally, CHPI attributed patients to a single PCP based on which one the patient had the most outpatient encounters with between January 1, 2013 and December 31, 2014.

Relative Versus Absolute Performance Scores. To account for differences in the distributions of the HEDIS measures, absolute HEDIS performance scores were converted to percentile ranks. Sensitivity analyses were conducted on absolute performance scores, for which interpretations of the clinical significance of differences in scores theoretically can be made.

Composite. The seven available HEDIS scores were averaged to create an index of PCP performance. Composites were created for both relative rank and absolute scores.

Physician Characteristics. We considered six PCP attributes: age; affiliation only with a group staff model, only with a solo practice, or with both; specialty certification in family versus internal medicine; participation in California's ACA exchange program; and geographic practice region of California. Data on physician characteristics were provided by the payer.

Covariates. We used percent of families with household income below poverty level, as defined by the census bureau for the 2014 American Community Survey, within practice zip code, as a measure of SES for the population served by a given PCP.

Statistical Analysis

Separate multiple linear regression models were used to assess the relationship between the composite or a given performance measure and physician characteristic, controlling for all other characteristics and the percent poverty covariate. Separate models with interactions were used to assess moderating effects of affiliation or region on the relationship between physician age and performance.

To investigate whether the relationship between physician age and performance might be nonlinear, piecewise regression

models were estimated using a cut point at age 50 based on patterns in the data and prior literature.³ These models suggested no difference in slopes for different intervals of physician age; thus, linear models were used.

Given this was an exploratory study, we kept all variables in the model regardless of their statistical significance according to the Wald test. Further, no formal adjustments were made for multiple testing across the seven performance and composite measures; instead, we emphasized repeated patterns of observed effect sizes. We checked for multi-collinearity between the percent poverty covariate and the geographic indicator using Tolerance = $1/\text{VIF} < .1$ as a criterion. Statistical significance was considered to be at the standard $p < .05$.

Effect sizes were estimated using absolute differences in adjusted means. Adjusted means were calculated as the estimated mean value for the dependent variable at a value of interest for an independent variable (e.g., for physician age = 40 or affiliation = group), holding other independent variables constant at their mean. Magnitude of change in performance scores across physician age was estimated using adjusted mean performance score for the beginning of each decade between 40 and 70, as well as for ages 35 and 75. Further, we calculated standardized Cohen's d effect size measures to enable comparison of effect sizes for composite rate versus composite rank, given they are measured on different scales.^{10, 11}

Sensitivity analyses considered the following: (1) the effects on absolute HEDIS performance scores to facilitate interpretation of clinical significance, (2) the impact of focusing on physicians with at least 3 out of the 7 and at least 4 out of the 7 individual contributing measures (sample size was too restricted by requiring scores on at least 5 or 6 measures), and (3) the effect of using an opportunity-based weights method¹² to create the composite performance score. In this method, each measure is weighted by the number of opportunities to get a score on that measure divided by the sum of the opportunities for all measures.

RESULTS

Study Sample

Table 1 shows the distribution of physician characteristics. A little over 25% of PCPs were 49 and under; a little over 30% were age 50–59 and 60–69; and just under 10% were age 70 and above. Twenty-six percent of PCPs were affiliated only with a solo practice; 57% only with a group practice; and 17% with both solo and group practices. Fifty-three percent were trained in internal medicine and 47% in family medicine. Just over a quarter of the PCPs participated in an ACA exchange. Forty-eight percent practiced in southern California; 35% in central, northern, and rural areas of the state; and 17% in the San Francisco Bay area. The mean number of patients eligible

Table 1 Descriptive Statistics for Study Sample

Physician characteristic	N*	%
Age		
< 40 years	125	3
40–49 years	1156	25
50–59 years	1508	32
60–69 years	1439	31
70+ years	426	9
Specialty		
Family practice	2638	47
Internal medicine	2930	53
Affiliation		
Solo	1285	26
Group	2822	57
Solo + group	824	17
Region		
Bay area	945	17
Central, rural, Northern CA	1976	35
Southern CA	2647	48
Participates in covered CA		
Yes	1452	26
No	4116	74

*Total N differs across characteristics due to missing values

for an individual measure varied from 53 to 207 (s.d. 33 to 68), with a minimum of 18 and maximum of 713.

Association Between Performance and PCP Characteristics

Results of regression analyses are shown in Table 2. The entries in the cells are the adjusted means estimated from the regression analyses on the percentile rank performance scores and are on a scale of 0 to 100 percentile points. Results of analyses conducted on the absolute HEDIS scores, expressed as percentages, are shown in Table A1.

Composite Scores (Rank). For the composite measure of rank performance scores, there was a statistically significant negative association with physician age, specialty (internal vs. family), affiliation (group vs. solo), and region (southern California vs. Bay area vs. central/north); association between performance and participation in ACA exchange programs and poverty were not statistically significant. The age effect size was 13 percentile points for PCPs age 35 versus 75 (60th percentile vs. 47th percentile; $p < .001$). The effect for group versus solo affiliation was 5 percentile points (55th vs. 50th percentile; $p < .001$); 3 percentile points for internal versus family medicine (54th vs. 51st percentile; $p < .001$); and 11 percentile points for geographic region (58th for southern CA vs. 47th percentile for central/north CA; $p < .001$).

Individual Scores (Rank). Table 2 shows that four of the individual HEDIS measures had statistically significant associations with PCP age, with effect sizes ranging from 8 to 18 percentile points. Four measures were significantly associated with practice affiliation, with effect sizes ranging from 2 to 6 percentile points. Three measures were

Table 2 Regression Results for Association Between PCP Relative Quality Performance Score (Rank of HEDIS score) and Physician Characteristics

	Composite	HbA1C	Att Nephrop	Appr. HTN	BSC 50–69	ACE ARB	Diuretic	Prop statin
	N = 4550	N = 2769	N = 3333	N = 103	N = 2383	N = 3443	N = 1719	N = 750
IV in final model	Adj mean	Adj mean	Adj mean	Adj mean	Adj mean	Adj mean	Adj mean	Adj mean
Physician age								
35 years	60	55	59	45	52	60	58	47
40 years	58	54	57	46	51	58	57	47
50 years	55	52	53	50	51	53	54	48
60 years	52	50	49	53	50	48	50	50
70 years	48	48	45	56	49	44	47	51
75 years	47	47	43	58	49	42	46	51
	$p < .001$	$p < .001$	$p < .001$	$p = .32$	$p = .20$	$p < .001$	$p < .001$	$p = .36$
Specialty								
Family	51	51	51	51	46	47	49	49
Internal	54	51	49	53	54	53	54	49
	$p < .001$	$p = .97$	$p = .06$	$p = .67$	$p < .001$	$p < .001$	$p < .001$	$p = .78$
Affiliation								
Solo	50	48	46	50	46	49	51	48
Group	55	54	54	54	54	51	53	51
Solo + group	48	45	45	49	42	47	48	45
	$p < .001$	$p < .001$	$p < .001$	$p = .76$	$p < .001$	$p < .001$	$p = .06$	$p = .08$
Region (CA)								
Bay area	51	49	51	66	49	42	51	55
Central/north	47	45	43	44	49	44	43	50
Southern CA	58	55	55	53	51	57	59	46
	$p < .001$	$p < .001$	$p < .001$	$p = .05$	$p = .07$	$p < .001$	$p < .001$	$p = .02$
ACA exchange								
Yes	51	48	49	54	51	49	52	49
No	53	51	50	51	50	50	51	49
	$p = .09$	$p = .01$	$p = .26$	$p = .76$	$p = .35$	$p = .35$	$p = .75$	$p = .92$

* p values based on the Wald t test or overall F test for significance of variable in model

significantly associated with specialty, with effect sizes ranging from 5 to 8 percentile point differences. All measures except breast cancer screening were significantly associated with geographic region, with effect sizes between 9 and 22 percentile points.

Interaction Effects. Only group practice moderated the effect of physician age on rank performance scores. Table 3 shows that while age had a statistically significant association within both group and solo PCPs, the effect size was 17 percentile points within group practitioners and only 6 percentile points

within solo practitioners. There was no significant interaction between physician age and geographic region.

Sensitivity Analyses

In the first sensitivity analysis, significance of results for composite absolute HEDIS scores mirrored those for composite rank scores. The biggest difference between absolute and rank effect sizes occurred for physician age, with $d = 0.18$ for the absolute measure and $d = 0.27$ for the rank measure; both are within the small effect size range (near $d = 0.2$).^{10, 11} Regression-adjusted means for the absolute HEDIS scores are shown in Appendix Tables A1 and A2.

In the second sensitivity analysis, results for the composites requiring at least 3 and at least 4 of the 7 individual measures were very close to those for the unrestricted composite rank, with effect sizes changing by no more than 3 percentile points. Finally, in the third sensitivity analysis, using an opportunity weights composite, effect sizes changed by no more than 1 percentile point.

DISCUSSION

The relative performance of PCPs on HEDIS scores systematically declined with age and was slightly lower for solo versus group-affiliated PCPs, for family versus internal

Table 3 Interaction Effect of Age by Affiliation on Composite Rank Performance Score

	Solo	Group	Solo + group
	N = 1232	N = 2528	N = 793
	Adj mean	Adj mean	Adj mean
Physician age			
35	53	64	52
40	52	62	51
50	50	58	49
60	49	54	47
70	47	50	45
75	47	47	44
	$p = .03$	$p < .001$	$p = .02$

* p values based on the Wald t test or overall F test for significance of variable in model

medicine PCPs, and for PCPs practicing in the more rural, central, and northern parts of California than in the Bay area and Southern California. Physician age has implications for policies, including mandatory retirement ages and additional required training. A recent study of hospitalists and general internists treating Medicare patients in acute care hospitals showed a negative association between physician age and quality as measured by 30-day mortality and readmission rates, although the relationship did not hold for physicians with high volume.³ On the other hand, another large study found no relationship between age of physician and quality outcomes, as measured by the RAND Quality Assessment Tool, across all subspecialties examined.⁴ One review considering a range of quality measures concluded there was a negative association between physician age and quality, but urged further study, especially as newer generations of physicians become older.¹³

Although our findings were statistically significant, we caution against overinterpretation. The numbers reported in Tables 2 and 3 represent relative performance. Analyses of absolute HEDIS performance scores (Appendix Tables A1 and A2) show much smaller effect sizes when measured by differences in adjusted mean scores representing the absolute percentage of a PCP's attributed patients who received the HEDIS recommended process of care. For example, the 13 percentile point difference between relative performance scores for 35- versus 75-year-old PCPs corresponded to only a 3 percentage point difference in absolute HEDIS composite scores (89% vs. 86%). The literature does not provide guidance on whether these HEDIS effect sizes should be considered clinically significant. One recent study, examining the relationship between Maintenance of Certification through the American Board of Internal Medicine and performance on a set of five HEDIS measures, found differences in adjusted mean absolute scores of 1.7 to 4.6 percentage points were statistically significant but noted they were unable to make assertions about the clinical significance of these effects.¹⁴ Understanding how incremental changes in HEDIS scores translate to incremental change in patient well-being warrants further research if these measures continue to be used for performance benchmarking.

Variability in PCP performance may be less of a concern when the average HEDIS performance rates are uniformly high. In this study, most providers achieved scores above 80%. This is considerably higher than in the widely discussed 2003 RAND Quality Assessment Tools study.¹⁵ With the exception of the breast cancer screening measure, the HEDIS rates observed in the present study also were higher than those found in the 2016 follow-up to the RAND study.¹⁶ This may be explained in part by the fact that California-based physicians' performance was influenced by a P4P program that included the HEDIS measures from the present study.¹⁷ Thus, the relationships observed in the present study may not apply to settings in which performance on these measures is not incentivized. Further, since the P4P program applied only to HMOs, it may have differentially influenced group practices.

Finally, Table 3 and Appendix Table A2 show that younger physicians did particularly well in group practices. Research into how to leverage the benefits of group practice for older physicians as well may be a fruitful area of focus.

Limitations

Our study is limited by a variety of factors. It focused only on performance according to a small number of claims/encounter-based HEDIS measures in a single US state. Further, the study was based on a single insurer's network within that state. However, according to recent statewide statistics,¹⁸ the PCPs included in this study represented approximately 25% of all PCPs practicing in California. It is possible, but unlikely, that our sample was biased in some way and not reflective of all practicing PCPs in the state.

While all of the payers agreed to a common set of HEDIS measures before contributing data to the CHPI system, there may have been differences in internal performance measurement priorities, incentives, and QI programs within each payer. The influence of these differences was not something we could measure or control for.

Limitations of using zip code rather than census tract data as a proxy for patient-related social determinants of health also have been noted.¹⁹ Additionally, although the CHPI-provided HEDIS measurements were risk-adjusted, there are always limitations with the risk adjustment process; some of the differences in scores may represent factors outside of individual physician control (e.g., patient lack of adherence).

Another limitation is that the study design was cross-sectional and observational, and, as such, correlations observed cannot be used to infer causality. For example, the study cannot address whether the association between the breast cancer screening measure and affiliation with group versus solo practice is due to group structures having better supports for implementing processes for promoting breast cancer screening, or to the selection of solo practitioners by women who are less likely to undertake breast cancer screening. Further, we were unable to control for selection bias of patients towards certain types of physicians.

Finally, the CHPI data are no longer publically available. Previously, CHPI provided a website that allowed anyone to look up a physician's performance. However, CHPI is no longer in operation. Moreover, while they were operational, the data were not posted in a way that researchers could easily analyze. We obtained access to an analytic dataset, to which physician characteristics were added, from one of the large insurers participating in CHPI. Unfortunately, the insurer considers the dataset proprietary and thus not available for public use.

CONCLUSION

In conclusion, we observed statistically significant differences in relative and absolute HEDIS performance scores by age, practice affiliation, speciality training, and region of practice

among California PCPs, but effect sizes for absolute HEDIS scores, representing percent of time recommended care actually was followed, were small (0 to 4 percentage points). There is a need to define a level of HEDIS score performance that signifies high-quality care. We recognize, however, that HEDIS is not a proficiency test, and given the limits of risk adjustment and claims-based quality measures, there may be not be a sufficient evidence base for setting strict thresholds. Given this, it may be time to identify a new set of quality performance measures that are easy to collect, associated with relevant patient outcomes, and sensitive to change.

The findings suggest one avenue for addressing shortages in high-quality primary care population health management is understanding how to help older physicians excel equally well compared with younger physicians in group practices. We found no evidence that primary care physicians choosing to participate in ACA narrow network plans systematically provide lower quality care than physicians choosing not to participate, reducing concerns over the potential effects of narrow networks on the quality of care. Similarly, we found little evidence that training in family versus internal medicine had any meaningful effect on performance of evidence-based care processes.

Acknowledgments: The authors would like to thank the California Healthcare Performance Information System for assembling one of the only multi-payer systems in California to provide validated information on provider quality performance ratings.

Corresponding Author: Jill R. Glassman, PhD, MSW; Clinical Excellence Research Center; School of Medicine Stanford University, 365 Lasuen Street, Stanford, CA 94305, USA (e-mail: jill.r.glassman@stanford.edu).

Compliance with Ethical Standards:

Conflict of Interest: The authors declare that they do not have a conflict of interest.

REFERENCES

- Ketcham JD, Baker LC, MacIsaac D. Physician Practice Size And Variations In Treatments And Outcomes: Evidence From Medicare Patients With AMI. *Health Aff (Millwood)* 2007;26(1):195–205.
- Casalino LP, Ramsay P, Baker LC, Pesko MF, Shortell SM. Medical Group Characteristics and the Cost and Quality of Care for Medicare Beneficiaries. *Health Serv Res* 2018;53(6):4970–96.
- Tsugawa Y, Newhouse JP, Zaslavsky AM, Blumenthal DM, Jena AB. Physician age and outcomes in elderly patients in hospital in the US: observational study. *BMJ*. 2017 16;j1797.
- Reid RO, Friedberg MW, Adams JL, McGlynn EA, Mehrotra A. Associations Between Physician Characteristics and Quality of Care. *Arch Intern Med*. [Internet]. 2010 13 [cited 2019 Jul 12];170(16). Available from: <http://archinte.jamanetwork.com/article.aspx?doi=10.1001/archinternmed.2010.307>
- Caballero AE, Murray R, Delbanco SF. Are Limited Networks What We Hope And Think They Are? 2018 Feb 12 [cited 2019 May 1]. Available from: <https://www.healthaffairs.org/doi/10.1377/hblog20180208.408967/full/>
- Shackelton-Piccolo R, McKinlay JB, Marceau LD, Goroll AH, Link CL. Differences Between Internists and Family Practitioners in the Diagnosis and Management of the Same Patient With Coronary Heart Disease. *Med Care Res Rev* 2011;68(6):650–66.
- Zoberi KA, Salas J, Morgan CN, Scherrer JF. Comparison of Family Medicine and General Internal Medicine on Diabetes Management. *Mo Med* 2017;114(3):187–94.
- Higgins A, Zeddies T, Pearson SD. Measuring The Performance Of Individual Physicians By Collecting Data From Multiple Health Plans: The Results Of A Two-State Test. *Health Aff (Millwood)* 2011;30(4):673–81.
- California Healthcare Performance Information System (CHPI): Methods for Rating Physicians and Practice Sites – Cycle 2 Prepared by: California Healthcare Performance Information System; 2016 Nov.
- Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers; 1988.
- Sedgwick P. Randomised controlled trials: understanding effect sizes. *BMJ*. 2015;350(mar27 2):h1690–h1690.
- Shwartz M, Restuccia JD, Rosen AK. Composite Measures of Health Care Provider Performance: A Description of Approaches: Composite Measures of Health Care Provider Performance. *Milbank Q* 2015;93(4):788–825.
- Choudhry NK, Fletcher RH, Soumerai SB. Systematic Review: The Relationship between Clinical Experience and Quality of Health Care. *Ann Intern Med* 2005;142(4):260–73.
- Gray B, Vandergrift J, Landon B, Reschovsky J, Lipner R. Associations Between American Board of Internal Medicine Maintenance of Certification Status and Performance on a Set of Healthcare Effectiveness Data and Information Set (HEDIS) Process Measures. *Ann Intern Med* 2018;169(2):97.
- McGlynn EA, Asch SM, Adams J, Keesey J, Hicks J, DeCristofaro A, et al. The Quality of Health Care Delivered to Adults in the United States. *N Engl J Med* 2003;348(26):2635–45.
- Levine DM, Linder JA, Landon BE. The Quality of Outpatient Care Delivered to Adults in the United States, 2002 to 2013. *JAMA Intern Med* 2016;176(12):1778.
- Yegian J, Yanagihara D. Value Based Pay for Performance in California [Internet]. Integrated Healthcare Association; 2013 [cited 2019 May 16]. Available from: <https://www.ihc.org/sites/default/files/resources/issue-brief-value-based-p4p-2013.pdf>
- Gaines R. California Maps: How Many Primary Care and Specialist Physicians Are in Your County? [Internet]. 2017 [cited 2019 May 16]. Available from: <https://www.chcf.org/publication/california-maps-primary-care-specialist-physicians-county/>
- Knighton AJ, Stephenson B, Savitz LA. Measuring the Effect of Social Determinants on Patient Outcomes: A Systematic Literature Review. *J Health Care Poor Underserved* 2018;29(1):81–106.

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