Original Paper

Web-Based Physician Ratings for California Physicians on Probation

Gregory P Murphy^{1*}, MD; Mohannad A Awad^{1,2*}, MD; E Charles Osterberg^{3*}, MD; Thomas W Gaither^{4*}, BS; Thanabhudee Chumnarnsongkhroh^{1*}, MD; Samuel L Washington^{1*}, MD; Benjamin N Breyer^{1,5}, MD

Corresponding Author:

Benjamin N Breyer, MD Zuckerberg San Francisco General Hospital Department of Urology University of California, San Francisco Suite 3A 1001 Potrero Ave San Francisco, CA, 94110 United States

Phone: 1 4152068805 Fax: 1 4152064499

Email: benjamin.breyer@ucsf.edu

Abstract

Background: Web-based physician ratings systems are a popular tool to help patients evaluate physicians. Websites help patients find information regarding physician licensure, office hours, and disciplinary records along with ratings and reviews. Whether higher patient ratings are associated with higher quality of care is unclear.

Objective: The aim of this study was to characterize the impact of physician probation on consumer ratings by comparing website ratings between doctors on probation against matched controls.

Methods: A retrospective review of data from the Medical Board of California for physicians placed on probation from December 1989 to September 2015 was performed. Violations were categorized into nine types. Nonprobation controls were matched by zip code and specialty with probation cases in a 2:1 ratio using the California Department of Consumer Affairs website. Web-based reviews were recorded from vitals.com, healthgrades.com, and ratemds.com (ratings range from 1-5).

Results: A total of 410 physicians were placed on probation for 866 violations. The mean (standard deviation [SD]) number of ratings per doctor was 5.2 (7.8) for cases and 4 (6.3) for controls (P=.003). The mean rating for physicians on probation was 3.7 (1.6) compared with 4.0 (1.0) for controls when all three rating websites were pooled (P<.001). Violations for medical documentation, incompetence, prescription negligence, and fraud were found to have statistically significant lower rating scores. Conversely, scores for professionalism, drugs or alcohol, crime, sexual misconduct, and personal illness were similar between cases and controls. In a univariate analysis, probation was found to be associated with lower rating, odds ratio=1.5 (95% CI 1.0-2.2). This association was not significant in a multivariate model when we included age and gender.

Conclusions: Web-based physician ratings were lower for doctors on probation indicating that patients may perceive a difference. Despite these statistical findings, the absolute difference was quite small. Physician rating websites have utility but are imperfect proxies for competence. Further research on physician Web-based ratings is warranted to understand what they measure and how they are associated with quality.

(J Med Internet Res 2017;19(8):e254) doi: 10.2196/jmir.7488



¹Zuckerberg San Francisco General Hospital, Department of Urology, University of California, San Francisco, San Francisco, CA, United States

²Department of Surgery, King Abdulaziz University, Rabigh, Saudi Arabia

³Dell Medical School, Department of Surgery, University of Texas, Austin, TX, United States

⁴School of Medicine, Department of Urology, University of California, San Francisco, San Francisco, CA, United States

⁵Department of Biostatistics and Epidemiology, University of California, San Francisco, San Francisco, CA, United States

^{*}these authors contributed equally

KEYWORDS

online physician ratings; probation; Internet; quality of care

Introduction

Web-based physician ratings systems are becoming an increasingly popular tool to help patients choose their hospitals and providers [1]. Physician rating websites contain information regarding physician licensure, office hours, and disciplinary records in addition to ratings and reviews that are helpful to health care consumers [2]. Studies show mixed results as to whether highly rated hospitals or physicians deliver superior care [3,4]. Okike et al found no correlation between Web-based ratings for cardiac surgeons and 30-day mortality following coronary artery bypass grafting [5]. Regarding hospitals, a review of social media and rating site literature indicated a relationship between improved hospital ratings and better mortality and infection rates [6]. A systematic review by Doyle et al suggested a positive correlation between higher ratings and patient safety and clinical effectiveness [7]. Similarly, Greaves et al demonstrated that Web-based hospital ratings are associated with improved mortality and infection rates [8]. Higher Yelp ratings were also associated with higher Hospital Consumer Assessment of Healthcare Providers and Systems ratings [9]. Regarding social media, Facebook ratings have been correlated to lower hospital readmission rates, whereas Twitter comments were not associated with quality metrics [10,11].

When studying the quality of physician care, many different outcome metrics have been used with differing strengths. These outcomes can be influenced by a multitude of factors, some directly related to physician skill whereas others, such as patient health, are outside of the doctor's control. Physician ratings have been correlated with such diverse outcomes as board certification, education, malpractice claims, mortality, infection, and readmission rates [2-12]. Though it can be difficult to equate physician competence with outcomes, doctors who violate codes of conduct and fail to meet the standard of care are placed on probation by their state medical board after a review of the evidence. It is unknown how patients perceive and rate physicians on probation as it has never been used as a quality metric. We sought to determine whether patients rated probated providers differently than physicians in good standing by comparing Web-based physician ratings from three consumer rating websites. We hypothesize that physicians on probation will have similar ratings to nonprobation controls.

Methods

We retrospectively reviewed publically available data from the Medical Board of California for physicians who were placed on probation from December 1989 to September 2015.

Rationales for probation were independently categorized into nine types of infractions by five independent reviewers (MAA, TWG, TC, SLW, and GPM). After reviewing all infractions, we used an inductive approach to create nine probation categories [13]. If there were any questions as to the appropriate categorization by an individual reviewer, this was brought to the group and a consensus decision was made. Nonprobation controls were matched by zip code and specialty with probation cases in a 2:1 ratio using the California Department of Consumer Affairs website. Web-based reviews were recorded from vitals.com, healthgrades.com, and ratemds.com. Ratings on these websites ranged from 1 to 5 and were weighted by the number of ratings so as not to overemphasize a small number of outlier ratings.

Statistical analysis was performed with STATA version 14 (College Station, TX). Parametric and nonparametric statistics were run despite the nonnormal statistical distribution. The results were similar between statistical analyses as the sample size was large [14]. For ease of reporting, parametric tests were reported. Specifically, a student's *t*-test was used to compare mean ratings for violations. A rating below 3 was considered low and logistic univariable and multivariable regression analyses were done to determine predictors for low rating. We determined covariates including age, gender, type of specialty, and type of violation to be included in the multivariable model a priori. All tests were two-sided and *P* values <.05 were considered significant.

Results

A total of 410 physicians (cases) were placed on probation for 866 violations and were matched with 818 controls. The mean (standard deviation [SD]) number of ratings per physician was 5.2 (7.8) for cases and 4.0 (6.3) for controls (P=.003). The mean rating for physicians on probation was 3.7 (SD 1.6) compared with 4.0 (SD 1.0) for controls when all three rating websites were pooled (P<.001). Figure 1 depicts the overall violations using a violin plot that shows the median, interquartile range, and distribution of the ratings for cases and controls.

Tables 1 and 2 show the differences in average weighted ratings between cases and controls for each website stratified by violation type. Violations for medical documentation, incompetence, prescription negligence, and fraud were found to have statistically significant differences in physician ratings between cases and controls. Conversely, physician ratings for cases and controls were not statistically different when the violation pertained to professionalism, drugs or alcohol, crime, sexual misconduct, and personal illness.



Table 1. Mean rating of probation cases and controls by violation type.

Violation	Total (%) ^a	Vitals cases	Vitals controls	P-value	Health Grades cases	Health Grades controls	P-value
Medical records, mean (SD ^b)	176 (42.9)	3.7 (1.41)	4.1 (0.94)	<.001	3.6 (2.1)	3.8 (1.29)	.02
Professionalism, mean (SD)	166 (40.5)	3.9 (1.35)	4.1 (0.82)	.02	3.9 (1.98)	3.9 (1.09)	.81
Incompetence, mean (SD)	161 (39.3)	3.6 (1.46)	4 (0,93)	<.001	3.6 (2.08)	3.8 (1.21)	.08
Prescription negligence, mean (SD)	106 (25.9)	3.9 (1.6)	4 (1.05)	.31	3.8 (2.11)	3.9 (1.23)	.49
Drug or alcohol addiction, mean (SD)	84 (20.5)	4 (1.57)	4.1 (0.88)	.63	3.8 (2.1)	3.9 (1.11)	.66
Committing a crime, mean (SD)	63 (15.4)	3.8 (1.53)	3.8 (0.83)	.91	3.8 (1.87)	3.9 (0.94)	.45
Fraud, mean (SD)	44 (10.7)	3.9 (1.28)	4.1 (0.82)	.11	3.7 (1.81)	3.9 (1.05)	.32
Sexual misconduct or battery, mean (SD)	42 (10.2)	4 (1.27)	4.1 (0.81)	.28	4 (2.02)	4 (1.11)	.86
Personal illness, mean (SD)	24 (5.9)	3.7 (1.32)	4.2 (0.83)	.02	3.6 (2.3)	3.8 (1.09)	.56
All violations, mean (SD)	866 (0)	3.8 (1.47)	4.1 (0.96)	<.001	3.7 (2.03)	3.8 (1.2)	.01

^aSome doctors have more than one violation. The percentage is for number of doctors on probation (cases group) having each violation out of the total number of doctors on probation (410).

Table 2. Mean rating of probation cases and controls by violation type.

Violation	RateMDs cases	RateMDs controls	P-value	Cases in all websites	Controls in all websites	P-value
Medical records, mean (SD ^a)	3.5 (2.06)	3.8 (1.4)	.04	3.5 (1.58)	3.9 (1.04)	<.001
Professionalism, mean (SD)	3.9 (2)	4 (1.14)	.78	3.9 (1.58)	4 (0.87)	.54
Incompetence, mean (SD)	3.3 (2.21)	3.8 (1.43)	.03	3.5 (1.55)	4 (1.01)	<.001
Prescription negligence, mean (SD)	3.8 (2.12)	4 (1.32)	.41	3.8 (1.59)	4 (1.05)	.04
Drug or alcohol addiction, mean (SD)	3.8 (1.93)	3.9 (1.09)	.86	4 (1.78)	4 (0.93)	.64
Committing a crime, mean (SD)	3.7 (2.1)	3.7 (1.05)	.91	3.9 (1.59)	3.8 (0.83)	.62
Fraud, mean (SD)	3.3 (2)	3.6 (1.23)	.35	3.6 (1.37)	3.9 (0.89)	.03
Sexual misconduct or battery, mean (SD)	3.5 (1.83)	3.8 (1.11)	.26	3.9 (1.53)	4 (0.94)	.59
Personal illness, mean (SD)	3.4 (1.65)	3.4 (1.32)	.95	3.6 (1.33)	3.9 (0.8)	.17
All violations, mean (SD)	3.5 (2)	3.9 (1.31)	.001	3.7 (1.56)	4 (1.01)	<.001

^aSD: standard deviation.

In the univariable analysis, probation was found to be associated with lower rating, odds ratio=1.5 (95% CI 1.001-2.2). This association was not significant in a multivariable model when we included age and gender, odds ratio=1.4 (95% CI 0.9-2.2). In addition, age, gender, type of specialty, and type of violation all did not predict a low rating.

Healthgrades.com included the probation status of the physician on their website in 328 of 389 (84%) cases in our cohort, yet

there was no significant difference in mean ratings for doctors who had their violations published compared with those who did not. A sensitivity analysis was done to determine if timing of probation effected ratings. The majority of our violations are from the recent past. When we excluded violations before 2005, only 17 physicians were excluded and no difference in results was seen.



^bSD: standard deviation.

healthgrades.com

Figure 1. Violin plot showing ratings of physicians on probation versus controls by website.

Discussion

Principal Findings

Web-based physician ratings from three websites were lower for doctors on probation, indicating that patients perceive a difference in physician performance. Despite this difference being statistically significant, the absolute difference is quite small with overlapping rating distributions. The vast majority of reviews were positive with small numbers of reviews per doctor. This is similar to other studies, making it difficult to draw strong conclusions about Web-based rating utility [6,15].

vitals.com

When categorizing the violations into subtypes, patients rate physicians lower when the probation infringements correspond to infractions for medical documentation, incompetence, prescription negligence, and fraud. In contrast, doctors on probation for professionalism, drugs or alcohol, sexual abuse, or personal illness received statistically similar ratings to controls. This finding may reflect what websites are measuring: patient perception and consumer experience, which is influenced mostly by bedside manner, wait times, and staffing issues [16]. Physicians with violations for medical documentation, for example, may run higher volume clinics, which negatively influences the patient experience. A German study showed higher physician ratings correlated to lower number of patients in the practice more than quality metrics [17].

Comparison With Prior Work

The relationship between physician skills, Web-based ratings and outcome metrics is complex. Polled physicians are concerned about the usefulness of Web-based ratings and how they could negatively impact their practice [18]. Comparing

websites scores of individual sports medicine doctors revealed a low degree of correlation between websites raising concerns about their reliability [19]. When reviews were correlated with scores, very high and very low ratings were correlated with the patient's perception of physician skill and quality, which would seem difficult for patients to truly know [20]. Similarly, Web-based reviews of hand surgeons showed very positive reviews were related to perceived competence [21]. This raises concerns among physicians as to the accuracy and potential harm of negative reviews [6].

All websites

ratemds.com

For providers, it is an uncomfortable position to be publically judged by patients, which could significantly damage a doctor's reputation and practice. Some physicians have responded with indifference or disdain for these reviews [12,22]. Others have embraced these ratings and made an effort to alter their behavior and practice to improve patient experience and scores [21]. When selecting a surgeon Web-based, polled patients listed Web-based reviews as a minor factor with insurance, office location, and hospital reputation as more important [23]. It seems most patients understand the limitations of Web-based reviews.

There are ethical pitfalls of Web-based reviews due to the possible financial gain from improved referrals. Most reviews appear genuine but careful study revealed a few anonymous reviews may be from the physicians themselves in an attempt to falsely raise their ratings [2]. Better regulation may be needed to prevent abuses by providers including asking or paying patients for positive reviews [23]. Some concerns have been raised by doctors that competitors or unhappy employees could easily pose as a patient and post a negative review [24].



Despite the negatives, Web-based reviews are likely to increase in importance and patient utilization [1]. Lee suggests Web-based reviews help improve physician-patient relationship by increasing transparency and trust [25]. Not only does this help patient decision making but also provides the physician with feedback that can improve patient experience.

Future Research

More research is needed to explore different physician quality metrics to see if there is an association with positive reviews. Determining more precisely what aspects of the health care system patient feedback can measure and improve is important. Prior studies explored many objective criteria such as mortality, infection, or readmission rates [4,6,7-10]. These are important metrics but do not necessarily reflect physician decision making. For example, a surgical site infection may occur despite excellent sterile technique and appropriate guidelines-based antibiotic practice. It raises questions as to whether physicians should be judged on outcomes or more on their decision making. Our study is the first to explore the link between Web-based reviews and probation status. Using probation as a proxy for clinical incompetence seems to be imperfect. When a physician is placed on probation, there are clearly issues to be remedied, but whether this negatively affects clinical care is difficult to determine. Whereas more research is needed to determine if probation is a reliable marker for clinical incompetence, it would seem to be a good indicator for poor physician decision making on some level. In addition, differing violations would seem to affect physician judgment and quality in diverse ways so lumping all probation cases has its limitations. More research is needed into which violations more severely affect physician quality.

Limitations

Our study has limitations. The timing of probation could be problematic as violations from many years ago could correlate poorly with recent reviews. However, a sensitivity analysis excluding 17 cases before 2005 showed similar results. We studied patient reviews of physicians practicing in California, so these results may not be generalizable to reviews from other states. Future research should be directed toward other populations to confirm our findings.

Conclusions

Web-based physician ratings were lower for doctors on probation indicating that patients may perceive a difference. Whereas statistically significant, the absolute difference was quite small. Physician rating websites have utility but are imperfect proxies for competence. Further research on physician Web-based ratings is warranted to understand what they measure and how they are associated with quality.

Acknowledgments

The Alafi Foundation provided funding for this research. Propublica.com provided the data regarding California physicians on probation.

Conflicts of Interest

None declared.

References

- 1. Hanauer DA, Zheng K, Singer DC, Gebremariam A, Davis MM. Public awareness, perception, and use of online physician rating sites. J Am Med Assoc 2014 Feb 19;311(7):734-735. [doi: 10.1001/jama.2013.283194] [Medline: 24549555]
- 2. Lagu T, Hannon NS, Rothberg MB, Lindenauer PK. Patients' evaluations of health care providers in the era of social networking: an analysis of physician-rating websites. J Gen Intern Med 2010 Sep;25(9):942-946 [FREE Full text] [doi: 10.1007/s11606-010-1383-0] [Medline: 20464523]
- 3. Fenton JJ, Jerant AF, Bertakis KD, Franks P. The cost of satisfaction: a national study of patient satisfaction, health care utilization, expenditures, and mortality. Arch Intern Med 2012 Mar 12;172(5):405-411. [doi: 10.1001/archinternmed.2011.1662] [Medline: 22331982]
- 4. Sacks GD, Lawson EH, Dawes AJ, Russell MM, Maggard-Gibbons M, Zingmond DS, et al. Relationship between hospital performance on a patient satisfaction survey and surgical quality. JAMA Surg 2015 Sep;150(9):858-864. [doi: 10.1001/jamasurg.2015.1108] [Medline: 26108091]
- 5. Okike K, Peter-Bibb TK, Xie KC, Okike ON. Association between physician online rating and quality of care. J Med Internet Res 2016 Dec 13;18(12):e324 [FREE Full text] [doi: 10.2196/jmir.6612] [Medline: 27965191]
- 6. Verhoef LM, Van de Belt TH, Engelen LJ, Schoonhoven L, Kool RB. Social media and rating sites as tools to understanding quality of care: a scoping review. J Med Internet Res 2014;16(2):e56 [FREE Full text] [doi: 10.2196/jmir.3024] [Medline: 24566844]
- 7. Doyle C, Lennox L, Bell D. A systematic review of evidence on the links between patient experience and clinical safety and effectiveness. BMJ Open 2013;3(1):231-233 [FREE Full text] [doi: 10.1136/bmjopen-2012-001570] [Medline: 23293244]
- 8. Greaves F, Pape UJ, King D, Darzi A, Majeed A, Wachter RM, et al. Associations between web-based patient ratings and objective measures of hospital quality. Arch Intern Med 2012 Mar 12;172(5):435-436. [doi: 10.1001/archinternmed.2011.1675] [Medline: 22331980]



- 9. Bardach NS, Asteria-Peñaloza R, Boscardin WJ, Dudley RA. The relationship between commercial website ratings and traditional hospital performance measures in the USA. BMJ Qual Saf 2013 Mar;22(3):194-202 [FREE Full text] [doi: 10.1136/bmjqs-2012-001360] [Medline: 23178860]
- 10. Glover M, Khalilzadeh O, Choy G, Prabhakar AM, Pandharipande PV, Gazelle GS. Hospital evaluations by social media: a comparative analysis of Facebook ratings among performance outliers. J Gen Intern Med 2015 Oct;30(10):1440-1446 [FREE Full text] [doi: 10.1007/s11606-015-3236-3] [Medline: 25749881]
- 11. Greaves F, Laverty A, Cano D, Moilanen K, Pulman S, Darzi A, et al. Tweets about hospital quality: a mixed methods study. BMJ Qual Saf Oct 2014;23(10):838-846. [doi: 10.1136/bmjqs-2014-002875] [Medline: 24748372]
- 12. Gao GG, McCullough JS, Agarwal R, Jha AK. A changing landscape of physician quality reporting: analysis of patients' online ratings of their physicians over a 5-year period. J Med Internet Res 2012;14(1):e38 [FREE Full text] [doi: 10.2196/jmir.2003] [Medline: 22366336]
- 13. Bolton JW. Varieties of clinical reasoning. J Eval Clin Pract 2015 Jun;21(3):486-489. [doi: 10.1111/jep.12309] [Medline: 25652845]
- 14. Fagerland MW. t-tests, non-parametric tests, and large studies--a paradox of statistical practice? BMC Med Res Methodol 2012 Jun 14;12:78 [FREE Full text] [doi: 10.1186/1471-2288-12-78] [Medline: 22697476]
- 15. Riemer C, Doctor M, Dellavalle RP. Analysis of online ratings of dermatologists. JAMA Dermatol 2016 Feb;152(2):218-219. [doi: 10.1001/jamadermatol.2015.4991] [Medline: 26677133]
- 16. López A, Detz A, Ratanawongsa N, Sarkar U. What patients say about their doctors online: a qualitative content analysis. J Gen Intern Med 2012 Jun;27(6):685-692 [FREE Full text] [doi: 10.1007/s11606-011-1958-4] [Medline: 22215270]
- 17. Emmert M, Adelhardt T, Sander U, Wambach V, Lindenthal J. A cross-sectional study assessing the association between online ratings and structural and quality of care measures: results from two German physician rating websites. BMC Health Serv Res 2015;15(1):414 [FREE Full text] [doi: 10.1186/s12913-015-1051-5] [Medline: 26404452]
- 18. Patel S, Cain R, Neailey K, Hooberman L. General practitioners' concerns about online patient feedback: findings from a descriptive exploratory qualitative study in England. J Med Internet Res 2015 Dec 08;17(12):e276 [FREE Full text] [doi: 10.2196/jmir.4989] [Medline: 26681299]
- 19. Nwachukwu BU, Adjei J, Trehan SK, Chang B, Amoo-Achampong K, Nguyen JT, et al. Rating a sports medicine surgeon's "quality" in the modern era: an analysis of popular physician online rating websites. HSS J 2016 Oct;12(3):272-277. [doi: 10.1007/s11420-016-9520-x] [Medline: 27703422]
- 20. Trehan SK, DeFrancesco CJ, Nguyen JT, Charalel RA, Daluiski A. Online patient ratings of hand surgeons. J Hand Surg Am 2016 Jan;41(1):98-103. [doi: 10.1016/j.jhsa.2015.10.006] [Medline: 26710742]
- 21. Emmert M, Meszmer N, Sander U. Do health care providers use online patient ratings to improve the quality of care? Results from an online-based cross-sectional study. J Med Internet Res 2016 Sep 19;18(9):e254 [FREE Full text] [doi: 10.2196/jmir.5889] [Medline: 27644135]
- 22. Ruan X, Bordelon GJ, Kaye AD. Online reviews of physicians. JAMA Dermatol 2016 Mar;152(3):350. [doi: 10.1001/jamadermatol.2015.6164] [Medline: 26964056]
- 23. Ziemba JB, Allaf ME, Haldeman D. Consumer preferences and online comparison tools used to select a surgeon. JAMA Surg 2017 Jan 04;152(4):410-411. [doi: 10.1001/jamasurg.2016.4993] [Medline: 28052162]
- 24. Samora JB, Lifchez SD, Blazar PE, American Society for Surgery of the Hand Ethics and Professionalism Committee. Physician-rating web sites: ethical implications. J Hand Surg Am 2016 Jan;41(1):104-10.e1. [doi: 10.1016/j.jhsa.2015.05.034] [Medline: 26304734]
- 25. Lee V. Transparency and trust online patient reviews of physicians. N Engl J Med 2017 Jan 19;376(3):197-199. [doi: 10.1056/NEJMp1610136] [Medline: 28099823]

Abbreviations

SD: standard deviation

Edited by T Kool; submitted 10.02.17; peer-reviewed by T van de Belt, M Emmert; comments to author 11.03.17; revised version received 09.04.17; accepted 11.05.17; published 22.08.17

<u>Please cite as:</u>

Murphy GP, Awad MA, Osterberg EC, Gaither TW, Chumnarnsongkhroh T, Washington SL, Breyer BN Web-Based Physician Ratings for California Physicians on Probation

J Med Internet Res 2017;19(8):e254 URL: <u>http://www.jmir.org/2017/8/e254/</u>

doi: <u>10.2196/jmir.7488</u> PMID: <u>28830852</u>



JOURNAL OF MEDICAL INTERNET RESEARCH

Murphy et al

©Gregory P Murphy, Mohannad A Awad, E Charles Osterberg, Thomas W Gaither, Thanabhudee Chumnarnsongkhroh, Samuel L Washington, Benjamin N Breyer. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 22.08.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

