

Original Paper

Physician Gender, Patient Risk, and Web-Based Reviews: Longitudinal Study of the Relationship Between Physicians' Gender and Their Web-Based Reviews

Danish Hasnain Saiffee¹, PhD; Matthew Hudnall¹, PhD; Uzma Raja², PhD

¹Department of Information Systems, Statistics, and Management Science, The University of Alabama, Tuscaloosa, AL, United States

²Department of Systems and Technology, Auburn University, Auburn, AL, United States

Corresponding Author:

Matthew Hudnall, PhD

Department of Information Systems, Statistics, and Management Science

The University of Alabama

801 University Blvd

Tuscaloosa, AL, 35487-0290

United States

Phone: 1 205 348 0856

Email: matthew.hudnall@ua.edu

Abstract

Background: Web-based reviews of physicians have become exceedingly popular among health care consumers since the early 2010s. A factor that can potentially influence these reviews is the gender of the physician, because the physician's gender has been found to influence patient-physician communication. Our study is among the first to conduct a rigorous longitudinal analysis to study the effects of the gender of physicians on their reviews, after accounting for several important clinical factors, including patient risk, physician specialty, and temporal factors, using time fixed effects. In addition, this study is among the first to study the possible gender bias in web-based reviews using statewide data from Alabama, a predominantly rural state with high Medicaid and Medicare use.

Objective: This study conducts a longitudinal empirical investigation of the relationship between physician gender and their web-based reviews using data across the state of Alabama, after accounting for patient risk and temporal effects.

Methods: We created a unique data set by combining data from web-based physician reviews from the popular physician review website, RateMDs, and clinical data from the Center for Medicare and Medicaid Services for the state of Alabama. We used longitudinal econometric specifications to conduct an econometric analysis, while controlling for several important clinical and review characteristics across four rating dimensions (helpfulness, knowledge, staff, and punctuality). The overall rating and these four rating dimensions from RateMDs were used as the dependent variables, and physician gender was the key explanatory variable in our panel regression models.

Results: The panel used to conduct the main econometric analysis included 1093 physicians. After controlling for several clinical and review factors, the physician random effects specifications showed that male physicians receive better web-based ratings than female physicians. Coefficients and corresponding SEs and *P* values of the binary variable *GenderFemale* (1 for female physicians and 0 otherwise) with different rating variables as outcomes were as follows: *OverallRating* (coefficient -0.194, SE 0.060; *P*=.001), *HelpfulnessRating* (coefficient -0.221, SE 0.069; *P*=.001), *KnowledgeRating* (coefficient -0.230, SE 0.065; *P*<.001), *StaffRating* (coefficient -0.123, SE 0.062; *P*=.049), and *PunctualityRating* (coefficient -0.200, SE 0.067; *P*=.003). The negative coefficients indicate a bias toward male physicians versus female physicians for aforementioned rating variables.

Conclusions: This study found that female physicians receive lower web-based ratings than male physicians even after accounting for several clinical characteristics associated with the physicians and temporal effects. Although the magnitude of the coefficients of *GenderFemale* was relatively small, they were statistically significant. This study provides support to the findings on gender bias in the existing health care literature. We contribute to the existing literature by conducting a study using data across the state of Alabama and using a longitudinal econometric analysis, along with incorporating important clinical and review controls associated with the physicians.

(*J Med Internet Res* 2022;24(4):e31659) doi: [10.2196/31659](https://doi.org/10.2196/31659)

KEYWORDS

web-based physician reviews; gender; gender bias; patient perception; Alabama; patient risk

Introduction

Background

Web-based reviews of physicians have been gaining significant popularity among health care consumers or patients over the past 2 decades. Some examples of popular websites for web-based physician reviews are RateMDs [1], Vitals [2], and HealthGrades [3]. The prominence of these reviews is enhanced as the health care landscape in the United States becomes more patient-centric. Patients are becoming more involved in the management of their own health care. Although the review websites were initially popular among certain demographics [4], over time, they have gained significant popularity across a substantial portion of patient population. In fact, a recent survey of web-based physician reviews found that approximately 95% of the respondents viewed web-based reviews to be *somewhat reliable* or *very reliable*, and approximately 70% of respondents said that their choice of a physician was affected by the ratings or reviews on web-based physician review websites [5].

The literature on web-based reviews of physicians has been growing in the past 10 years. Using data from the United States and other countries, numerous studies have examined the content and valence of web-based physician or hospital reviews and the factors that could explain their variance [6-15]. A substream of this literature examined the relationship between the clinical outcomes or performance of physicians and their web-based reviews. The results were quite mixed [13]. Some studies have found a statistically significant association between physicians' clinical performance and their web-based reviews [16-18]. On the other hand, some studies have found that physicians with better clinical practices or outcomes do not receive better web-based reviews [19-21].

Another substream has investigated the influence of web-based physician reviews on patients' choices. There has been a significant interest among health care researchers and practitioners in the health care consumers' awareness of web-based physician reviews [22]. Several studies have investigated whether web-based physician reviews impact patients' choices and whether there are certain characteristics of these reviews that impact the choice. These studies found that high number of reviews and high valence of reviews were associated with a more positive attitude toward the rated physicians and their selection by patients [23-25].

The increasing reliance on web-based physician reviews is indicated by other surveys also [26]. These surveys of web-based reviews also reveal that a significant portion of patients checks the web-based reviews of physicians, even if they were referred to these physicians by their health care providers. Collectively, these findings reveal the extent to which web-based reviews of physicians have become prominent among patients or health care consumers.

As web-based health care information, including physician reviews, is publicly available and easily accessible, there has

been a long-standing concern among the health care providers and research communities about the quality and clinical relevance of web-based health care information [27]. The interaction between health care providers and their patients can affect the patients' opinions of them. In turn, these opinions can become web-based reviews that are accessible to anyone searching for their physicians' information on the web.

There has been a long-established interest among researchers in the impact of physician gender on patient communication and patients' choice of physicians. Extant literature has found that female physicians tend to engage in patient-centered communication [28-30] and do not receive ratings as high as their male counterparts [31,32]. It has also been proposed that the relationship between physicians and their patients might be affected by the physician's gender and different expectations of patients from male and female physicians [33-35]. The dynamics of patients' communication and relationship with physicians of different genders have received significant attention in the extant literature [36,37].

Questions about whether patients have a preference for male physicians over female physicians, and vice versa, and whether their opinions of physicians are affected by the physicians' gender have also received substantial attention from health care professionals and researchers. For instance, in a survey of 185 patients, Fennema et al [38] found that 43% of women and 12% of men preferred a female physician, whereas 31% of men and 9% of women preferred a male physician and that patients who preferred male physicians reported technical competence to be a more prominent characteristic of male physicians. In a different survey, Kerssens et al [39] did not find a preference for surgeons or anesthesiologists of a particular gender, but found preferences for female physicians as gynecologists in 8.5:1 ratio and general practitioners in 2.32:1 ratio among female respondents. In another survey of 125 women, Plunkett et al [40] found that the gender of a physician was not of primary importance when selecting an obstetrician or gynecologist. Some of these studies have also attempted to identify the mechanisms that may have led to their findings. There have also been calls for suggestions on making health care workplaces more equitable for female physicians [41].

With the proliferation of web-based physician reviews among patients or health care consumers, a natural and important question is, "Whether and to what extent is a physician's gender related to their online reviews after accounting for patient risk and time shocks (time fixed effects)?"

After a careful review of the existing literature, we found that the potential effect of physician gender on web-based reviews of physicians has not received sufficient attention. In the few studies that have examined the relationship between physicians' gender and their web-based reviews, the findings have been mixed. For example, Dunivin et al [32] and Thawani et al [42] found that female physicians receive lower ratings than male physicians. On the other hand, Emmert and Meier [43] found that female physicians receive better ratings than their male

counterparts. Marrero et al [44] found that female surgeons receive more positive ratings for social interaction, whereas male surgeons receive better ratings for technical aspects. Clearly, the possible effect of physicians' gender on their web-based reviews, or lack thereof, requires more thorough examination.

In the examination of the aforementioned relationship, it is important to account for the characteristics of patients, such as patient risk, in some form. It is also important to account for the variation in the reviews over time to determine the direct relationship between physicians' gender and their web-based reviews. Including patient risk allows us to account for the health characteristics of a significant patient population under the care of physicians. Not controlling for such characteristics can potentially bias the results because a physician's interaction can be affected by the existing health condition of their patients. Therefore, we examine the effect of physician gender on web-based patient reviews, while controlling for patient health risks over time.

Objective

To the best of our knowledge, our study is among the first to examine the effect of physicians' gender on their web-based reviews over time and after accounting for patient risk. Furthermore, our study is the first to conduct such an investigation using physician data across Alabama, a state that has received very little attention in the literature on web-based physician reviews. We accomplish our analysis by using a unique data set that we created by combining data from web-based physician reviews from a popular physician review website, RateMDs, and clinical data from the Center for Medicare and Medicaid Services (CMS) for the state of Alabama.

Methods

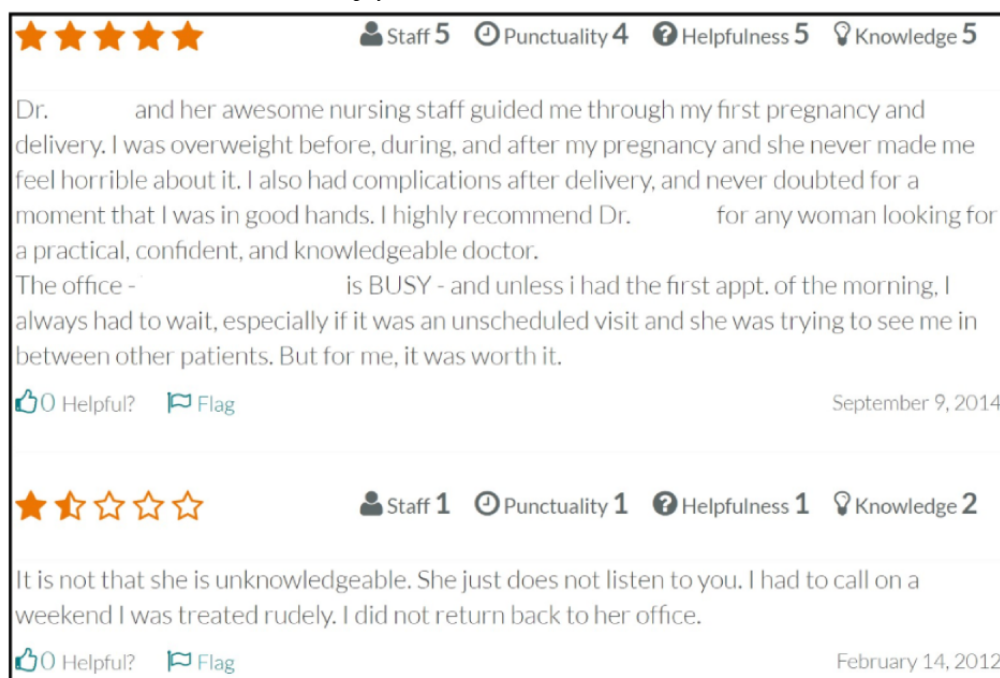
Ethics Approval

No ethics board review or approval was required for this study. All the raw data that were collected for this study are publicly available on the web.

Data

To study whether web-based reviews of physicians are more favorable toward male or female physicians, we constructed a panel data set of physicians in Alabama using data from 2 sources. The unit of analysis in our study was a physician, and the time periods in the panel were years. We collected data on web-based reviews and the gender of physicians from RateMDs to construct our web-based review data set spanning from 2012 to 2018. We used Python (Python Software Foundation) to collect data from RateMDs. We also obtained clinical data on physicians from Medicare Provider Utilization and Payment Data: Physician and Other Supplier [45], which traversed the same time frame of 2012 to 2018. We combined the data from these sources using a combination of physicians' first names, last names, specialty, and years. Our final unbalanced panel data set had 1093 matching physicians over a 7-year time span (2012 to 2018) that matched both data sets. There were 5912 physicians in the RateMDs data set who had at least one review, and there were a total of approximately 26,600 reviews across these physicians. Among these 5912 physicians, 2673 (45.21%) physicians had reviews in at least two years. We were able to match 40.89% (1093/2673) of these physicians with our data from CMS, and this 40.89% (1093/2673) of the physicians constituted the panel used to conduct the panel analysis in this study.

Each physician in our final panel has a unique national provider identification number that was collected from CMS. This ensured that all the physicians in our final panel were unique. Figure 1 shows an anonymized selection of reviews from RateMDs for a physician in our data set. As shown in Figure 1, a physician can receive numeric ratings on four different dimensions (staff, punctuality, helpfulness, and knowledge). Along with these numeric ratings, a physician can also receive textual comments. The dates on which the reviews were provided on RateMDs is also shown in Figure 1. Patient reviews on RateMDs and optional responses by the physicians are free of charge. Paid tiers for physicians exist on RateMDs, but they do not allow for the alteration of reviews. The paid tiers allow for physicians to be notified of new ratings, the ability to feature a rating, appointment requests, photos, and other features, but no paid feature inhibits the ability of a person to post a review on the site.

Figure 1. Example screenshot of RateMDs reviews for a physician.

Measures

As we were examining whether the web-based reviews of physicians are favorable to male or female physicians, we constructed our dependent variables using the numeric physician ratings from RateMDs. Physicians on RateMDs can be rated on four dimensions: helpfulness, knowledge, staff, and punctuality. The ratings for each of these dimensions are on a scale of 1 to 5, with 5 being the best possible score and 1 being the lowest score. To capture the information in each of these four dimensions, we constructed the following four dependent variables: *HelpfulnessRating*, *KnowledgeRating*, *StaffRating*, and *PunctualityRating*. *HelpfulnessRating* was the average of the ratings received by a physician on the helpfulness dimension in a year. Similarly, *KnowledgeRating*, *StaffRating*, and *PunctualityRating* were the averages of the ratings received by a physician on the knowledge, staff, and punctuality dimensions, respectively. To capture the combined information across these dimensions, we constructed a panel variable, *OverallRating*. For this purpose, initially, we constructed a variable *NetRating* using the average of the ratings received on the four aforementioned dimensions. Then, we constructed *OverallRating* by calculating the average of *NetRating* in each year, similar to how we constructed *HelpfulnessRating*, *KnowledgeRating*, *StaffRating*, and *PunctualityRating*.

Our key explanatory variable was a time-invariant variable, *GenderFemale*, which equals 1 for female physicians and 0 for male physicians. We obtained data on the gender of the physicians from RateMDs. We also used several control variables to account for the clinical aspects associated with the physician and with the textual comments that go alongside numeric *RateMDs* ratings. Our control variables included *RiskScore*, *TopicCare*, *TopicSurgery*, *TopicStaff*, and *Specialty*.

RiskScore was the average yearly hierarchical condition category (HCC) risk score calculated by CMS using data on Medicare

beneficiaries [45]. HCC coding can provide information about patient complexity and a description of the medical complications a patient is experiencing. HCC relies on the International Classification of Diseases–10th Edition coding to assign risk scores to patients [46]. A physician with high *RiskScore* would have Medicare beneficiaries with high risk scores (above-average spending). This variable allowed us to control for the patient risk score of the Medicare patients under the care of a physician. As Medicare is one of the largest health care insurers or payers in the United States, *RiskScore* helped us to account for the patient risk of a significant proportion of the patient population under the care of physicians.

TopicCare was the proportion of textual reviews received by a physician each year, in which the dominant underlying theme was care provided by the physician. *TopicStaff* was the proportion of textual reviews in which the dominant underlying theme was the office or staff of the physician. *TopicSurgery* was the proportion of textual reviews in which the dominant underlying theme was the surgical proficiency of the physician.

To construct these topics (latent topics), we used topic modeling techniques based on Latent Dirichlet Allocation (LDA) [47,48]. LDA has been used extensively for topic modeling in the extant literature on web-based reviews of products and services, including several studies involving web-based physician reviews [9,49–53]. The following sections provide a brief description of the main steps through which we used topic modeling to construct the aforementioned topic variables. We used R (R Foundation for Statistical Computing) for topic modeling.

We created a corpus of all the reviews using an R text-mining package(TM) within RStudio, after which we converted the corpus to lower case [54–56]. We also replaced punctuation, numbers, and stop words. We *stemmed* the corpus to allow us to reduce words with a common root to the root word, such as *nurse* and *nursing* to the word fragment *nurs*. Next, we created

the *document-term matrix*, which stored the frequencies of stemmed words in our textual comment corpus by each textual comment. Then, we leveraged the LDA algorithm and used an R package (*topicmodels*) to extract topics from our textual comments [57-59]. These R packages have been widely used in the literature mentioned previously to construct latent topics or themes from textual data. For each comment, a probability was assigned to each of the identified latent themes or topics, and the probabilities summed up to 1 for each comment. We

classified each comment based on the topic that had the highest probability. We identified the most common words within each of the 3 target latent topics, as shown in [Textbox 1](#). We chose these 3 topics because it was the minimum number of topics that we could use to clearly categorize the experiences with the physicians and their staff reported in textual reviews [20,21]. [Textbox 1](#) shows the stemmed words most closely (probabilistically) associated with each of the 3 review comment topics.

Textbox 1. Most prominent words (after stemming) by topic.

TopicCare

- care, doctor, staff, recommend, patient, time, knowledg, help, friend, love, wonder, high, listen, excel, and feel

TopicStaff

- time, office, doctor, wait, staff, patient, appoint, call, nurs, rude, visit, day, question, hour, and talk

TopicSurgery

- doctor, surgeri, pain, care, medic, life, patient, treat, recommend, time, day, surgeon, procedur, treatment, and feel

We had physicians from across 34 specialties in our final panel data set. The 15 specialties with most physicians (in descending order of the number of physicians) were as follows: general (family) practice, obstetrics and gynecology, internal medicine, orthopedic surgery, neurology, otolaryngology, cardiology, ophthalmology and optometry, psychiatry, dermatology, general surgery, podiatry, urology, endocrinology, and rheumatology. Physicians in these 15 specialties accounted for approximately 85.73% (937/1093) of all the physicians in our panel data set. Table S1 in [Multimedia Appendix 1](#) lists the number of male and female physicians across specialties in the panel data set. The physician specialties were time-invariant binary variables. Controlling for the specialties allowed us to compare the effect of the physicians' gender on their reviews after accounting for the numerous unobservable time-invariant clinical aspects that could influence physicians of both genders within each specialty. We also conducted further robustness checks by including additional clinical review control variables. These control measures helped us significantly distinguish our research from previous studies.

Analysis

We used physician random effects panel regression, along with year fixed effects to account for time shocks. A time shock in the context of this paper can be considered as an event or collection of events that can impact physicians across the board in the duration of a year. For example, a statewide or nationwide health care policy change would likely have an impact on physicians across different specialties. As the analysis used panel data, it was important to account for such time shocks. We did so by including year fixed effects in our regression specifications. We used Stata (StataCorp) for conducting our econometric analysis.

We leveraged the physician random effects model instead of the physician fixed effects model to estimate the effect of physician gender because of the following reasons: (1) our main

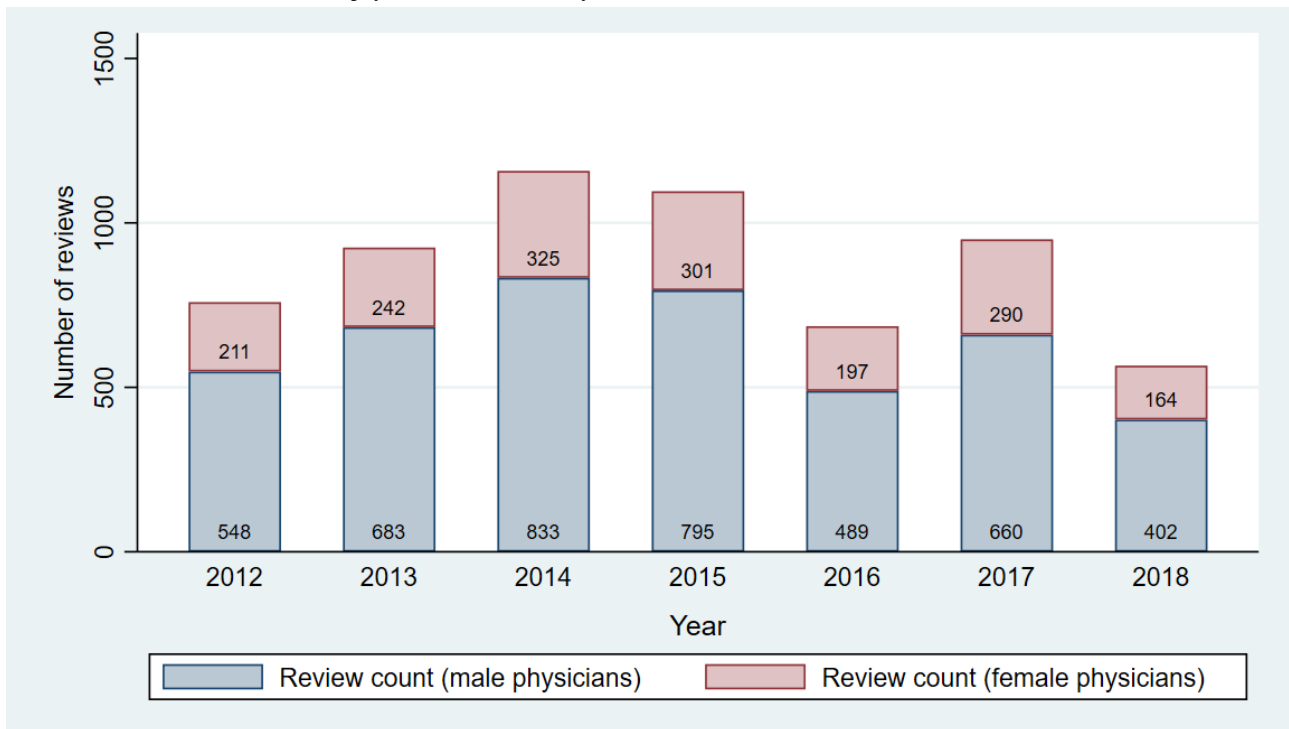
explanatory variable, *PhysicianGender*, was time-invariant, and physician fixed effects would have subsumed the *PhysicianGender* variable and (2) a physician's gender can be safely assumed to be randomly assigned in the context of our study, and thus, it was very unlikely that there were unobserved variables that could simultaneously drive or influence both the physician gender and their web-based reviews. The year fixed effects allowed us to account for the time shocks in the health care industry or web-based physician review websites that can influence physicians across the state of Alabama. The SEs shown in all the panel regression specifications were robust. For brevity, we do not report the coefficients, SEs, and *P* values of the different specialties and year fixed effects. The sum of *TopicCare*, *TopicSurgery*, and *TopicStaff* was equal to 1. In our specifications, *TopicStaff* was the base topic variable, and thus, not included in the regressions. One of the specialties and one of the years acted as the base specialty and base year, respectively, and thus, were not included in the regression specifications.

Results

Descriptive Statistics

[Figure 2](#) shows the distribution of the number of reviews for male and female physicians across the years from the original RateMDs data set. This chart and the subsequent figures were created using the 1093 physicians who were present in our panels across CMS and RateMDs data used for panel regressions. Our panel consisted of a 7-year period spanning from 2012 to 2018 to include a broad set of historical data that were also relatively current. As shown in [Figure 2](#), the year 2014 had the highest number of reviews, whereas 2018 had the lowest number of reviews across the physicians in our panel, and there were ample number of physician reviews across all years in our panel.

Figure 2. Distribution of total number of physician reviews across years.



Figures 3-7 show plots of the average annual values of *OverallRating*, *HelpfulnessRating*, *KnowledgeRating*, *StaffRating*, and *PunctualityRating* for male and female physicians for the 1093 physicians from RateMDs who were in the panel. As shown in Figure 3, the average *OverallRating* of male physicians was consistently higher than that of female physicians. The average annual ratings on all 4 dimensions were more favorable for male physicians across most years. The time trends depicted in these figures revealed reviews more favorable toward male physicians than female physicians. The variation

in the difference in the average values is visible in these figures and warrants a thorough longitudinal investigation of the effect of physician gender on the web-based ratings. Accordingly, we conducted a longitudinal or panel empirical investigation of the effect of a physician’s gender on their web-based ratings. As stated previously, we controlled for several clinical and review characteristics associated with physicians, and by doing so, we isolated the direct effect of a physician’s gender on their web-based ratings.

Figure 3. Comparison of average overall ratings for female and male physicians across years.

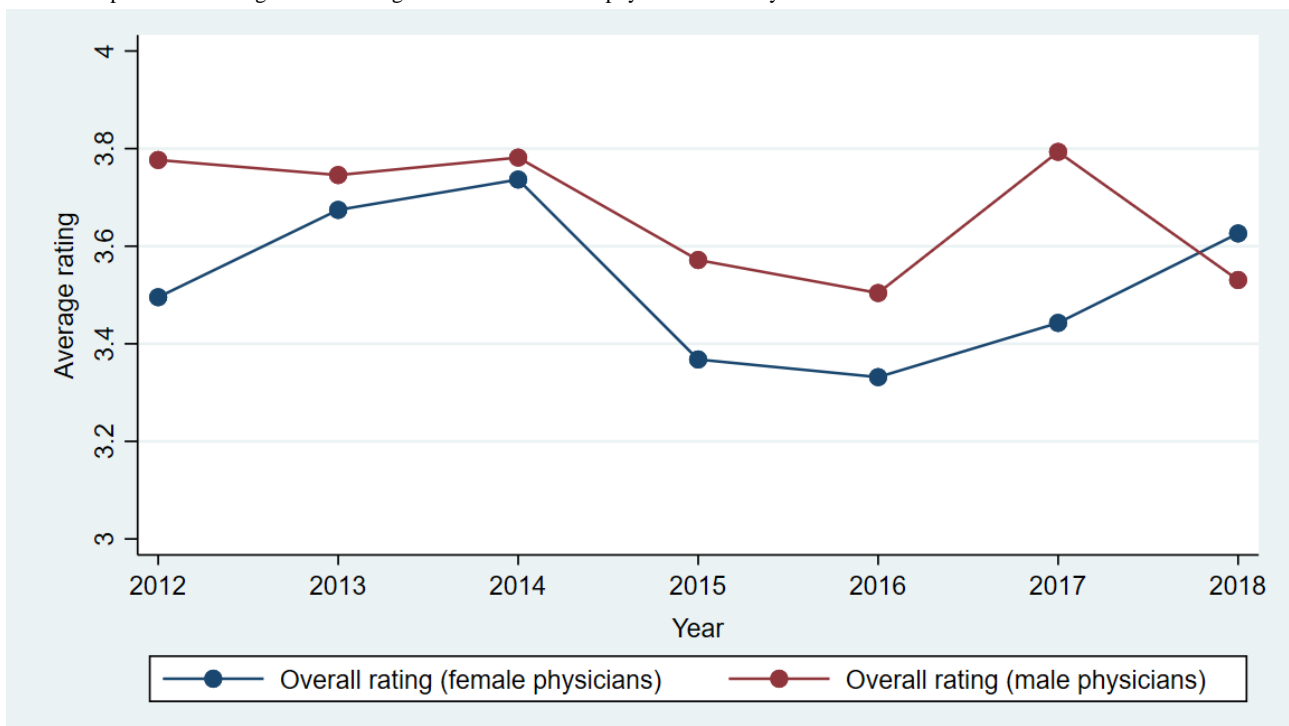


Figure 4. Comparison of average helpfulness ratings for female and male physicians across years.

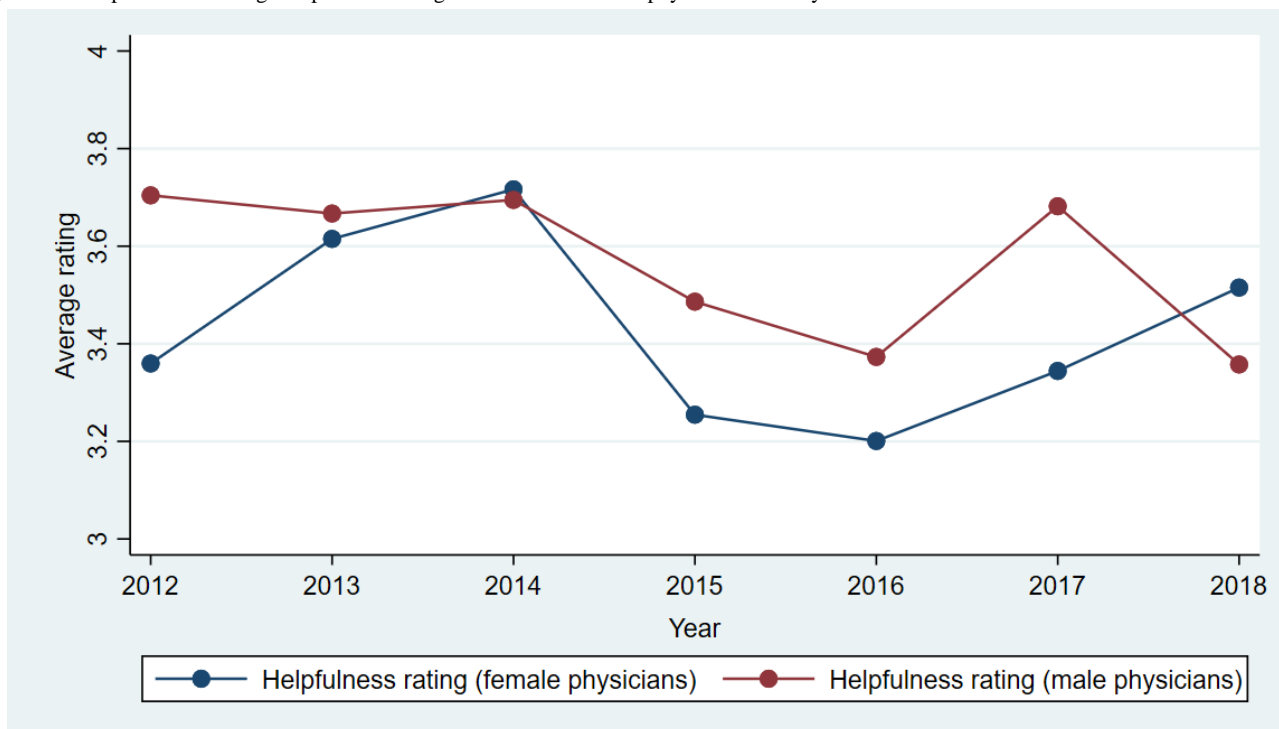


Figure 5. Comparison of average knowledge ratings for female and male physicians across years.

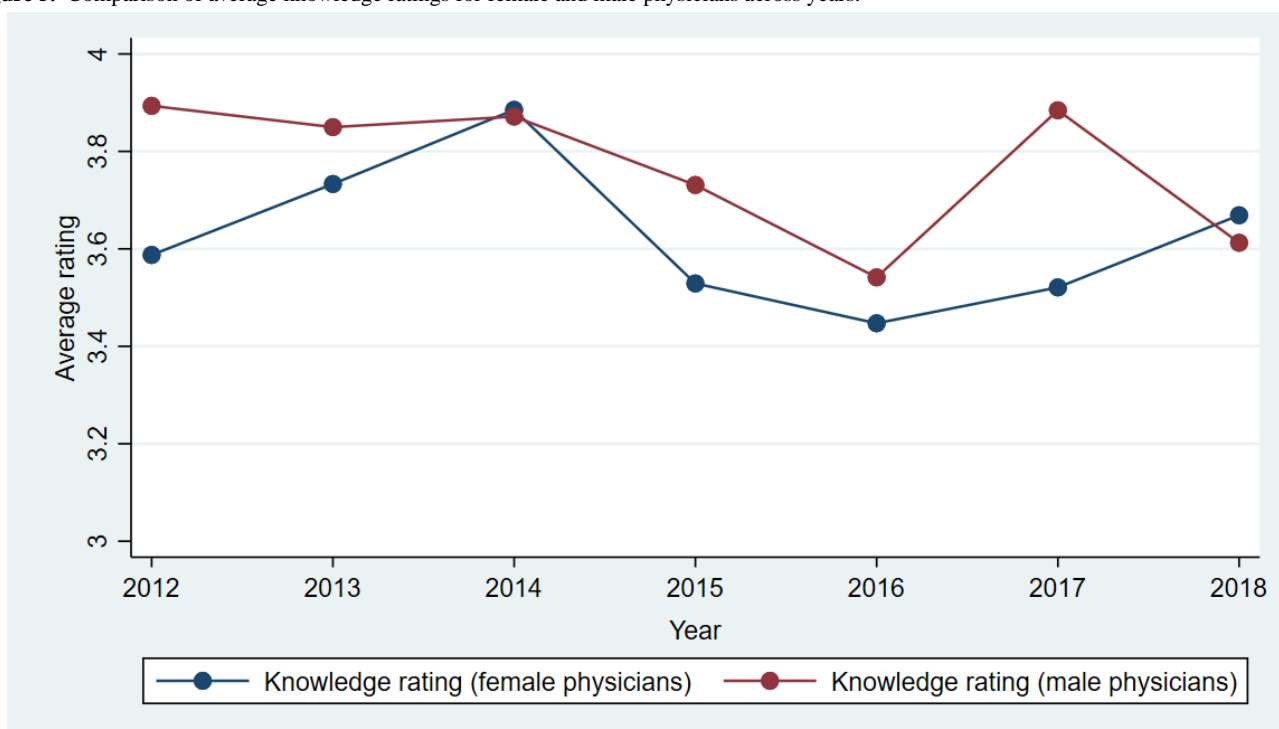


Figure 6. Comparison of average staff ratings for female and male physicians across years.

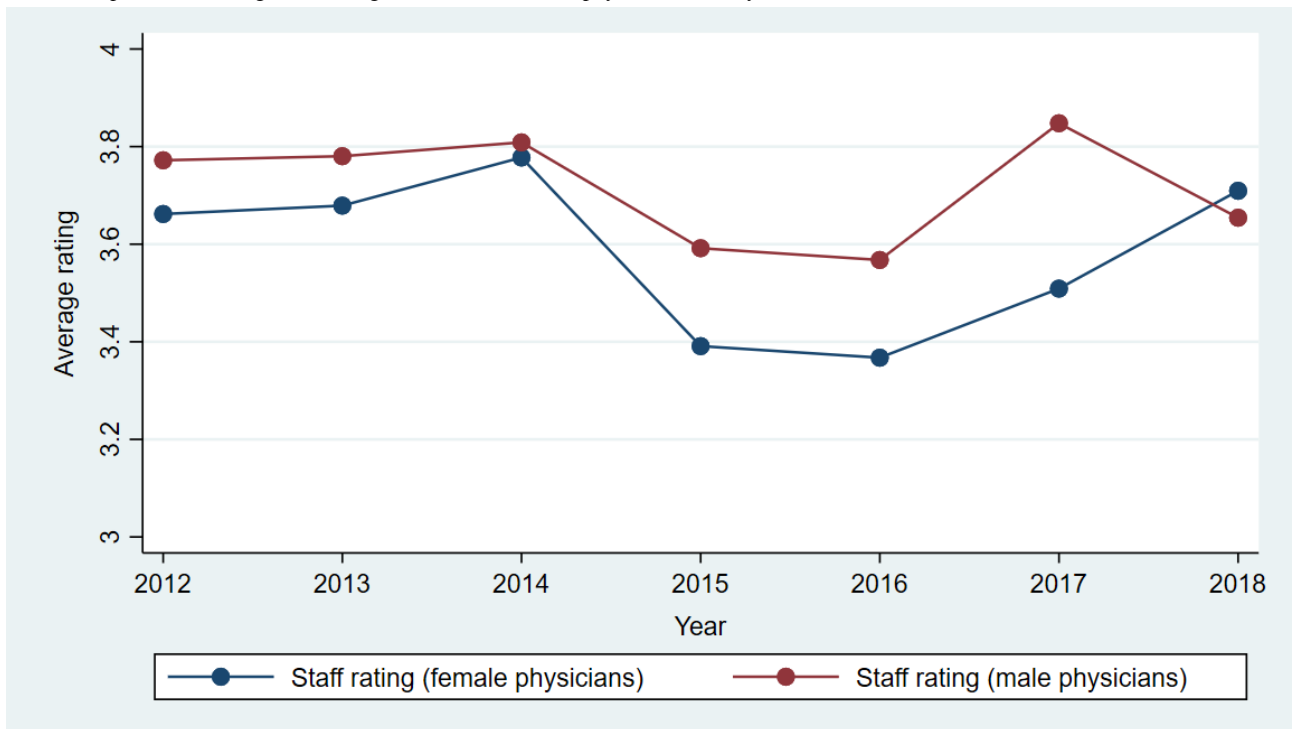


Figure 7. Comparison of average punctuality ratings for female and male physicians across years.

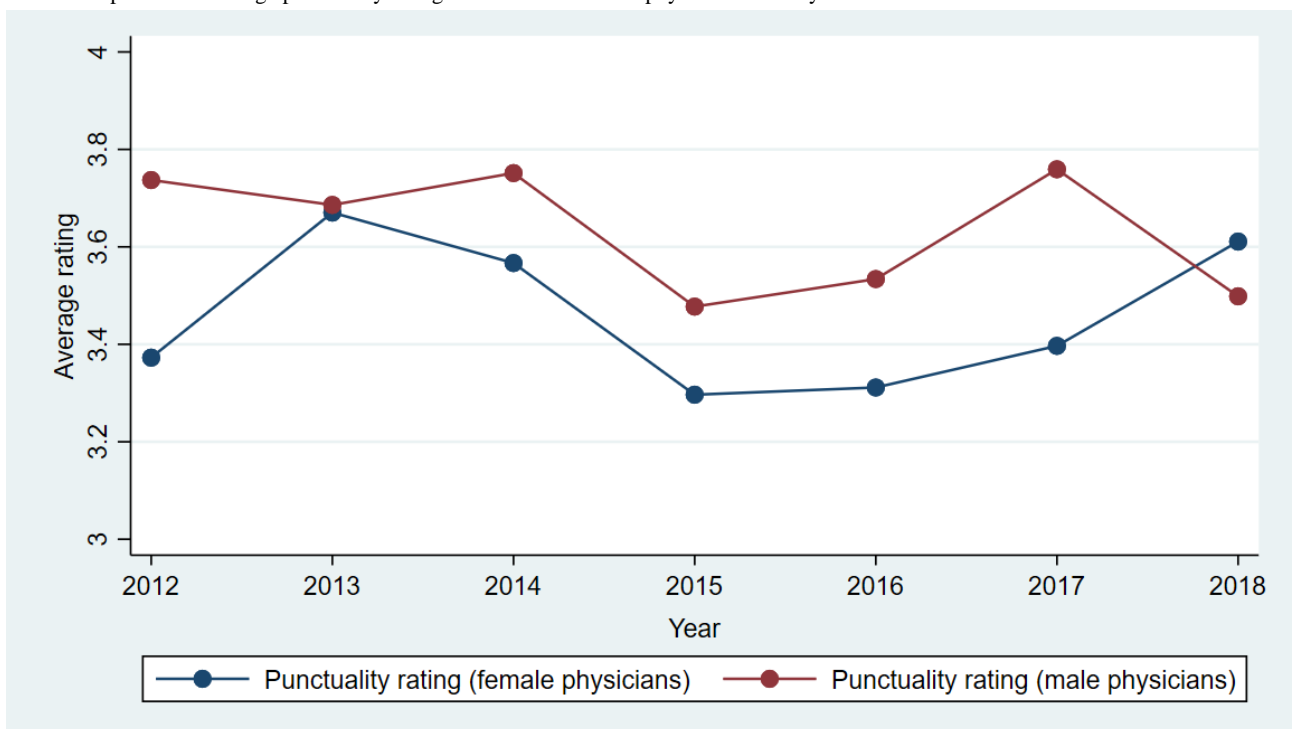


Table 1 shows the descriptive statistics of the various dependent variables, topic controls, and *RiskScore* control. The average values of the rating variables were between 3.5 and 3.6. In our panel, approximately 25.34% (277/1093) of the physicians were women. In Alabama, female physicians account for

approximately 28.5% (3025/10,614) of the overall physician population [60]. This suggests that the overall distribution of physician gender in our panel was fairly representative of that in Alabama.

Table 1. Descriptive statistics (number of observations=3446).

Variable	Values, mean (SD)	Values, median	Values, minimum	Values, maximum
<i>OverallRating</i>	3.64 (1.43)	4.25	1	5
<i>HelpfulnessRating</i>	3.54 (1.65)	4.37	1	5
<i>KnowledgeRating</i>	3.74 (1.54)	5	1	5
<i>StaffRating</i>	3.69 (1.48)	4	1	5
<i>PunctualityRating</i>	3.60 (1.49)	4	1	5
<i>TopicCare</i>	0.41 (0.45)	0	0	1
<i>TopicSurgery</i>	0.27 (0.40)	0	0	1
<i>TopicStaff</i>	0.32 (0.42)	0	0	1
<i>RiskScore</i>	1.23 (0.41)	1.14	0.53	5.62

Effect of Gender

Table 2 provides the results of random effects panel regression, with *OverallRating* as the dependent variable. We included physician specialties as controls and year fixed effects in each of the regression specifications. The SEs of each specification were robust. As shown in **Table 2**, the coefficient of *GenderFemale* was negative and statistically significant, implying that female physicians tend to receive worse overall web-based ratings than their male counterparts. The coefficient of *RiskScore* was statistically insignificant in all the specifications, implying that physicians who treat Medicare patients of high risk tend to not receive better or worse overall ratings than their counterparts who treat Medicare patients of low risk. The coefficients of *TopicCare* and *TopicSurgery* were positive and statistically significant, implying that the physicians who receive a high proportion of review comments with an

underlying theme of physician care and surgical aspects tend to have better overall ratings than those who receive a high proportion of review comments with an underlying theme of their office and staff. In **Tables 3** and **4**, the coefficient of *GenderFemale* was negative and statistically significant for *HelpfulnessRating*, *KnowledgeRating*, and *PunctualityRating*, but not for *StaffRating*.

The coefficient of *RiskScore* was statistically insignificant for each of the four rating dimensions, whereas that of *TopicCare* and *TopicSurgery* were positive and statistically significant. The magnitude of the coefficient of *GenderFemale* was close to 0.2. This means that, on average, female physicians receive ratings lower by 0.2 points than their male counterparts. For example, on average, if male physicians receive a rating of 4 out of 5, their female counterparts would receive a rating of 3.8 out of 5.

Table 2. Estimation for OverallRating (N=1093)^a.

Variable	Coefficient (SE)	P value
<i>GenderFemale</i>	-0.162 (0.060)	.007
<i>RiskScore</i>	-0.056 (0.086)	.52
<i>TopicCare</i>	1.557 (0.058)	<.001
<i>TopicSurgery</i>	0.739 (0.071)	<.001

^aSpecialty controls=yes; year fixed effects=yes; robust SE=yes; overall R-squared=0.267; within R-squared=0.168; between R-squared =0.339.

Table 3. Estimation for HelpfulnessRating and KnowledgeRating (N=1093).

Variable	<i>HelpfulnessRating</i> ^a		<i>KnowledgeRating</i> ^b	
	Coefficient (SE)	P value	Coefficient (SE)	P value
<i>GenderFemale</i>	-0.185 (0.069)	.008	-0.198 (0.065)	.002
<i>RiskScore</i>	0.003 (0.098)	.97	-0.057 (0.094)	.54
<i>TopicCare</i>	1.702 (0.069)	<.001	1.492 (0.064)	<.001
<i>TopicSurgery</i>	0.688 (0.084)	<.001	0.513 (0.080)	<.001

^aSpecialty controls=yes; year fixed effects=yes; robust SE=yes; overall R-squared=0.239; within R-squared=0.153; between R-squared=0.310.

^bSpecialty controls=yes; year fixed effects=yes; robust SE=yes; overall R-squared=0.220; within R-squared=0.137; between R-squared=0.282.

Table 4. Random effects panel regression (StaffRating and PunctualityRating; N=1093).

Variable	<i>StaffRating</i> ^a		<i>PunctualityRating</i> ^b	
	Coefficient (SE)	P value	Coefficient (SE)	P value
<i>GenderFemale</i>	-0.095 (0.062)	.13	-0.172 (0.067)	.01
<i>RiskScore</i>	-0.045 (0.087)	.61	-0.127 (0.105)	.23
<i>TopicCare</i>	1.547 (0.063)	<.001	1.488 (0.063)	<.001
<i>TopicSurgery</i>	0.923 (0.076)	<.001	0.832 (0.074)	<.001

^aSpecialty controls=yes; year fixed effects=yes; robust SE=yes; overall R-squared=0.247; within R-squared=0.155; between R-squared=0.315.

^bSpecialty controls=yes; year fixed effects=yes; robust SE=yes; overall R-squared=0.234; within R-squared=0.130; between R-squared=0.318.

Robustness Checks

We added additional control variables to check whether our findings would change. The three additional variables were *BeneficiaryCount*, *ServicesCount*, and *WordCount*. *BeneficiaryCount* was the number of Medicare beneficiaries under the care of a physician in a year. *ServicesCount* was the number of services provided by a physician in a year. *WordCount* was the average number of words in the review comments received by a physician in a year. Tables S2-S4 in [Multimedia Appendix 1](#) provide the results of panel specifications with additional control variables. Table S2 in [Multimedia Appendix 1](#) provides the results with *OverallRating* as the dependent variable. Table S3 in [Multimedia Appendix 1](#) provides the results with *Helpfulness* and *KnowledgeRating* as the dependent variables, and Table S4 in [Multimedia Appendix 1](#) provides the results with *StaffRating* and *PunctualityRating* as the dependent variables. As can be observed in Tables S2-S4 in [Multimedia Appendix 1](#), the coefficients of *GenderFemale* were negative and statistically significant for *OverallRating* and each of the four rating dimensions, including *StaffRating*. The magnitude of coefficient of *GenderFemale* was close but slightly higher than those in [Tables 2-4](#).

We conducted further robustness checks by removing the specialties in our panel in which both genders were not represented. This helped us mitigate the concern that a possible bias may arise owing to the absence of physicians of one of the genders in any of the specialties in our panel. The results displayed in Tables S5-S7 in [Multimedia Appendix 1](#) are consistent with our original findings that female physicians receive lower ratings than their male counterparts.

In our next robustness check, we conducted our main regression analysis without topic controls. This test was conducted to examine whether the topic variables may have introduced a systemic bias in the specifications owing to the manner in which they were constructed and whether the negative coefficient of *GenderFemale* variable may have been an artifact. As can be observed from the results in Tables S8-S10 in [Multimedia Appendix 1](#), the coefficient of *GenderFemale* was negative and statistically significant across the specifications, even after topic controls were excluded. This further supports our main finding that female physicians tend to receive worse web-based reviews than their male counterparts. The topic controls play an important role in our specifications because they help to explain part of the variance in the web-based ratings. This can be further

understood by comparing the overall R-squared, within R-squared, and between R-squared values in [Tables 2, 3, and 4](#) with those in Tables S8, S9, and S10 in [Multimedia Appendix 1](#), respectively. The 3 R-squared values were substantially higher in [Tables 2-4](#), which means that the topic controls explained a considerable part of the variance in the web-based rating variables.

In summary, we conducted three additional robustness checks as explained above: (1) included additional control variables, (2) removed the specialties that did not include physicians of both genders, and (3) removed the topic controls. After conducting these robustness checks, we can conclude that female physicians tend to receive worse web-based reviews than their male counterparts. This finding is consistent across the regression specifications used in this study.

A concern could be about how representative the data in our panel are of the original data collected from RateMDs and Medicare (CMS). To address this concern, we calculated the descriptive statistics of the variables shown in [Table 1](#) using the original longitudinal data collected from RateMDs and Medicare. The descriptive statistics are shown in Table S11 in [Multimedia Appendix 1](#). A comparison of the statistical values in Table S11 in [Multimedia Appendix 1](#) shows that the panel data used for the econometric analysis in our study are fairly representative of the original data collected from the 2 aforementioned sources.

Discussion

Overview

Our study provides an important contribution to the growing literature on web-based physician reviews and physician gender. A possible concern could be that the differences observed in the reviews between physicians of different genders could be driven by the differences in the quality of care or outcomes delivered by physicians of different genders. To address this concern, we performed a substantial search of the existing literature examining the differences between the quality of clinical care or outcomes delivered by male and female physicians. We found several research papers in this context [61-65], but we could not find significant evidence from extant research that male physicians deliver better care than female physicians.

Principal Findings

We found that male physicians receive better web-based reviews than female physicians after controlling for their clinical characteristics such as specialty and patient risk. Although the difference between the web-based ratings for male and female physicians was statistically significant, the average magnitude of the difference was not substantial. Our findings support that of Dunivin et al [32] and Thawani et al [42], but do not support the findings of Emmert and Meier [43], who found that during the examined time frame, female physicians had better reviews than male physicians. Their results indicated a slight but statistically significant preference for female physicians (2% differential in the percentage of reviews below the mean for each gender) compared with our results that found a 0.2 differential on a 5-point scale in favor of male physicians (4% difference). Possible reasons for these differences could be attributed to cultural variations between the patient populations in Alabama and Germany and that the reviews collected by Emmert and Meier [43] included more female respondents than male respondents. It is also possible that the relationship between patients and their physicians were not in favor of male physicians in Germany, and temporal shifts in patient-physician relationships over the time frames examined could also impact the result differences (2012 vs 2012-2018 in our data).

Implications

Our findings have important implications for health care researchers, professionals, and policy makers. First, the empirical evidence of web-based reviews is less favorable toward female physicians, after accounting or controlling for several clinical aspects (including specialty and Medicare patient risk), and temporal effects should inform health care professionals and policy makers that patients' opinions are consistently more favorable toward male physicians than toward female physicians. This cannot be overlooked even though the magnitude of the effect of gender on web-based reviews is not sizable.

Policy and Design Suggestions

Gender bias in reviews has been reported across multiple domains, including academia. Murray et al [66] found that male faculty tended to receive higher ratings for overall teaching quality than female faculty, and Turrentine et al [67] and Rojek et al [68] found implicit bias in the narrative evaluations, with a bias toward men receiving more superlative praise. Studies have shown that measures can be taken to help reduce gender bias in reviews. Peterson et al [69] found that simply informing students of potential gender biases can have significant effects on the evaluation of female instructors, and Rivera and Tilcsik [70] found that by changing the rating scale from a 10-point to a 6-point rating system, gender bias can be reduced.

Large societal-level aspects may also be in effect; however, that would seemingly be very hard to account for within a single portal. Sprague and Massoni [71] found that male teachers are more likely to be held to an entertainer standard, whereas female teachers are held to a nurturer standard. These biases are formed throughout an individual's life, and therefore, are harder to adjust for, even when directly informing users of the potential

for bias. By leveraging the lessons learned from gender bias studies, web-based physician review sites could help to mitigate, but not eliminate, gender bias within their systems.

Concentrated efforts to educate and inform patients about female physicians' competence are needed. This can help to reduce implicit bias among patients toward the competence of female physicians compared with their male counterparts. These websites serve as an important resource for both reviewers and readers of the reviews, and the information needs to flow well. At the same time, readers of the reviews may be served better if the reviewers are asked to provide opinions about physicians of different genders before they provide a review for a physician. To solicit reviewers' predisposed opinions about physicians of different genders, the questions can be framed in a manner that does not make the reviewers feel that they are being investigated for their opinions. After collecting their opinions on this issue, the websites may consider filtering the reviews provided by reviewers with an overt bias against physicians of one gender. The question of how to design the website to reduce the possible gender bias is complex and requires serious thought and consideration from both researchers and website designers. By leveraging previous research efforts targeted at informing users of bias potential, review portals can better collect and present information about physicians.

Limitations

Our study has a few limitations. First, we constructed our patient risk scores using the HCC risk score from Medicare data. Although Medicare is among the largest health care payers or insurers in the United States, further studies can attempt to validate the findings of our study using clinical data from other insurers. For instance, a significant proportion of the patient population in the United States has insurance from private insurers. Future studies can attempt to validate our findings by constructing clinical variables, such as risk scores, using clinical data from one or more private insurers. Second, we focused on the physician data from Alabama. Although it is 1 state, it provides a good mix of rural and urban counties. Future studies could extend this work to other states and compare the findings across a broader set of patients and health care providers.

Future Studies

The findings of this study suggest that gender bias in web-based reviews needs to be examined more closely. Additional studies that identify factors impacting this gender bias could help us develop strategies to mitigate gender bias in web-based reviews. Given the shortage of health care providers and the need for a robust and diverse health care workforce, such studies can help not only the service providers but also policy makers, educators, and administrators. If the administrators of hospitals and clinics are made aware of this bias and acknowledge it accordingly, institutional changes can be implemented to support and empower women to take up more leadership roles in clinical settings. As Sandberg [72] points out in her *New York Times* best seller, as fewer women are in leadership roles than men, it can be challenging for junior women to have mentorship opportunities. A possible solution to this problem could be the performance evaluations of male leadership personnel to include the number of women mentored and focused initiatives and

incentive opportunities for women to take on pathways to leadership roles.

These focused efforts can provide a strong signal to patients about the competence of female physicians and, in turn, increase their confidence in the care provided by female physicians. This can further help to improve the overall care delivered to patients,

as the increase in patients' confidence can improve their communication with physicians, irrespective of the physicians' gender. However, an open research question is whether the bias observed in web-based physician reviews is also observable in offline physician surveys. To examine this question, studies that compare reviews of male and female physicians in web-based and offline media need to be conducted.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Tables depicting the results of additional analysis including robustness checks.

[\[DOCX File , 39 KB-Multimedia Appendix 1\]](#)

References

1. RateMDs. URL: <https://www.ratemds.com> [accessed 2021-06-28]
2. Vitals. URL: <https://www.vitals.com/> [accessed 2021-06-28]
3. Healthgrades. URL: <https://www.healthgrades.com/> [accessed 2021-06-28]
4. Terlutter R, Bidmon S, Röttl J. Who uses physician-rating websites? Differences in sociodemographic variables, psychographic variables, and health status of users and nonusers of physician-rating websites. *J Med Internet Res* 2014 Mar 31;16(3):e97. [doi: [10.2196/jmir.3145](https://doi.org/10.2196/jmir.3145)] [Medline: [24686918](https://pubmed.ncbi.nlm.nih.gov/24686918/)]
5. Binary Fountain. 2018. URL: <https://www.binaryfountain.com/wp-content/uploads/2018/09/Binary-Fountain-2018-Consumerism-Survey-Infographic.pdf> [accessed 2021-06-24]
6. Detz A, López A, Sarkar U. Long-term doctor-patient relationships: patient perspective from online reviews. *J Med Internet Res* 2013 Jul 02;15(7):e131 [FREE Full text] [doi: [10.2196/jmir.2552](https://doi.org/10.2196/jmir.2552)] [Medline: [23819959](https://pubmed.ncbi.nlm.nih.gov/23819959/)]
7. Emmert M, Halling F, Meier F. Evaluations of dentists on a German physician rating website: an analysis of the ratings. *J Med Internet Res* 2015 Jan 12;17(1):e15 [FREE Full text] [doi: [10.2196/jmir.3830](https://doi.org/10.2196/jmir.3830)] [Medline: [25582914](https://pubmed.ncbi.nlm.nih.gov/25582914/)]
8. 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 Feb 24;14(1):e38 [FREE Full text] [doi: [10.2196/jmir.2003](https://doi.org/10.2196/jmir.2003)] [Medline: [22366336](https://pubmed.ncbi.nlm.nih.gov/22366336/)]
9. Hao H, Zhang K. The voice of Chinese health consumers: a text mining approach to web-based physician reviews. *J Med Internet Res* 2016 May 10;18(5):e108 [FREE Full text] [doi: [10.2196/jmir.4430](https://doi.org/10.2196/jmir.4430)] [Medline: [27165558](https://pubmed.ncbi.nlm.nih.gov/27165558/)]
10. Hong YA, Liang C, Radcliff TA, Wigfall LT, Street RL. What do patients say about doctors online? A systematic review of studies on patient online reviews. *J Med Internet Res* 2019 Apr 08;21(4):e12521 [FREE Full text] [doi: [10.2196/12521](https://doi.org/10.2196/12521)] [Medline: [30958276](https://pubmed.ncbi.nlm.nih.gov/30958276/)]
11. Kadry B, Chu LF, Kadry B, Gammas D, Macario A. Analysis of 4999 online physician ratings indicates that most patients give physicians a favorable rating. *J Med Internet Res* 2011 Nov 16;13(4):e95 [FREE Full text] [doi: [10.2196/jmir.1960](https://doi.org/10.2196/jmir.1960)] [Medline: [22088924](https://pubmed.ncbi.nlm.nih.gov/22088924/)]
12. Liu JJ, Matelski JJ, Bell CM. Scope, breadth, and differences in online physician ratings related to geography, specialty, and year: observational retrospective study. *J Med Internet Res* 2018 Mar 07;20(3):e76 [FREE Full text] [doi: [10.2196/jmir.7475](https://doi.org/10.2196/jmir.7475)] [Medline: [29514775](https://pubmed.ncbi.nlm.nih.gov/29514775/)]
13. Segal J, Sacopulos M, Sheets V, Thurston I, Brooks K, Puccia R. Online doctor reviews: do they track surgeon volume, a proxy for quality of care? *J Med Internet Res* 2012 Apr 10;14(2):e50 [FREE Full text] [doi: [10.2196/jmir.2005](https://doi.org/10.2196/jmir.2005)] [Medline: [22491423](https://pubmed.ncbi.nlm.nih.gov/22491423/)]
14. Huerta TR, Hefner JL, Ford EW, McAlearney AS, Menachemi N. Hospital website rankings in the United States: expanding benchmarks and standards for effective consumer engagement. *J Med Internet Res* 2014 Feb 25;16(2):e64 [FREE Full text] [doi: [10.2196/jmir.3054](https://doi.org/10.2196/jmir.3054)] [Medline: [24568892](https://pubmed.ncbi.nlm.nih.gov/24568892/)]
15. Emmert M, Sander U, Pisch F. Eight questions about physician-rating websites: a systematic review. *J Med Internet Res* 2013 Feb 01;15(2):e24 [FREE Full text] [doi: [10.2196/jmir.2360](https://doi.org/10.2196/jmir.2360)] [Medline: [23372115](https://pubmed.ncbi.nlm.nih.gov/23372115/)]
16. 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](https://doi.org/10.2196/jmir.5889)] [Medline: [27644135](https://pubmed.ncbi.nlm.nih.gov/27644135/)]
17. Murphy GP, Awad MA, Osterberg EC, Gaither TW, Chumnarnsongkhroh T, Washington SL, et al. Web-based physician ratings for California physicians on probation. *J Med Internet Res* 2017 Aug 22;19(8):e254 [FREE Full text] [doi: [10.2196/jmir.7488](https://doi.org/10.2196/jmir.7488)] [Medline: [28830852](https://pubmed.ncbi.nlm.nih.gov/28830852/)]
18. Lu SF, Rui H. Can we trust online physician ratings? Evidence from cardiac surgeons in Florida. *Manag Sci* 2018 Jun;64(6):2557-2573. [doi: [10.1287/mnsc.2017.2741](https://doi.org/10.1287/mnsc.2017.2741)]

19. 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](https://doi.org/10.2196/jmir.6612)] [Medline: [27965191](https://pubmed.ncbi.nlm.nih.gov/27965191/)]
20. Saifee DH, Bardhan IR, Lahiri A, Zheng ZE. Adherence to clinical guidelines, electronic health record use, and online reviews. *J Manag Inf Syst* 2019 Oct 09;36(4):1071-1104. [doi: [10.1080/07421222.2019.1661093](https://doi.org/10.1080/07421222.2019.1661093)]
21. Saifee DH, Zheng ZE, Bardhan IR, Lahiri A. Are online reviews of physicians reliable indicators of clinical outcomes? A focus on chronic disease management. *Inf Syst Res* 2020 Dec;31(4):1282-1300. [doi: [10.1287/isre.2020.0945](https://doi.org/10.1287/isre.2020.0945)]
22. Emmert M, Meier F, Pisch F, Sander U. Physician choice making and characteristics associated with using physician-rating websites: cross-sectional study. *J Med Internet Res* 2013 Aug 28;15(8):e187 [FREE Full text] [doi: [10.2196/jmir.2702](https://doi.org/10.2196/jmir.2702)] [Medline: [23985220](https://pubmed.ncbi.nlm.nih.gov/23985220/)]
23. Grabner-Kräuter S, Waiguny MK. Insights into the impact of online physician reviews on patients' decision making: randomized experiment. *J Med Internet Res* 2015 Apr 09;17(4):e93 [FREE Full text] [doi: [10.2196/jmir.3991](https://doi.org/10.2196/jmir.3991)] [Medline: [25862516](https://pubmed.ncbi.nlm.nih.gov/25862516/)]
24. Han X, Qu J, Zhang T. Exploring the impact of review valence, disease risk, and trust on patient choice based on online physician reviews. *Telemat Inform* 2019 Dec;45:101276. [doi: [10.1016/j.tele.2019.101276](https://doi.org/10.1016/j.tele.2019.101276)]
25. Li S, Lee-Won RJ, McKnight J. Effects of online physician reviews and physician gender on perceptions of physician skills and primary care physician (PCP) selection. *Health Commun* 2019 Oct;34(11):1250-1258. [doi: [10.1080/10410236.2018.1475192](https://doi.org/10.1080/10410236.2018.1475192)] [Medline: [29792519](https://pubmed.ncbi.nlm.nih.gov/29792519/)]
26. Brohan M. Even with a referral, 80% of patients still check out doctors online. *Digital Commerce 360*. 2018. URL: <https://www.digitalcommerce360.com/2018/05/22/even-with-a-referral-80-of-patients-still-check-out-doctors-online/> [accessed 2022-04-04]
27. Murray E, Lo B, Pollack L, Donelan K, Catania J, Lee K, et al. The impact of health information on the internet on health care and the physician-patient relationship: national U.S. survey among 1.050 U.S. physicians. *J Med Internet Res* 2003;5(3):e17 [FREE Full text] [doi: [10.2196/jmir.5.3.e17](https://doi.org/10.2196/jmir.5.3.e17)] [Medline: [14517108](https://pubmed.ncbi.nlm.nih.gov/14517108/)]
28. Roter DL, Hall JA, Aoki Y. Physician gender effects in medical communication: a meta-analytic review. *JAMA* 2002 Aug 14;288(6):756-764. [doi: [10.1001/jama.288.6.756](https://doi.org/10.1001/jama.288.6.756)] [Medline: [12169083](https://pubmed.ncbi.nlm.nih.gov/12169083/)]
29. Shin DW, Roter DL, Roh YK, Hahm SK, Cho B, Park HK, Board Certification Committee of The Korean Academy of Family Medicine. Physician gender and patient centered communication: the moderating effect of psychosocial and biomedical case characteristics. *Patient Educ Couns* 2015 Jan;98(1):55-60. [doi: [10.1016/j.pec.2014.10.008](https://doi.org/10.1016/j.pec.2014.10.008)] [Medline: [25457177](https://pubmed.ncbi.nlm.nih.gov/25457177/)]
30. Nicolai J, Demmel R. The impact of gender stereotypes on the evaluation of general practitioners' communication skills: an experimental study using transcripts of physician-patient encounters. *Patient Educ Couns* 2007 Dec;69(1-3):200-205. [doi: [10.1016/j.pec.2007.08.013](https://doi.org/10.1016/j.pec.2007.08.013)] [Medline: [17942269](https://pubmed.ncbi.nlm.nih.gov/17942269/)]
31. Hall JA, Roter DL, Blanch-Hartigan D, Mast MS, Pitegoff CA. How patient-centered do female physicians need to be? Analogue patients' satisfaction with male and female physicians' identical behaviors. *Health Commun* 2015;30(9):894-900. [doi: [10.1080/10410236.2014.900892](https://doi.org/10.1080/10410236.2014.900892)] [Medline: [25175277](https://pubmed.ncbi.nlm.nih.gov/25175277/)]
32. 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 Jul 30;22(7):e14455 [FREE Full text] [doi: [10.2196/14455](https://doi.org/10.2196/14455)] [Medline: [32729844](https://pubmed.ncbi.nlm.nih.gov/32729844/)]
33. Weisman CS, Teitelbaum MA. Physician gender and the physician-patient relationship: recent evidence and relevant questions. *Soc Sci Med* 1985;20(11):1119-1127. [doi: [10.1016/0277-9536\(85\)90189-3](https://doi.org/10.1016/0277-9536(85)90189-3)] [Medline: [3895448](https://pubmed.ncbi.nlm.nih.gov/3895448/)]
34. Roter DL, Hall JA. Physician gender and patient-centered communication: a critical review of empirical research. *Annu Rev Public Health* 2004;25:497-519. [doi: [10.1146/annurev.publhealth.25.101802.123134](https://doi.org/10.1146/annurev.publhealth.25.101802.123134)] [Medline: [15015932](https://pubmed.ncbi.nlm.nih.gov/15015932/)]
35. Hall JA, Blanch-Hartigan D, Roter DL. Patients' satisfaction with male versus female physicians: a meta-analysis. *Med Care* 2011 Jul;49(7):611-617. [doi: [10.1097/MLR.0b013e318213c03f](https://doi.org/10.1097/MLR.0b013e318213c03f)] [Medline: [21478782](https://pubmed.ncbi.nlm.nih.gov/21478782/)]
36. Hall JA, Roter DL. Do patients talk differently to male and female physicians? A meta-analytic review. *Patient Educ Couns* 2002 Dec;48(3):217-224. [doi: [10.1016/s0738-3991\(02\)00174-x](https://doi.org/10.1016/s0738-3991(02)00174-x)] [Medline: [12477606](https://pubmed.ncbi.nlm.nih.gov/12477606/)]
37. Sandhu H, Adams A, Singleton L, Clark-Carter D, Kidd J. The impact of gender dyads on doctor-patient communication: a systematic review. *Patient Educ Couns* 2009 Sep;76(3):348-355. [doi: [10.1016/j.pec.2009.07.010](https://doi.org/10.1016/j.pec.2009.07.010)] [Medline: [19647969](https://pubmed.ncbi.nlm.nih.gov/19647969/)]
38. Fennema K, Meyer DL, Owen N. Sex of physician: patients' preferences and stereotypes. *J Fam Pract* 1990 Apr;30(4):441-446. [Medline: [2324696](https://pubmed.ncbi.nlm.nih.gov/2324696/)]
39. Kerssens JJ, Bensing JM, Andela MG. Patient preference for genders of health professionals. *Soc Sci Med* 1997 May;44(10):1531-1540. [doi: [10.1016/s0277-9536\(96\)00272-9](https://doi.org/10.1016/s0277-9536(96)00272-9)] [Medline: [9160442](https://pubmed.ncbi.nlm.nih.gov/9160442/)]
40. Plunkett BA, Kohli P, Milad MP. The importance of physician gender in the selection of an obstetrician or a gynecologist. *Am J Obstet Gynecol* 2002 May;186(5):926-928. [doi: [10.1067/mob.2002.123401](https://doi.org/10.1067/mob.2002.123401)] [Medline: [12015516](https://pubmed.ncbi.nlm.nih.gov/12015516/)]
41. Hasebrook J, Hahnenkamp K, Buhre WF, de Korte-de Boer D, Hamaekers AE, Metelmann B, et al. Medicine goes female: protocol for improving career options of females and working conditions for researching physicians in clinical medical research by organizational transformation and participatory design. *JMIR Res Protoc* 2017 Aug 02;6(8):e152 [FREE Full text] [doi: [10.2196/resprot.7632](https://doi.org/10.2196/resprot.7632)] [Medline: [28768613](https://pubmed.ncbi.nlm.nih.gov/28768613/)]

42. Thawani A, Paul MJ, Sarkar U, Wallace BC. Are online reviews of physicians biased against female providers? In: Proceedings of the 4th Machine Learning for Healthcare Conference. 2019 Presented at: PMLR '19; August 8-10, 2019; Ann Arbor, MI, USA p. 406-423.
43. Emmert M, Meier F. An analysis of online evaluations on a physician rating website: evidence from a German public reporting instrument. *J Med Internet Res* 2013 Aug 06;15(8):e157 [FREE Full text] [doi: [10.2196/jmir.2655](https://doi.org/10.2196/jmir.2655)] [Medline: [23919987](https://pubmed.ncbi.nlm.nih.gov/23919987/)]
44. Marrero K, King E, Fingeret AL. Impact of surgeon gender on online physician reviews. *J Surg Res* 2020 Jan;245:510-515. [doi: [10.1016/j.jss.2019.07.047](https://doi.org/10.1016/j.jss.2019.07.047)] [Medline: [31446193](https://pubmed.ncbi.nlm.nih.gov/31446193/)]
45. Centers for Medicare & Medicaid Services. 2020. URL: <https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Medicare-Provider-Charge-Data/Downloads/Medicare-Physician-and-Other-Supplier-PUF-Methodology.pdf> [accessed 2022-04-04]
46. Hierarchical condition category coding. American Academy of Family Physicians. 2017. URL: <https://www.aafp.org/family-physician/practice-and-career/getting-paid/coding/hierarchical-condition-category.html> [accessed 2021-06-24]
47. Blei DM, Ng AY, Jordan MI. Latent dirichlet allocation. *J Mach Learn Res* 2003;3:993-1022.
48. Schweinberger M. Topic modeling with R. Language Technology and Data Analysis Laboratory. 2021. URL: <https://slcladal.github.io/topicmodels.html> [accessed 2021-06-05]
49. Guo Y, Barnes SJ, Jia Q. Mining meaning from online ratings and reviews: tourist satisfaction analysis using latent dirichlet allocation. *Tourism Management* 2017 Apr;59(C):467-483. [doi: [10.1016/j.tourman.2016.09.009](https://doi.org/10.1016/j.tourman.2016.09.009)]
50. Hao H, Zhang K, Wang W, Gao G. A tale of two countries: international comparison of online doctor reviews between China and the United States. *Int J Med Inform* 2017 Mar;99:37-44. [doi: [10.1016/j.ijmedinf.2016.12.007](https://doi.org/10.1016/j.ijmedinf.2016.12.007)] [Medline: [28118920](https://pubmed.ncbi.nlm.nih.gov/28118920/)]
51. Tirunillai S, Tellis GJ. Mining marketing meaning from online chatter: strategic brand analysis of big data using latent dirichlet allocation. *J Mark Res* 2014 Aug 01;51(4):463-479. [doi: [10.1509/jmr.12.0106](https://doi.org/10.1509/jmr.12.0106)]
52. Wallace BC, Paul MJ, Sarkar U, Trikalinos TA, Dredze M. A large-scale quantitative analysis of latent factors and sentiment in online doctor reviews. *J Am Med Inform Assoc* 2014;21(6):1098-1103 [FREE Full text] [doi: [10.1136/amiajnl-2014-002711](https://doi.org/10.1136/amiajnl-2014-002711)] [Medline: [24918109](https://pubmed.ncbi.nlm.nih.gov/24918109/)]
53. Zhang Y, Chen M, Huang D, Wu D, Li Y. iDoctor: personalized and professionalized medical recommendations based on hybrid matrix factorization. *Future Gener Comput Syst* 2017 Jan;66:30-35. [doi: [10.1016/j.future.2015.12.001](https://doi.org/10.1016/j.future.2015.12.001)]
54. Feinerer I. Introduction to the tm package text mining in R. Microsoft R Application Network. 2013. URL: <https://mran.microsoft.com/snapshot/2017-12-04/web/packages/tm/vignettes/tm.pdf> [accessed 2022-04-04]
55. Feinerer I, Hornik K. Package 'tm'. The Comprehensive R Archive Network. 2020. URL: <http://cran.rapporter.net/web/packages/tm/tm.pdf> [accessed 2022-04-04]
56. Grün B, Hornik K. topicmodels: an R package for fitting topic models. *J Stat Soft* 2011;40(13):1-30. [doi: [10.18637/jss.v040.i13](https://doi.org/10.18637/jss.v040.i13)]
57. Amado A, Cortez P, Rita P, Moro S. Research trends on big data in marketing: a text mining and topic modeling based literature analysis. *Eur Res Manag Bus Econ* 2018 Jan;24(1):1-7. [doi: [10.1016/j.iiedeen.2017.06.002](https://doi.org/10.1016/j.iiedeen.2017.06.002)]
58. Calheiros AC, Moro S, Rita P. Sentiment classification of consumer-generated online reviews using topic modeling. *J Hosp Mark Manag* 2017 Apr 27;26(7):675-693. [doi: [10.1080/19368623.2017.1310075](https://doi.org/10.1080/19368623.2017.1310075)]
59. Vacher JP, Chave J, Ficetola FG, Sommeria - Klein G, Tao S, Thébaud C, et al. Large - scale DNA - based survey of frogs in Amazonia suggests a vast underestimation of species richness and endemism. *J Biogeogr* 2020 May 18;47(8):1781-1791. [doi: [10.1111/jbi.13847](https://doi.org/10.1111/jbi.13847)]
60. Alabama Physician Workforce Profile. Association of American Medical Colleges. 2019. URL: <https://www.aamc.org/media/37726/download> [accessed 2021-06-24]
61. Berthold HK, Gouni-Berthold I, Bestehorn KP, Böhm M, Krone W. Physician gender is associated with the quality of type 2 diabetes care. *J Intern Med* 2008 Oct;264(4):340-350 [FREE Full text] [doi: [10.1111/j.1365-2796.2008.01967.x](https://doi.org/10.1111/j.1365-2796.2008.01967.x)] [Medline: [18397244](https://pubmed.ncbi.nlm.nih.gov/18397244/)]
62. Dahrouge S, Seale E, Hogg W, Russell G, Younger J, Muggah E, et al. A comprehensive assessment of family physician gender and quality of care: a cross-sectional analysis in Ontario, Canada. *Med Care* 2016 Mar;54(3):277-286. [doi: [10.1097/MLR.000000000000480](https://doi.org/10.1097/MLR.000000000000480)] [Medline: [26765146](https://pubmed.ncbi.nlm.nih.gov/26765146/)]
63. Meier A, Yang J, Liu J, Beitler JR, Tu XM, Owens RL, et al. Female physician leadership during cardiopulmonary resuscitation is associated with improved patient outcomes. *Crit Care Med* 2019 Jan;47(1):e8-13 [FREE Full text] [doi: [10.1097/CCM.0000000000003464](https://doi.org/10.1097/CCM.0000000000003464)] [Medline: [30303843](https://pubmed.ncbi.nlm.nih.gov/30303843/)]
64. Tsugawa Y, Jena AB, Figueroa JF, Orav EJ, Blumenthal DM, Jha AK. Comparison of hospital mortality and readmission rates for Medicare patients treated by male vs female physicians. *JAMA Intern Med* 2017 Feb 01;177(2):206-213 [FREE Full text] [doi: [10.1001/jamainternmed.2016.7875](https://doi.org/10.1001/jamainternmed.2016.7875)] [Medline: [27992617](https://pubmed.ncbi.nlm.nih.gov/27992617/)]
65. Wallis CJ, Ravi B, Coburn N, Nam RK, Detsky AS, Satkunasivam R. Comparison of postoperative outcomes among patients treated by male and female surgeons: a population based matched cohort study. *BMJ* 2017 Oct 10;359:j4366 [FREE Full text] [doi: [10.1136/bmj.j4366](https://doi.org/10.1136/bmj.j4366)] [Medline: [29018008](https://pubmed.ncbi.nlm.nih.gov/29018008/)]

66. Murray D, Boothby C, Zhao H, Minik V, Bérubé N, Larivière V, et al. Exploring the personal and professional factors associated with student evaluations of tenure-track faculty. *PLoS One* 2020;15(6):e0233515 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0233515](https://doi.org/10.1371/journal.pone.0233515)] [Medline: [32492028](#)]
67. Turrentine FE, Dreisbach CN, St Ivany AR, Hanks JB, Schroen AT. Influence of gender on surgical residency applicants' recommendation letters. *J Am Coll Surg* 2019 Apr;228(4):356-65.e3. [doi: [10.1016/j.jamcollsurg.2018.12.020](https://doi.org/10.1016/j.jamcollsurg.2018.12.020)] [Medline: [30630084](#)]
68. Rojek AE, Khanna R, Yim JW, Gardner R, Lisker S, Hauer KE, et al. Differences in narrative language in evaluations of medical students by gender and under-represented minority status. *J Gen Intern Med* 2019 May;34(5):684-691 [[FREE Full text](#)] [doi: [10.1007/s11606-019-04889-9](https://doi.org/10.1007/s11606-019-04889-9)] [Medline: [30993609](#)]
69. Peterson DA, Biederman LA, Andersen D, Ditonto TM, Roe K. Mitigating gender bias in student evaluations of teaching. *PLoS One* 2019;14(5):e0216241 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0216241](https://doi.org/10.1371/journal.pone.0216241)] [Medline: [31091292](#)]
70. Rivera LA, Tilcsik A. Scaling down inequality: rating scales, gender bias, and the architecture of evaluation. *Am Sociol Rev* 2019 Mar 12;84(2):248-274. [doi: [10.1177/0003122419833601](https://doi.org/10.1177/0003122419833601)]
71. Sprague J, Massoni K. Student evaluations and gendered expectations: what we can't count can hurt us. *Sex Roles* 2005 Dec;53(11-12):779-793. [doi: [10.1007/s11199-005-8292-4](https://doi.org/10.1007/s11199-005-8292-4)]
72. Sandberg S, Scovell N. *Lean in: women, work, and the will to lead*. New York, NY, USA: Random House; 2013.

Abbreviations

CMS: Center for Medicare and Medicaid Services

HCC: hierarchical condition category

LDA: Latent Dirichlet Allocation

Edited by A Mavragani; submitted 29.06.21; peer-reviewed by B Metelmann, R Krukowski, D Verran, M Salimi; comments to author 18.07.21; revised version received 02.09.21; accepted 16.03.22; published 08.04.22

Please cite as:

Saifee DH, Hudnall M, Raja U

Physician Gender, Patient Risk, and Web-Based Reviews: Longitudinal Study of the Relationship Between Physicians' Gender and Their Web-Based Reviews

J Med Internet Res 2022;24(4):e31659

URL: <https://www.jmir.org/2022/4/e31659>

doi: [10.2196/31659](https://doi.org/10.2196/31659)

PMID:

©Danish Hasnain Saifee, Matthew Hudnall, Uzma Raja. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 08.04.2022. 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 <https://www.jmir.org/>, as well as this copyright and license information must be included.