Patient Satisfaction of General Dermatologists: A Quantitative and Qualitative Analysis of 38,008 Online Reviews by Gender and Years of Experience

Megan H. Trager[1,4], Dawn Queen[2,4], Weijia Fan[3] and Faramarz H. Samie[2]

Online reviews are the newest method for patients to evaluate their providers. However, insufficient studies focus on the role of inherent physician characteristics, such as gender and years of experience, on patient satisfaction. We analyzed both quantitative and qualitative online reviews of 350 general dermatology providers at 121 Accreditation Council for Graduate Medical Education—accredited dermatology programs across the country to determine the effect of gender and years of experience. There were 38,008 online reviews of general dermatology providers. There was no significant difference in male and female overall ratings. Ratings were overall equally positive for both genders. Female providers were more likely to have positive written comments regarding time spent with patients ($P = 0.027$). New providers received highest overall, promptness, and time spent with patient ratings ($P < 0.001$). Medium experience providers received highest scores in bedside manner ($P < 0.001$), accurate diagnosis ($P = 0.018$), and ability to answer questions ($P = 0.005$). Advanced providers scored the lowest across all categories. In conclusion, gender did not significantly affect ratings, although females received more positive written comments on time spent with patients. Years of experience, however, is a significant factor in patient ratings, with new or medium experience providers scoring higher than advanced providers in every category.

INTRODUCTION

Online ratings have become the newest method for patient evaluation of physician performance. A number of subjective factors have been shown to affect ratings, including bedside manner, perceived experience, and ability to answer questions and communicate clearly (Queen et al., 2021). However, few studies have addressed the effect of inherent provider characteristics, such as gender and years of experience, on online ratings.

Indeed, the role of gender, if any, on ratings has yet to be elucidated. One study previously noted that female doctors across all specialties are more likely to be assigned negative descriptors, including lack of candor, amicability, and disrespect (Dunivin et al., 2020). In contrast, another study of

outpatient medicine providers found that patients of female physicians were more satisfied (Bertakis et al., 2003). Similarly, among rheumatologists, a specialty similar to dermatology, female gender also positively impacted satisfaction (Ku et al., 2015).

The influence of years of experience on ratings must also be clarified. In the Ku et al. (2015) study, more years of experience was associated with higher satisfaction scores. Likewise, another multispecialty study found that older age was most influential on positive satisfaction (Chen et al., 2017). In contrast, in a study by Gao et al. (2012) across multiple specialties, new physicians (<12 years in practice) had significantly higher ratings than advanced physicians.

Even fewer studies have been published focusing on the field of dermatology specifically. A recent study by Waqas et al. (2020) found no association between ratings and gender but did find greater satisfaction with newer dermatologists. To further explore these issues, we present an extensive analysis of 38,008 online reviews of 350 general dermatology providers from 121 Accreditation Council for Graduate Medical Education—accredited academic dermatology programs across the country. Expanding on our previous study (Queen et al., 2021), we analyze the reviews to examine whether differences exist in patient ratings by gender and years of experience for general dermatology providers.

RESULTS

There was a total of 205 female physicians (59%) and 145 male physicians (41%) who were reviewed online in this

[1]Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, New York, USA; [2]Department of Dermatology, Vagelos College of Physicians and Surgeons, Columbia University Irving Medical Center, New York, New York, USA; and [3]Department of Biostatistics, Mailman School of Public Health, Columbia University Irving Medical Center, New York, New York, USA

[4]These authors contributed equally to this work.

Correspondence: Faramarz H. Samie, Department of Dermatology, Vagelos College of Physicians and Surgeons, Columbia University Irving Medical Center, 12th Floor, Herbert Irving Pavilion, 161 Fort Washington Avenue, New York, New York 10032, USA. E-mail: fs2614@cumc.columbia.edu

Cite this article as: JID Innovations 2022;2:100089

Table 1. Quantitative Reviews Separated by Gender

Category	Female (n = 205)	Male (n = 145)	Total (N = 350)	P-Value
Overall rating				
Mean (SD)	4.146 (0.730)	4.250 (0.773)	4.189 (0.748)	0.206
Range	1.000−5.000	1.000−5.000	1.000−5.000	
Bedside manner				
Mean (SD)	4.033 (1.047)	4.073 (1.022)	4.050 (1.035)	0.740
Range	1.000−5.000	1.500−5.000	1.000−5.000	
Answered questions				
Mean (SD)	4.115 (1.031)	4.260 (0.900)	4.177 (0.978)	0.207
Range	1.000−5.000	1.000−5.000	1.000−5.000	
Accurate diagnosis				
Mean (SD)	4.016 (0.968)	4.079 (0.985)	4.044 (0.974)	0.596
Range	1.000−5.000	1.000−5.000	1.000−5.000	
Promptness				
Mean (SD)	3.916 (0.942)	3.954 (1.013)	3.932 (0.972)	0.750
Range	1.000−5.000	1.000−5.000	1.000−5.000	
Time spent with patients				
Mean (SD)	3.906 (1.009)	3.913 (1.066)	3.909 (1.033)	0.958
Range	1.000−5.000	1.000−5.000	1.000−5.000	
Years of experience				
Mean (SD)	20.365 (11.469)	26.583 (14.840)	22.968 (13.325)	**<0.001**
Range	3.000−54.000	3.000−60.000	3.000−60.000	

Bold font indicates statistical significance.

sample. Quantitative reviews (scores) of general dermatology providers totaled 28,216 (4,466 Vitals, 13,899 US News, 4,044 WebMD, 1,392 Google, 4,415 Healthgrades). Qualitative reviews (written) totaled 9,792 comments.

When comparing the quantitative ratings between genders, none of the six categories were significantly different between males and females (Table 1). We then analyzed males and females separately to determine whether there were any differences in ratings among the six quantitative categories. Pair-wise comparisons between any two measurements found that average overall rating was significantly higher than both promptness (n = 154, 4.13 vs. 3.92, P = 0.01) and time spent with patients (n = 154, 4.13 vs. 3.91, P < 0.006). Answered questions ratings were higher than promptness (n = 136, 4.14 vs. 3.88, P = 0.031) and time with patients (P < 0.01), and accurate diagnosis was higher than time spent with patients (n = 137, 4.13 vs. 3.85, P = 0.007). Similarly, within males, average overall rating was higher than promptness (n = 119, 4.19 vs. 3.95, P = 0.025) and time spent with patients (n = 120, 4.17 vs. 3.91, P = 0.002). Bedside manner was higher than answered questions (n = 114, 4.05 vs. 4.25, P = 0.03), answered questions was higher than both promptness (n = 110, 4.23 vs. 3.89, P < 0.001) and time spent with patients (n = 111, 4.20 vs. 3.87, P < 0.001), and accurate diagnosis was higher than time spent with patients (n = 120, 4.08 vs. 3.91, P = 0.02).

Next, we compared quantitative ratings by years of experience. Overall, 59 physicians had 6−10 years of experience (new physicians), 122 had 11−20 years of experience (medium experience physicians), and 155 had over 20 years of experience (advanced physicians), ranging from 21 to 60 years (median 34 years). Males had significantly higher years of experience than females (26.58 vs. 20.37 years, P < 0.001). There were significant differences in all categories

based on years of experience (Table 2). For overall rating, promptness, and time spent with patients, rating was highest for those with 6−10 years of experience and lowest for those with over 20 years of experience. For bedside manner, diagnosis, and answered questions, those with 11−20 years of experience had the highest ratings and those with over 20 years of experience had the lowest rating. In all categories, those with over 20 years of experience had the lowest rating.

There was a total of 9,792 written qualitative comments. More comments were regarding female dermatologists than male dermatologists, with 5,538 (57%) comments pertaining to female dermatologists and 4,254 to males (43%). When comparing the number of positive comments for each gender, females and males had a similar percentage of positive comments (78% positive for females vs. 76% positive for males, P = 0.117). Table 3 shows the percentages of comments and ORs regarding female and male physicians for each qualitative category. The majority of comments were regarding bedside manner and perceived experience. The OR of having a positive comment for time spent with patients was significantly higher in females (OR = 1.42, P = 0.027), but there were no differences in OR between females and males for any of the other categories.

DISCUSSION

Online reviews are an evolving tool for evaluating physicians and are influenced by a number of subjective factors as well as inherent physician characteristics. Other studies have provided contradictory evidence on the role of gender and years of experience on these ratings. Here, we build on our previous study, which demonstrated the importance of bedside manner and perceived experience of the physician on ratings, and further analyze the role of gender and years of experience (Queen et al., 2021).

Table 2. Quantitative Reviews Separated by Years of Experience

Category	Advanced Providers >20 (n = 155)	Medium Experience Providers 11–20 (n = 122)	New Providers 6–10 (n = 59)	All Providers Total (n = 336)	*P*-Value
Overall rating					
Mean (SD)	4.015 (0.714)	4.302 (0.687)	4.368 (0.792)	4.180 (0.733)	**<0.001**
Range	2.000–5.000	1.000–5.000	2.500–5.000	1.000–5.000	
Bedside manner					
Mean (SD)	3.771 (1.097)	4.336(0.827)	4.167 (1.153)	4.038 (1.041)	**<0.001**
Range	1.000–5.000	1.000–5.000	1.000–5.000	1.000–5.000	
Answered questions					
Mean (SD)	3.961 (0.994)	4.357 (0.867)	4.316 (1.157)	4.159 (0.985)	**0.005**
Range	1.000–5.000	1.000–5.000	1.000–5.000	1.000–5.000	
Accurate diagnosis					
Mean (SD)	3.874 (1.011)	4.231 (0.842)	4.100 (1.155)	4.037 (0.978)	**0.018**
Range	1.000–5.000	1.000–5.000	1.000–5.000	1.000–5.000	
Promptness					
Mean (SD)	3.694 (0.983)	4.131 (0.807)	4.267 (1.172)	3.927 (0.969)	**<0.001**
Range	1.000–5.000	2.000–5.000	1.000–5.000	1.000–5.000	
Time spent with patients					
Mean (SD)	3.659 (1.059)	4.135 (0.928)	4.207 (1.013)	3.903 (1.031)	**<0.001**
Range	1.000–5.000	1.000–5.000	1.000–5.000	1.000–5.000	

Bold font indicates statistical significance.

More qualitative comments in our study were directed at female dermatologists (57%) than male dermatologists (43%), but this closely reflects our study population (59% female, 41% male) and suggests no gender bias by reviewers. Unlike in other studies, we did not find that women were more likely to receive negative comments than men (Dunivin et al., 2020). Indeed, both genders had similarly positive comments (78% for females, 76% for males).

Furthermore, with regard to quantitative ratings, we found no significant difference between female and male overall ratings. Likewise, we found the same breakdown of importance for various other factors whether it was a male or female provider. For both genders, overall rating, ability to answer questions, and accurate diagnosis were consistently rated higher than promptness and time spent with patients. Two differences did arise between genders: first, women were significantly more likely than men to receive a positive qualitative comment in the time spent with patient category, and second, for male practitioners, bedside manner was rated significantly lower than their ability to answer questions. Of

Table 3. Qualitative Reviews Separated by Gender

Category	Comment Type	Female		Male		OR (95% CI)	*P*-Value
		n	% (95% CI)	n	% (95% CI)		
Bedside manner	Positive	817	14.8 (13.8–15.7)	529	12.4 (11.5–13.5)	1.18 (0.92–1.51)	0.227
	Negative	176	3.2 (2.7–3.7)	134	3.1 (2.7–3.7)		
Perceived experience: NOS	Positive	327	5.9 (5.3–6.6)	258	6.1 (5.4–6.8)	0.93 (0.69–1.26)	0.695
	Negative	139	2.5 (2.1–3.0)	102	2.4 (2.0–2.9)		
Perceived experience: Staff	Positive	361	6.5 (5.9–7.2)	320	7.5 (6.8–8.4)	1.06 (0.74–1.50)	0.827
	Negative	79	1.4 (1.1–1.8)	74	1.7 (1.4–2.2)		
Perceived experience: Physician	Positive	1,120	20.2 (19.2–21.3)	899	21.1 (19.9–22.4)	1.03 (0.85–1.26)	0.788
	Negative	269	4.9 (4.3–5.5)	223	5.2 (4.6–6.0)		
Communication	Positive	480	8.7 (7.9–9.4)	326	7.7 (6.9–8.5)	1.19 (0.87–1.64)	0.309
	Negative	106	1.9 (1.6–2.3)	86	2.0 (1.6–2.5)		
Finance	Positive	31	0.6 (0.4–0.8)	23	0.5 (0.4–0.8)	1.20 (0.65–2.22)	0.669
	Negative	92	1.7 (1.3–2.0)	82	1.9 (1.5–2.4)		
Wait time	Positive	163	2.9 (2.5–3.4)	141	3.3 (2.8–3.9)	0.88 (0.63–1.22)	0.485
	Negative	153	2.8 (2.4–3.2)	116	2.7 (2.3–3.3)		
Time spent with patients	Positive	418	7.5 (6.9–8.3)	283	6.7 (6.0–7.5)	1.42 (1.05–1.93)	**0.027**
	Negative	113	2.0 (1.7–2.5)	109	2.6 (2.1–2.1)		
Competence	Positive	578	10.4 (9.7–11.3)	462	10.9 (9.9–11.8)	0.94 (0.69–1.27)	0.739
	Negative	116	2.1 (1.7–2.5)	87	2.0 (1.7–2.5)		

Abbreviations: CI, confidence interval; NOS, not otherwise specified; OR, odds ratio.
OR represents the OR of a positive comment for females and 95% CI.
Bold font indicates statistical significance.

note, this has been observed in other specialties, such as primary care, where female physicians were found to spend 15% more time with patients (Ganguli et al., 2020). This finding is relevant because it suggests that patients notice and appreciate this part of their experience and there is in fact a gendered difference in this aspect of patient care.

Interestingly, years of experience did have a significant effect on ratings. Unlike other studies, we divided providers into three groups of years of experience: 6—10 years (new provider), 11—20 years (medium experience), and over 20 years (advanced provider, with a median of 34 years of experience). These year ranges were the same as those used by US News and World Report to categorize physician experience. Overall rating was highest among new providers and lowest among advanced providers. New providers also scored highest in promptness and time spent with patients. Medium experience providers had highest ratings on bedside manner, ability to answer questions, and accurate diagnosis. In contrast, advanced providers had the lowest ratings across all categories. Although there may be multiple reasons for this difference, one possible conclusion is that there has been a shifted focus in younger providers on improving the patient—physician relationship. Additionally, these findings may suggest that overall rating is most influenced by promptness and time spent with patient, which are two categories in which older providers received lower marks. Of note, gender was controlled for in the statistical analysis to prevent confounding of the years of experience analysis. Our results are similar to the study by Gao et al. (2012) showing that younger physicians graduating from medical school within 12 years had higher ratings than older physicians. This has also been observed in other specialties, including orthopedic surgery, where younger surgeons have higher ratings than more experienced surgeons (Damodar et al., 2019; Jack et al., 2018; Nwachukwu et al., 2016). The reason for this finding is likely multifactorial. Damodar et al. (2019) suggest that because recent graduates are less established, their clinics may have fewer patients, allowing them to spend more time with individual patients without sacrificing wait times, leading to more favorable reviews. Our data showing significantly higher ratings of time spent with patients for younger providers supports this theory. However, this discussion is more nuanced because older, more experienced physicians likely do not need the same length of time to make diagnoses as their younger counterparts. Additionally, younger providers may be more cognizant of the power of social media and online reviews and thus cater the patient experience with this knowledge in mind.

This study has several limitations, including inability to assess reviewer biases (such as gender and age) and confirm their identity as patients. As well, we included a number of different subspecialties of dermatology but did not include others, such as Mohs surgery, that include more male practitioners. Although the results are applicable to the field of academic dermatology, they may not be representative of the larger healthcare field. Additionally, written comments may have been counted independently when they are in fact interrelated, as described further in the methods. Despite these limitations, our large sample size of over 38,000 reviews from 121 dermatology programs across the country provides significant power to our conclusions.

With the growth in popularity of online reviews, it is important to understand the inherent biases in ratings of physicians to correctly interpret reviews. This has implications both for patients choosing a physician and for physicians working to improve the satisfaction of their patients. Analysis of online reviews reveals that gender does not have a significant effect on overall provider ratings. Most comments were positive, and a gender bias toward negative reviews of female providers was not found. However, there are multiple categories that represent opportunities to augment the patient experience. For both genders, promptness had lower ratings and represents an opportunity for improvement. Male practitioners specifically may benefit from spending more time with patients in the exam room. With regards to years of experience, new providers scored significantly higher in overall ratings, promptness, and time spent with patients. Medium experience providers led in all other categories. These findings suggest an overall patient preference toward new to medium experience providers. As discussed, the reason for this finding is likely complex and multifactorial, including awareness of the power of social media, practice size, and expertise. We view these findings presented as opportunities for continued improvement of the patient experience.

MATERIALS AND METHODS

The search was conducted (12 August 2020 to 22 August 2020) using the online platforms Vitals, US News, WebMD, Google Reviews, and Healthgrades. A total of 121 Accreditation Council for Graduate Medical Education—accredited academic dermatology programs in the United States and Puerto Rico were included. Several programs were excluded if faculty lists were not available or faculty were not general dermatology providers. In total, we excluded eight programs (four Midwest, three South, and one West Coast). General dermatology providers encompassed a broad category, including general, allergy and immunology, cosmetic, and cutaneous oncology, whereas pediatric dermatology, dermatopathology, and Mohs surgery were excluded. A total of 350 physicians at these programs were identified using a random number generator from faculty lists available on their websites. Given potential regional differences, programs were grouped geographically (Southwest = 22, South = 8, West Coast = 56, Midwest = 87, Northeast = 97, Southeast/Puerto Rico = 80). A total of 28,216 quantitative ratings and 9,792 qualitative (written comments) were identified/analyzed. The study was exempt from Institutional Review Board review.

Quantitative (ratings on a scale of 1 to 5) categories are listed in Table 1. All physicians were rated by patients across the same online platforms, and the quantitative categories on all online platforms use a scale of 1 to 5. Although not all platforms evaluated the same categories and not every individual physician was rated by each platform, we averaged scores in similar categories across multiple platforms to get the composite quantitative rating for each physician. All websites were averaged to get the overall rating score. Vitals and US News both had categories corresponding to bedside manner, and these scores were averaged. US News and WebMD both had categories corresponding to answered questions, and these scores

Table 4. Representative Qualitative Reviews from the Online Rating Platforms

Review	Comment Excerpt
Review 1	Melanoma runs in my family and I need to get regular screenings. Dr. X always performs a very **thorough** examination from head to toe, making sure she won't miss any suspicious lesion. She is very **knowledgeable, professional and kind**. She **takes time** to explain her findings and recommendations, **allowing me to ask any questions** I have. She is an **outstanding doctor and I strongly recommend her** to any one. I trust her with my life.
Review 2	Dr. X treated my son's acne. I was very **impressed with her "bedside manner"** in dealing with my teenager. She **explained the science** behind all available treatment options and was very supportive of his choice. She is an extremely **knowledgeable** physician who clearly has her patients' best interest at heart. **We highly recommend her**!
Review 3	**Rushed. Abrupt. Impersonal**. But that's not the worst. He **MISDIAGNOSED** my wife. We went to him to diagnose a rash spreading all over her body. He **didn't listen** to her comments. He said one area was due to moisture, another due to fungal infection, and her eyes from RUBBING THEM TOO MUCH. In other words, the rash was her own fault. Turns the ENTIRE rash (including the eyes) is due to inverse psoriasis. Dr. X **is the worst**.

Pertinent words and phrases are bolded and further categorized in Table 5.

averaged. Accurate diagnosis, promptness, and time spent with patients were only rated on Vitals.

A standardized qualitative coding methodology was applied to all written comments. Six positive and six negative categories were agreed on by investigators (Table 3). Physician's bedside manner included comments regarding the attitude and personality of the provider (e.g., positive: caring, kind, empathetic; negative: rude, uncaring). Patient's perceived experience included comments remarking on the overall experience, intent to return for future care, and recommending others to seek treatment with this provider. Within perceived experience, the comments were categorized as pertaining to the physician (e.g., love this doctor), staff, or not otherwise specified (comments pertaining to office location, décor, cleanliness, etc.). Communication included the ability of the physicians to answer questions, listen, and provide clarity on procedures. Finance encompassed comments describing general costs.

Wait time included comments describing time spent in the office before seeing the provider or difficulty making appointments. Time spent with patients included comments that described the provider as thorough versus rushed. Competence/knowledge referred to comments regarding the provider's ability to correctly diagnose disease and perceived skill level. Reviews were coded using the scissor and sort method, in which each reviewer read the comments and sorted based on their assessment into the predefined categories. Each written review was sorted into at least one of the six categories (first into a category and then simply as positive or negative). If a review fit more than two categories, it was counted as two individual comments. Tables 4 and 5 include representative quotes using our established coding system. To limit bias, we chose specific words or phrases that would be coded as a positive or negative evaluation and defined these categories ahead of time. All comments were coded by two investigators (DQ, MHT) to ensure internal validity. There was

Table 5. Coding for the Representative Reviews in Table 4

Category	Comment Type	Review 1	Review 2	Review 3
Bedside manner	Positive	X (professional, kind)	X (bedside manner)	—
	Negative	—	—	X (impersonal)
Perceived experience NOS	Positive	—	—	—
	Negative	—	—	—
Perceived experience Staff	Positive	—	—	—
	Negative	—	—	—
Perceived experience physician	Positive	X (strongly recommend her)	X (highly recommend her)	—
	Negative	—	—	X (the worst)
Communication	Positive	X (allows me to ask any questions)	X (explained the science)	—
	Negative	—	—	X (didn't listen)
Finance	Positive	—	—	—
	Negative	—	—	—
Wait time	Positive	—	—	—
	Negative	—	—	—
Time with patients	Positive	X (thorough, takes time)	—	—
	Negative	—	—	X (rushed, abrupt)
Competence	Positive	X (knowledgeable)	X (knowledgeable)	—
	Negative	—	—	X (misdiagnosed)

Abbreviation: NOS, not otherwise specified.
Review 1 was coded as five positive comments. Review 2 was coded as four positive comments. Review 3 was coded as five negative comments.

95.2% agreement between the two raters (95% confidence interval = 92.4−97.2%) on category content and whether the comment was positive or negative.

Physicians were categorized according to gender and years of experience (years since graduation from medical school). Years of experience were further categorized into three groups (6−10 years, 11−20 years, and over 20 years). These were designated as new physicians, medium experience physicians, and advanced physicians. The eight physicians with 0−5 years of experience were excluded from this section of the analysis owing to limited sample size.

Statistical analysis was conducted using R studio (1.1.453). Two sample *t*-test was conducted to compare mean ratings and years of experience between males and females. Paired *t*-test was then conducted to assess whether the mean ratings were different between any two of the quantitative scores for males and females separately. Bonferroni correction was conducted to adjust for multiple comparisons. Adjusted *P*-values are reported. ANOVA was used to compare the mean quantitative scores across groups of years of experience. For qualitative comments, the counts and proportion with 95% confidence interval of positive and negative comments for each category per gender were reported. The proportions were calculated as the number of comments in that category divided by the number of total comments for that gender. OR with 95% confidence interval for receiving a positive comment comparing female to male were calculated. A chi-squared test with Yates' continuity correction was used to assess the association between gender and receiving a positive comment.

Data availability statement

Datasets related to this article can be found at https://data.mendeley.com//datasets/rgk5kn8y9k/1, hosted at Mendeley Digital Commons Data. Citation: Queen et al. (2021), "Patient Satisfaction of General Dermatologists: A Quantitative and Qualitative Analysis of 38,008 Online Reviews by Gender and Years of Experience", Mendeley Data, V2, https://doi.org/10.17632/rgk5kn8y9k.2.

ORCIDs
Megan H. Trager: http://orcid.org/0000-0002-7330-1627
Dawn Q. Queen: http://orcid.org/0000-0001-5665-102X
Weijia Fan: http://orcid.org/0000-0002-0803-4972
Faramarz H. Samie: http://orcid.org/0000-0003-0595-1457

CONFLICT OF INTEREST
The authors state no conflict of interest.

REFERENCES
Bertakis KD, Franks P, Azari R. Effects of physician gender on patient satisfaction. J Am Med Womens Assoc (1972) 2003;58:69−75.

Chen JG, Zou B, Shuster J. Relationship between patient satisfaction and physician characteristics. J Patient Exp 2017;4:177−84.

Damodar D, Donnally CJ 3rd, McCormick JR, Li DJ, Ingrasci GV, Roche MW, et al. How wait-times, social media, and surgeon demographics influence online reviews on leading review websites for joint replacement surgeons. J Clin Orthop Trauma 2019;10:761−7.

Dunivin Z, Zadunayski L, Baskota U, Siek K, Mankoff J. Gender, soft skills, and patient experience in online physician reviews: a large-scale text analysis. J Med Internet Res 2020;22:e14455.

Ganguli I, Sheridan B, Gray J, Chernew M, Rosenthal MB, Neprash H. Physician work hours and the gender pay gap - evidence from primary care. N Engl J Med 2020;383:1349−57.

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:e38.

Jack RA 2nd, Burn MB, McCulloch PC, Liberman SR, Varner KE, Harris JD. Does experience matter? A meta-analysis of physician rating websites of orthopaedic surgeons. Musculoskelet Surg 2018;102:63−71.

Ku JH, Danve A, Pang H, Choi D, Rosenbaum JT. Determinants of patient satisfaction in an academic rheumatology practice. J Clin Rheumatol 2015;21:256−62.

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;12:272−7.

Queen D, Trager MH, Fan W, Samie FH. Patient satisfaction of general dermatology providers: a quantitative and qualitative analysis of 38,008 online reviews. JID Innov 2021;1:100049.

Waqas B, Cooley V, Lipner SR. Association of sex, location, and experience with online patient ratings of dermatologists. J Am Acad Dermatol 2020;83:954−5.