Decision Support Systems 166 (2023) 113897

Contents lists available at ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

What ails physician review websites? A study of information needs of patients

Soham Ghosh [a,*], Soumyakanti Chakraborty [a], Narain Gupta [b], Sumanta Basu [a]

[a] Indian Institute of Management, Calcutta, D. H. Road, Kolkata, W Bengal 700104, India
[b] Management Development Institute Gurgaon, Gurugram, Haryana 122007, India

ARTICLE INFO

Keywords:
Physician review websites (PRWs)
Healthcare analytics
Topic modelling
Online reviews
Patient experience
Experimental design

ABSTRACT

Physician Review Websites (PRWs) help users select physicians by providing both structured and unstructured data on physicians and patients' experiences with physicians. In this paper, the adequacy of the available information on PRWs is investigated. An empirical study is conducted to understand the information that patients seek from PRWs through a survey instrument administered to 344 patients. A topic modelling approach is then employed on 666,551 text reviews of patients using data from a popular PRW in India to discover information available on the PRWs. Combining both approaches, it is found that 12 out of 25 topics are either unavailable or only partially available, indicating an information gap. Two topics which constitute the information gap are then utilised to conduct a full factorial experiment on 153 respondents using a Latin Square design. Experimental results demonstrate that adding relevant information improves the physician selection process. Our work contributes to the healthcare IS literature on online physician selection and the role of information availability. The results will help PRWs and hospitals decide how to restructure information on their websites and devise strategies to nudge patients to write reviews highlighting the desired information.

1. Introduction

Online healthcare portals that provide a platform for matching patients with physicians are popular in many countries, for example, RateMDs (USA and Canada), Vitals (USA), Good Doctor Online (China) and Practo (India). These portals, referred to as Physician Review Websites (PRWs) [1], have amassed large userbases. RateMDs is home to over 2.6 million physician reviews, 1.7 million healthcare providers and 161 million users, and Vitals features 1 million physician profiles and >9 million reviews and ratings. The primary services provided by these portals are physician search and booking physician appointments for consultation. Practo, one of the popular PRWs in India, has over 100,000 physicians registered on its platform, whereas Good Doctor Online (haodf.com) in China has >170,000 physicians registered on its platform. One of the key focus areas of these portals is to elicit reviews from patients regarding their consultation experience with the physician. Insightful reviews of physicians expedite the physician selection process and enrich customer (patient) experience.

In the USA, RateMDs, a popular PRW, has been used by 161 million people (~49% of the population). Developing countries such as India show lower numbers in terms of PRW usage. In India, 20 million people use Practo every month (yourstory.com, 2021), with most users belonging to the top seven cities. However, the average population in each of the most populated cities is 7.6 million, and with a total population of 1.38 billion (Worldometers, 2021), there is scope for substantial growth. Ping An Good Doctor has over 11.2 million active users per month in China (Statista, 2021). These numbers are not as high as they could be, given that the population in major Chinese cities ranges between 7.9 and 23.4 million people (chinahighlights.com, 2021).

One possible reason for the less than expected usage of PRWs has been mentioned as the lack of relevant information on these websites [2]. In a qualitative study with 22 participants [3], it has been found that PRWs are unable to cater to the different information needs of health consumers. The authors conclude that the design and content of PRWs need to be modified to address this challenge. A study of 280 patients in Germany [4] indicates that most health consumers consider PRWs important. However, the information inequality is higher for consumers and therefore, most consumers rely less on PRWs for physician selection compared to referrals from friends and other physicians. In a study of approximately 1400 patients in China, it has been found that very few

* Corresponding author at: C-238, Lake View Hostel, Indian Institute of Management Calcutta, Joka, D.H. Road, Kolkata, West Bengal PIN-700104, India.
E-mail addresses: sohamg17@iimcal.ac.in (S. Ghosh), soumyakc@iimcal.ac.in (S. Chakraborty), narain.gupta@mdi.ac.in (N. Gupta), sumanta@iimcal.ac.in (S. Basu).

https://doi.org/10.1016/j.dss.2022.113897
Received 5 April 2022; Received in revised form 3 November 2022; Accepted 4 November 2022
Available online 13 November 2022
0167-9236/

people contribute to the reviews on PRWs [5]. The above studies indicate that prospective patients do not have access to all the relevant information needed to procure healthcare services like physician consultation.

The Covid-19 pandemic has proved to be a proverbial shot in the arm for PRWs. Many PRWs have augmented their usual offerings with teleconsulting services and have consequently witnessed a sharp spike in users during the last year. Between March and May 2020, Practo reported that 80% of its users were first-time telemedicine users, implying that its userbase had grown five times due to the pandemic [6]. In the USA, Blue Cross Blue Shield of Massachusetts reported a peak increase of telehealth claims of over 2000 times, and a year later, it still reported an increased claim rate of 150 times the previous rate [7]. This has also increased the reach of PRWs – in the case of Practo, 44% of the new bookings were reported to come from Tier 2 or 3 cities in India, while previously, such numbers were low. As patients are almost entirely dependent on the information available on the PRWs for teleconsultation services, providing relevant and high-quality information may become even more important for the PRWs in future.

Extant literature has established that the availability of information on websites influences consumers' decision-making process. When consumers buy a product or a service, the decision-making process involves three major phases – intelligence, design, and choice [8]. Information search and information processing, which fall under the intelligence phase, are two steps in the decision-making process influenced by the information available to the consumer [9]. Higher information availability lowers search costs for the consumer, confirming the impact of information available on these websites [10]. Information on PRWs is typically provided in structured (specialisation, fees, etc.) and unstructured formats, as seen in Table 1. As seen in several studies, patient reviews are an important form of unstructured information that impacts the decision-making process. In one study [11], the authors distinguish between reviews from marketer and non-marketer-generated websites and establish the influence of reviews on product choice. Prior research has demonstrated the impact of both quality and quantity of reviews and source credibility on purchase intention [12]. Review credibility plays an important role in the attitude towards the physician [13]. The paper aims to address the gaps in understanding (1) what information is available on PRWs and what information patients seek for the selection of physicians, and (2) how PRWs can address the information gap. The research questions that we attempt to answer in our work are:

RQ1: What information do patients seek while selecting physicians or booking physician appointments online?

RQ2: Is the information sought by patients available on the PRWs as structured or unstructured data?

RQ3: Does providing information that is sought by patients but not available on PRWs positively influence patients' choice of physician?

To address RQ1, a survey instrument is designed to determine the information that patients look for while selecting a physician online on PRWs, and text mining is applied to patient reviews to determine the information available on the PRWs to address RQ2. For RQ3, an experiment is conducted to verify that additional information does indeed help patients and improves the physician selection process. To determine the information that patients seek, a study is conducted in two phases – 1) In the first phase, a survey instrument is developed and modified for the online selection of physicians, and 2) In the second phase, the survey instrument is administered to respondents. The responses are analysed to find the factors patients perceive as important. The patients' reviews of physicians from a leading PRW in India are then analysed to determine the information available on the PRWs.

The data collected includes unstructured data (text reviews), with 666,551 reviews, analysed using a well-established text mining technique, Latent Dirichlet Allocation (LDA). Eight topics that occur throughout the reviews are obtained via this analysis. We generate themes that fit the topics, which indicate aspects of the healthcare experience patients have written about in their reviews. The items in the survey instrument, their availability on PRWs as structured data and whether they match a corresponding topic in the reviews (unstructured data) are cross-referenced. The following information is thus determined – 1) items considered important by patients, according to the results of the empirical study, and 2) the items corresponding to the topics mentioned by patients in reviews that are not available or partially available on the PRWs are ascertained. As the findings from the empirical study are compared with the analysis of the text reviews, a substantial information gap is observed between the information that patients look for and the information present on the PRWs.

We then proceed to address the second research gap using an experimental study. The experimental verification is carried out on a different set of users to corroborate and confirm our findings from the empirical study and therefore requires a multi-method approach [14]. Two items on which patients seek information (determined from the survey) and are unavailable on PRWs are selected. A 2×2 full factorial design is utilised, followed by a Latin Square design with four simulated physician profiles to study the effect of providing these additional items during physician selection. Statistical analysis of the experimental data reveals that addressing the information gap significantly improves the process of physician selection.

Our study makes several contributions. To the best of our knowledge, this is the first study to investigate the gap between the information desired by the prospective patients and the information available on the PRWs. Extant literature has studied the impact of patient reviews on physician selection. In contrast, we study patient reviews to extract the information that patients may find useful for physician selection. We demonstrate the benefits of providing previously uncaptured, relevant information to patients in the physician selection process. The nature of the questions prompts us to adopt a multi-method approach which synthesises the tools and techniques from empirical research, machine learning, and experimental research to identify and address the information gap on these websites. Based on the findings from the study, business managers who are responsible for these websites may decide to modify the form and content of the websites to make them more useful for patients seeking relevant information for physician selection. IS researchers can use the findings to gain more insight into the role of information availability in decision-making in the context of PRWs. This multi-method approach allows us to create a generalised framework for studying online information systems which can be applied to other contexts.

The remainder of the article is organised as follows. In the next section, we present a review of the relevant literature. We then discuss the empirical study and its findings, followed by an explication of our analysis of patients' reviews. The next section discusses the experimental study to verify whether addressing the information gaps improves the physician selection process. We conclude with a discussion on the academic and managerial implications of the study, the limitations, and possible future extensions of our work.

Table 1
Information available on PRWs.

Structured Information	Unstructured Information
Demographic details of the physician	
Timings for consultation	Text reviews by patients (Practo, RateMDs, Vitals)
Specialisation	
Degrees	
Years of Experience	Consult Q&A (Practo)
Ratings	
Hospital/clinic location	Healthfeed (Practo)[a]
Profile Picture	

[a] Healthfeed is a section where physicians can write articles related to their areas of expertise on their profiles.

2. Literature review

2.1. Evaluation of healthcare quality

One component of clinic or hospital service quality is the degree of disparity between the services patients want and the services they can discern when visiting a physician or a hospital [15]. Researchers have developed models for evaluating healthcare service quality, such as the HEALTHQUAL framework [16]. This framework has been used in studies to determine the influence of the dimensions of healthcare service quality on patient quality satisfaction and quality [17]. Most studies have focused on determining the impact of these dimensions on patient satisfaction and loyalty. Studies on factors affecting the choice of a physician based on healthcare service quality dimensions are less common [18]. This study bridges the gap by adopting the dimensions used to determine healthcare quality from multiple frameworks [16,19,20]. We then combine the same with information available on PRWs to determine the factors affecting the choice of a physician in an online setting.

2.2. Impact of online reviews on physician selection

Online reviews have been studied extensively in the Information Systems (IS) literature, with several studies focused on PRWs. A relatively early study [21] established the causal impact of overall ratings on physician demand. The effect of online reviews on the perception of physician skills is examined in [1] using an experimental design. They conclude that through the reviews, perceptions get influenced, which impacts the choice of primary care physician. Recently, service quality features were derived from unstructured patient reviews using econometric models and text mining to predict doctor demand [22]. They demonstrate that their method improves predictive power. The relationship between perceived physician quality and the likelihood of being rated online is also explored in one study where authors find that the "vocal minority" disproportionately influences the choice of physician [23]. The link between patient mortality and the ratings of the surgeons from rating data has been examined by [24] with a focus on cardiac surgeons. Data from a popular PRW is used to find the relationship between patient readmission risk and chronic disease care [25]. The study combines ratings with review data (unstructured information) and builds a predictive model for readmissions risk, future ER visit rate and changes in severity and mortality levels. The study concludes that the impact of patient feedback is limited due to the credence nature of healthcare services. A similar study uses data from a leading PRW to predict a physician's suitability to practice medicine based on online rating data [26]. The effect of online ratings on the choice of physicians is examined through an instrumental variable approach in [27], where the leniency of reviewers while rating other businesses is taken into account. They find that the ratings then positively affect the choice of physicians.

2.3. Information availability and physician selection

Controlled experiments to study the impact of information availability [28] are quite common in the IS literature. Some studies exist in the context of online healthcare [29]. Decision aids in the selection of physicians have been studied using controlled experiments [30]. The clickstream data from one of the largest PRWs in India has been experimentally analysed to determine how online word of mouth influences the patients' decision-making process [31]. The motivation for using online healthcare platforms like PRWs has also been experimentally studied with a focus on developing long-term participation [32]. The key difference between our work and the extant literature is that while the existing papers have discussed the impact, usage, and credibility of the information available on the PRWs, we investigate the adequacy of the information provided on the PRWs. This research thus addresses how PRWs can provide more relevant information to patients to enrich the experience of physician selection.

3. Empirical study

In this section, the empirical study conducted to understand information that patients seek on PRWs while selecting physicians online is discussed in detail. The methodology is based on existing healthcare service quality frameworks, like SERVQUAL and HEALTHQUAL. These frameworks contain widely adopted survey scales consisting of multiple factors or items applicable in the context of the study. Data is collected for the study in two phases to answer RQ1. The workings of each phase of the empirical study are highlighted in Fig. 1.

For the first phase of the empirical study, a survey instrument is developed to identify the factors those patients consider while booking a physician's appointment online. We refer to physicians as doctors in the survey instrument, as this term is more widely used in India. The constructs Assurance, Doctor's Credibility [33], Efficiency, Empathy, Tangibles, and Degree of Care Improvement [16] are adapted from the extant literature. A few patients and physicians are also consulted on the important factors. The constructs are chosen based on the information present on the physicians' profiles on the PRWs, along with the feedback from the patients and physicians. For example, it is noted that online reviews and referrals are important factors. Thus, items related to reviews, ratings and physician's experience are added in the construct, Doctor's Credibility. Items related to waiting time and availability of booking slots are added to the construct Tangibles. The constructs are modified to include the items appropriately and create the survey instrument.

The dimensions of the constructs are measured using a seven-point Likert scale, where "1" is defined as "Somewhat important" to "7" as "Very important". The construct, Degree of Care Improvement, is taken as the dependent variable of choice here, indicating the healthcare service quality. Therefore, we can assume that a patient who receives high-quality healthcare services, i.e., a patient with a high degree of care improvement, will recommend the physician to others. This is in line with previous literature, where healthcare service quality has impacted the willingness of a patient to recommend a physician [34]. The survey consists of 50 items, including 5 demographic items (age, gender, income, occupation, and education) and 4 general questions. Scholars typically recommend a minimum of 1:5 sample size for factor analysis [35]. For models with 5 or fewer constructs, with communalities higher than 0.6, a sample size of 100 is considered sufficient.

The scales are based on the physician's inputs (see Table A.1 for added items in italics). A pilot survey and Exploratory Factor Analysis (EFA) are conducted to confirm the item groupings in the instrument and to establish the contextual validity of the measures. Eighty-four responses are collected on the Qualtrics platform and analysed to investigate the possible factorial structure. Ten responses have >50% of the responses missing and, as a result, are considered incomplete and removed. Little's MCAR (Missing Completely at Random) test ($p > 0.05$) shows that the remaining missing values have no specific pattern and as a result, are replaced with the series mean. The EFA is carried out using maximum likelihood extraction (MLE) in SPSS 25, consistent with the literature on the identification of latent dimensions. Bartlett's Test of Sphericity ($p < 0.05$), which indicates a high correlation between at least two variables in the model and the Kaiser-Meyer-Olkin (KMO) test (0.652), which refers to sampling adequacy, are used for preliminary testing. Cronbach's alpha is 0.7 and above for four of the five constructs. The reliability and validity of the constructs are tested and found satisfactory. Some items are altered based on the lower loadings and communalities and cross-loadings. For example, the item "How many years of experience does the doctor have" is simplified as "Years of experience of doctor". The EFA helped improve the survey instrument for the final data collection phase.

The second phase involves collecting patients' opinions using the final survey instrument based on the learnings from the first phase of

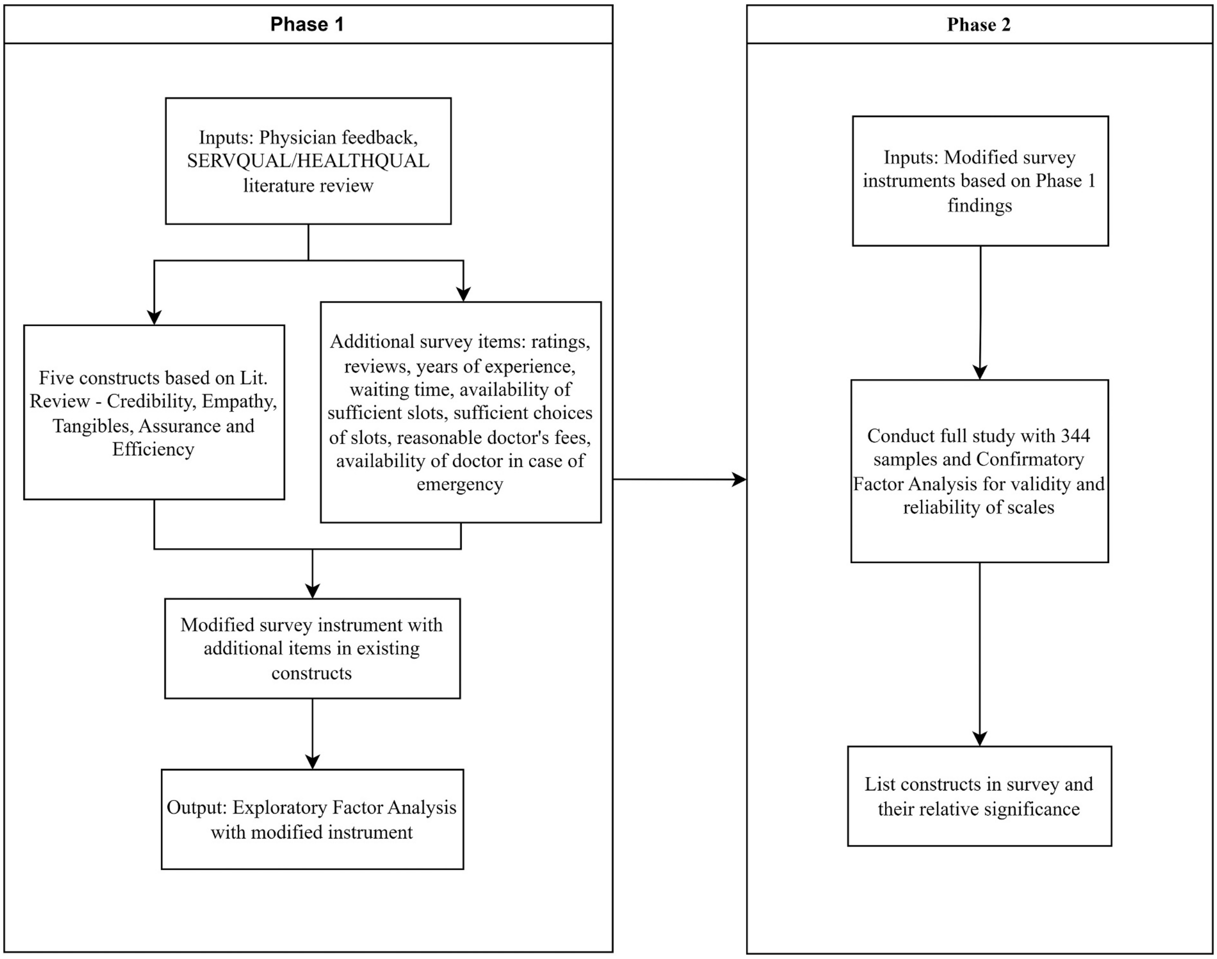

Fig. 1. Two phases of the empirical study.

data collection. As the popularity of PRWs has soared during the Covid 19 pandemic, we realised that it would be difficult to ensure that respondents do not have any prior exposure to PRWs. Therefore, responses have been collected only from patients undergoing treatment to reduce the bias in the responses from any previous exposure to PRWs. Three hundred eighty patients' responses are collected, with 52% male respondents, and most respondents lie within the age range of 25–44 (72.7%). 51.74% of the respondents reported that they had previously booked a physician online. Outliers are removed using the Mahalanobis distance technique [36], where 45 out of 389 responses are removed with a threshold of $p < 0.05$. Unengaged respondents are tracked by calculating the standard deviation of responses with a threshold of 0.5. After pre-processing, a valid sample of 344 is finally used for developing the model, with a response rate of 90.53%.

3.1. Reliability and validity

Confirmatory Factor Analysis (CFA) is carried out using the SmartPLS v3.0 software to check the validity and reliability of the scales used in the study. The study is exploratory, and the constructs used for the scales are formative, leading to the choice of a component-based PLS method. Initially, the constructs Tangibles and Credibility are observed to have an Average Variance Extracted (AVE) of <0.5 due to the low loadings of some items. Scale purification is carried out by removing the items with low loadings [37] – e.g., "Availability of preferred time slots" and "The doctor's specialisation", which improves the AVE to the threshold of 0.49. The removed items are marked with a star (*) in Table A.1 in the appendix. The Fornell-Larcker criterion [38] is applied to test the validity and reliability of the constructs. The results are summarised in Table 2. The convergent validity is confirmed by observing the Composite Reliability (> 0.7), Cronbach's Alpha (>0.7) and AVE (> 0.49), which are in line with the recommended thresholds for all the constructs. Tangibles has a Cronbach's alpha value of 0.694.

For the construct, Tangibles, we note that the composite reliability is >0.7, inter-item correlations for the construct are >0.3, and the discriminant validity of the construct holds. The discriminant validity is ensured by comparing the square root of AVE and shared correlations among the constructs, as shown in Table 3. The square root of AVE (in the diagonal of Table 3) is noted to be greater than the respective bivariate correlations of the constructs. The model statistics are presented in Table A.1.

3.2. Common method bias

The possibility of Common Method Bias (CMB) in this research is low. The patients have no incentive to inflate the dependent variable, so

Table 2
Reliability and Validity test results.

Constructs	Cronbach's Alpha	Composite Reliability	Average Variance Extracted
Assurance	0.816	0.880	0.648
Credibility	0.721	0.830	0.554
Efficiency	0.856	0.894	0.586
Empathy	0.874	0.904	0.576
Tangibles	0.694	0.812	0.526

Table 3
Bivariate correlations of the constructs.

Correlations	Assurance	Credibility	Efficiency	Empathy	Tangibles
Assurance	0.805	0.229	0.651	0.450	0.647
Credibility	0.229	0.744	0.286	0.294	0.345
Efficiency	0.651	0.286	0.766	0.512	0.583
Empathy	0.450	0.294	0.512	0.759	0.550
Tangibles	0.647	0.345	0.583	0.550	0.725

the possibility of spurious correlation is minimal. Nonetheless, process and statistical control are carried out to address this issue. CMB is traditionally tested using Harman's single factor test. The variance extracted is 35.42% (<50%), indicating that CMB is not a problem in this study. The main effects of each construct on the Degree of Care Improvement are tested using a Structural Equation Model (SEM). Four out of five constructs impact the Degree of Care Improvement with *p*-values <0.05, which largely supports the motivation for using these constructs to understand patient expectations regarding physician selection. The results for the same are given in Table 4. The coefficients and loadings for the SEM is provided in Fig. A.1 in the appendix.

3.3. Ranking

A ranking of items used in the survey instrument is developed. Multiple Linear Regression (MLR) is used on the data using SPSS 25, with the dependent variable being the mean of the items constituting the construct, Degree of Care Improvement. The independent variables consist of the items in the constructs Assurance, Doctor's Credibility, Efficiency, Empathy, and Tangibles, with 25 items. Since the independent variables are on the same scale, and the Variance Inflation Factor (VIF) is <5 for each [35], the standardised beta coefficient values are used to determine the importance of the independent variables in the model. The model has an R^2 value of 0.68 and an adjusted R^2 value of 0.655, indicating the model's goodness of fit. Appendix B shows the ranks of the items obtained with respect to the standardised beta coefficients.

3.4. Findings from the empirical study

The empirical study helps identify the factors, or items, that patients consider important while booking a physician's online appointment, thus answering the first research question. Table A.1 shows the items used for all the constructs in the model. It is found that most factors are not available as structured information on the PRWs (see Table 6). However, a major source of information regarding physicians is online reviews by patients, which are unstructured and textual. Therefore, information on some of the items in the survey instrument may be available as part of patient reviews, necessitating a thorough analysis of the reviews.

4. Analysis of patient reviews

In addition to structured information on the websites, unstructured text reviews are a rich source of information for patients [22]. Thus, to further investigate the adequacy of information available on PRWs, we go beyond the structured summary information available on physicians' profiles and mine the reviews to extract topics from unstructured text. Most popular PRWs have a section dedicated to patient reviews after the consultation. In the context of the study, reviews are collected from a PRW and analysed to obtain topics. Each physician for whom reviews have been analysed has verified medical registration, and all the reviews are posted only by verified patients. The PRW has a verification process which ensures that only patients with valid contact numbers (or email IDs) can book an appointment, and it does not allow the registered patient to post the review until the consultation is over. Therefore, the data does not have fake reviews by unverified patients or reviews of unverified doctors. Latent Dirichlet Allocation (LDA) [39] is selected for analysis, which is a popular topic modelling algorithm in the IS literature [40]. This allows the identification of topics from the keywords used in reviews. LDA has become a popular methodology in the healthcare information systems space and has often been combined with econometric models to study the effect of waiting times on physician ratings [41] or to understand the determinants of patient satisfaction [42].

Table 4
SEM results.

Independent Variables	Coeff.	t-values
Assurance- > Degree of Care Improvement	0.259***	4.20
Doctor's Credibility - > Degree of Care Improvement	0.018	0.51
Efficiency - > Degree of Care Improvement	0.251***	4.60
Empathy - > Degree of Care Improvement	0.335***	5.80
Tangibles - > Degree of Care Improvement	0.132*	2.18
Adjusted R-Sq (Degree of Care Improvement)	65.1%	
* p < 0.05, **, *p* < 0.01, *** *p* < 0.001		

Unstructured data (textual reviews) is collected from a major PRW in India. We obtain 666,551 reviews for seven major cities in India. The R statistical software is used for analysis, where multiple rounds of data cleaning are carried out. Short reviews are removed (< 4 characters), and standard pre-processing procedures for text analysis, like conversion to lower case and removing punctuation, standard stopwords, and numbers from the text data, are carried out. Based on the algorithm's output, a list of custom stopwords is added to filter out uninformative words in the context of the study.

LDA is defined as "a three-level hierarchical Bayesian model, in which each item of a collection is modelled as a finite mixture over an underlying set of topics" [39]. It is considered a generative probabilistic model, which indicates the output as a set of probabilities associated with each topic appearing in the corpus of reviews. The assumption is that topics generate words through conditional distributions, which are infinitely exchangeable within a document. A set of topics is thus matched with several words associated with each topic. LDA is advantageous in ascertaining topics from a large set of documents due to 1) the ability to scan many documents automatically, rather than manually, which saves time, and 2) the topic generation happens automatically, i.e., the topic keywords are generated through statistical modelling, rather than through manual inspection.

The first step involves deciding the number of topics for evaluation (K). The LDA algorithm is tuned with multiple topic numbers from 5 to 15, and the change in diagnostic values for different topic numbers is compared, as shown in Fig. 2. Semantic coherence measures the co-occurrence of words within the documents to ensure selected keywords belong to a single concept, thus preserving the interpretability or quality of the topic [43]. Held-out likelihood estimates the probability of keywords appearing in documents (excluded or held-out of the training) to indicate the generalisation capability of the topic model [44]. A word's exclusivity to a topic is its usage rate relative to a set of comparison topics [45]. The held-out likelihood, residuals and exclusivity keep improving as the number of topics increases. However, the semantic coherence decreases beyond ten topics. Combining these diagnostics, K = 10 is selected as the ideal number of topics for the study. The second step is generating a list of topics for evaluation. After selecting the number of topics, an LDA model is generated and applied to the pre-processed and cleaned document corpus. The topic groups are determined using the Score metric, inspired by the popular term frequency-inverse document frequency (TF-IDF) score of vocabulary terms used in information retrieval [46]. Three physicians are consulted and asked to go through the keywords for each of the 10 topics and assign a theme to each topic. The physicians assign a theme to the topics independently, and the inter-rater reliability is calculated to be 0.79, indicating a high degree of agreement among the three physicians. We

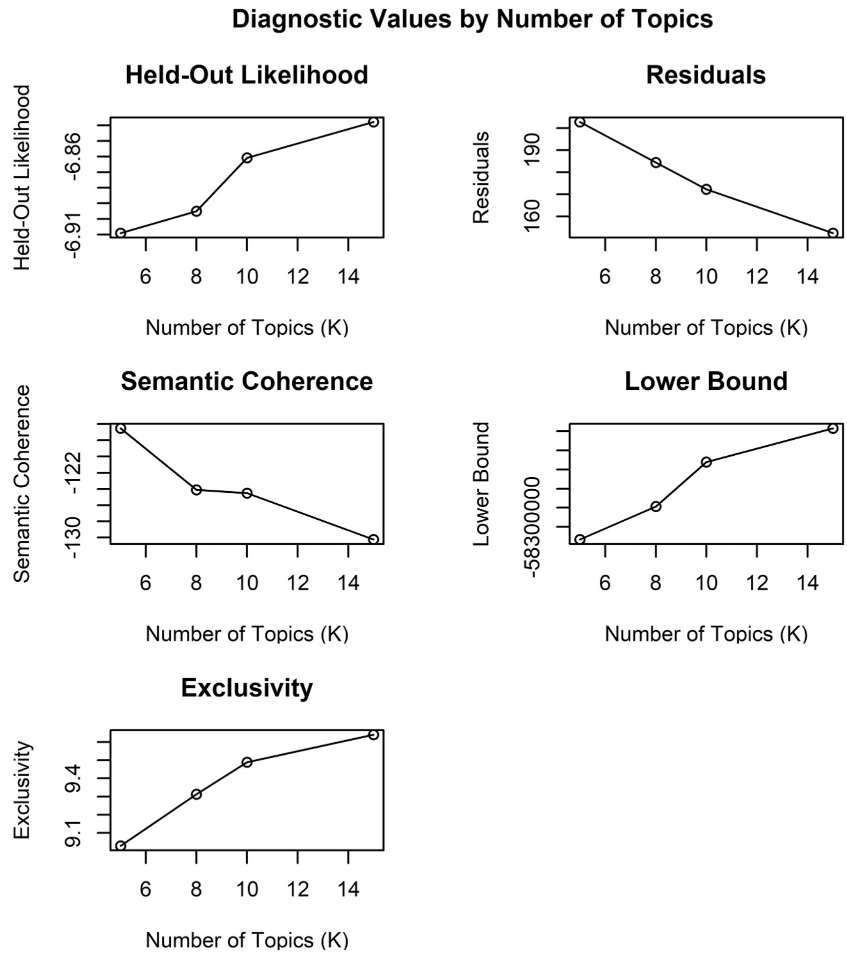

Fig. 2. LDA diagnostic values for topic selection.

then note that three of the topics represent the responsiveness/helpfulness of physicians. We represent this as one topic, with the result being eight topics, which are presented in Table 5.

Table 5
Topics and associated keywords in reviews.

Topic No.	Topics	Keywords
1	Hospital administration	nice, staff, clinic, excellent, friendly, experience, clean, service, professional, treatment
2	Treatment experience – Gynaecology	pregnancy, baby, delivery, normal, wife, god, life, support, positive, healthy
3	Treatment experience – Dental	dental, tooth, teeth, root, canal, procedure, dentist, pain, painless, extraction
4	Responsiveness/ helpfulness of physicians	listens, explains, patiently, queries, friendly, questions, kids, friendly, recommend, satisfied, friends, treatment, explained, family, patients, understand, understanding, health, explain, explains, friendly, patients, doubt, approach, soft, nature, polite, issues, explanation, understands
5	Treatment experience –dermatology	skin, treatment, hair, acne, happy, result, month, medicines, started, laser
6	Waiting time	appointment, time, wait, told, waiting, tests, consultation, tests, reports, hospital
7	Treatment experience –ophthalmology	issue, explained, eye, friendly, treatment, disease, listened, health, explanation, satisfied
8	Treatment experience – surgery	surgery, pain, suffering, days, mother, severe, day, suggested, operation, knee

Topic 1 talks about the hospital's administration regarding insurance and other services. The keywords indicate positive experiences, with keywords like "excellent" - highlighting the importance of proper administrative staff behaviour in reviews. Topics 2, 3, 5, 7 and 8 relate to the patients' treatment experience, divided into gynaecology, dental, dermatology, ophthalmology, and surgery. This indicates that a major portion of the reviews refers to specific experiences – e.g., a patient going for surgery would talk specifically about surgery and associated topics like the surgeon's skill. Topic 4 is about the physicians, and their responsiveness or helpfulness, with positive keywords indicating happy patients. Topic 6 highlights issues regarding the waiting time at the reception of the physician's office or getting an appointment with the physician.

The survey items from the empirical study are matched with the generated topics from the LDA model and categorised. For example, the items "The doctor gave an adequate amount of time to me" and "The doctor diagnosed my medical problem satisfactorily" are categorised under the topic "Treatment Experience". Similarly, the survey item "Waiting time at the doctor's office" is categorised under "Waiting Time". As the study aims to highlight the information gap on PRWs, the topic reviews can be better mapped with the items at the individual level rather than the construct level.

The review topics in Table 5 are very specific in nature, while the construct labels in Table 4 are more generic. In order to compare them appropriately, topics are combined where feasible - topics 2,3,5 and 7 are combined into one overarching topic– Treatment Experience. Table 5 and the items in Table A.1 are combined to list the items in Table 6. The availability of information regarding the items on the PRWs is then verified. The availability of information is classified as Y

Table 6
Comparison between survey items and review topics.

Constructs (1)	Availability on PRWs as structured information (2)	Availability on PRWs as unstructured information (3)	Availability on PRWs from both structured and unstructured information (4)
F1. Credibility of the doctor			
The doctor's ratings from the website (s)	Y	N	Y
The doctor's reviews from the website (s)	Y	N	Y
Years of experience of doctor	Y	N	Y
The doctor's achievements: – quality certification, accreditations, awards received, membership of medical societies	Y	N	Y
F2. Doctor's interaction with the patient – Empathy			
The doctor behaved/behaves politely with me	N	N	N
The doctor gave/gives an adequate amount of time to me	N	N	N
The doctor listened/listens to my symptoms	N	N	N
The doctor explained/explains my medical problem to me	N	N	N
The doctor was understanding of my financial or time constraints/limitations	N	P	P
The doctor spent/spends adequate time to understand my problem	N	Y	Y
The doctor diagnosed/diagnoses my medical problem satisfactorily	N	Y	Y
F3. Information about the doctor and hospital – Tangibles			
The expertise of the doctor	Y	N	Y
Overall cleanliness of the hospital/clinic	N	N	N
Neatness of the doctor	Y	N	Y
Waiting time at the doctor's office	N	Y	Y
F4. Assurance regarding the medical staff – Assurance			
Medical care was/is performed carefully	N	Y	Y
The hospital staff was/is always honest and genuine	N	Y	Y
The hospital staff respected/respects my feelings	N	Y	Y
The hospital staff was/is friendly and cheerful	N	P	P
F5. Treatment of the doctor – Efficiency			
The doctor provided/provides necessary medications to me	N	P	P
The doctor provided/provides necessary diagnostic tests to me	N	P	P
The doctor provided/provides necessary procedures to me	N	P	P
The cost of medication was/is reasonable	N	N	N
The doctor's fee was/is reasonable	Y	N	Y
The doctor was/is available for consultation in an emergency	N	N	N

(Available) / N (Not Available) / P (Partially Available) by proceeding as follows.

4.1. Findings from the analysis of patient reviews

We first determine 7 items for which information is present on the PRW as structured information, for example, the doctor's ratings, the doctor's years of experience, consultation fees, etc. These items have been marked as Y in Column (2), Table 6, with the others being marked N. Three doctoral students at a reputed business school in India are then asked to go through the topics in Table 5. The doctoral students independently associate a topic to each of the items. For example, for the item – The doctor diagnosed/diagnoses my medical problem satisfactorily (construct - Empathy), we find the students associated it with the topic theme – Responsiveness / Helpfulness of physicians. The inter-rater reliability is found to be 0.644. Relevant topics are found for 15 of the 25 items. The 10 items for which no match has been found by the doctoral students are marked as N in Column (3), Table 6.

The next step is to label the items as available (Y), not-available (N) and partially-available (P) under unstructured information availability. Manual matching is carried out for the 15 items for which matches have been found, to identify more detailed linkages between keywords in the topics with the items. The items of the survey are checked against the top 100 keywords of each of the topics analysed by the LDA (Table 5). Studies employing LDA have generally considered the top 20–30 keywords [47]. We use the term score [39] for each of the 50,589 keywords (obtained from the reviews) provided by the LDA algorithm, to decide on the number of keywords to be analysed. We found that the term score of the 100th keyword (arranged in descending order of term scores) is <5% of the term score of the top keyword for all the 8 topics.

Two Teaching Associates (TAs) with post graduate degrees at a reputed business school in India are then asked to independently label the items as Y or N in terms of information availability. The TAs are provided the items (from the questionnaire) and are asked to refer to the 100 keywords of the topics which the items had previously been matched with. Term scores are not provided to them. We particularly instructed the TAs to consider synonyms or words with similar semantic concepts. The TAs then look at the keywords and mark the item as Y if the set of keywords have an overall association with the item, or N otherwise. The inter-rater reliability is found to be 0.627. For each of the items, the match between the findings of the two TAs is considered. If both TAs agree on the match (Y or N), it is retained accordingly. If there is a disagreement between the TAs' findings, we mark the items as partially-available (P). This results in 6 items marked as Y, 4 as N and 5 as P in Column (3), Table 6 out of the 15 items. Column (4), Table 6 lists the information availability from structured and unstructured data. There is a total of 7 items which are not present in either structured or unstructured format and labelled as N. There are thus 12 out of 25 items which are found to be partially available or unavailable on PRWs. This answers the second research question, as stated in the introduction.

Information on the physician's availability in an emergency is not available as either structured or unstructured data. Patients often consult physicians for serious health issues, which contributes to the importance of this item. Information on the hospital's overall cleanliness is not available either as structured or unstructured information. One possible reason for the presence of this item can be due to the high sanitation expectations brought about by the impact of Covid-19. As the pandemic continues to bring about significant behaviour shifts worldwide, the overall cleanliness of a facility may become an important issue in physician selection. PRWs can add this as a star rating (out of 5) for each hospital. Waiting time is also worth mentioning, as it appears as a topic in the reviews (Topic 6) and as an item in the survey instrument (Table 6). This is in line with the open-ended questionnaire filled by physicians as part of the empirical study, wherein the majority list waiting time as an important factor in booking an appointment online, as opposed to offline bookings. The cost of medications is also not available in either form, which is important since the patients come from different economic backgrounds and may be unable to afford expensive medication. Most items related to Empathy and Efficiency are only partially available or unavailable on the PRWs.

5. Experimental study

The previous two phases of the study have established the information gap on the PRWs and identified the items that constitute the information gap. Previous studies have indicated that while making purchase decisions, consumers may not be able to process the information provided, even though they may have identified the information as desirable [48]. Therefore, the efficacy of providing additional information to address the gap is verified by conducting a controlled experiment. Controlled experimental studies are quite common in the IS literature and have been carried out across problem domains like e-commerce usage [49] and online healthcare [29].

For the controlled experiment, two of the items from Table 6 are taken – "The doctor gave/gives adequate time to me" and "Overall cleanliness of hospital/clinic". This selection is based on qualitative interviews used to gain insight into the relative importance of the constructs used in the study. Each construct is ranked from 1 to 5 by ten respondents. The relative importance is based on the ranking, with 1 being the most important. Respondents who had previously filled out the survey in study 1 are chosen to ensure that there is no confusion among respondents in the interview. Feedback is obtained regarding the same in the context of physician selection. It is found that the physician's Empathy is considered important by most respondents, citing patient-physician interaction, which is in line with the results in Table 4. Tangibles is also considered important by prospective patients, possibly due to a heightened appreciation for hygiene requirements during the prevalent pandemic. The factor loadings for each item in these two constructs is then considered from Table A.1, as they determine an item's absolute contribution to its assigned construct [35]. For Empathy, many items are subjective in nature and may not be easily implemented by PRWs.

Thus, the first item is specifically selected based on two things: the relative ease with which PRWs can provide this information as structured information, despite being ranked 5th in terms of loadings (0.761), and the fact that implementing this item may also partially explain the impact of the items "The doctor spent/spends adequate time to understand my problem" (0.825) and "The doctor was understanding of my financial or time constraints/limitations" (0.581). Extant healthcare literature recognises that the time that a physician spends with their patient is an important factor in patient satisfaction [9]. Information on the cleanliness of the hospital/clinic is not present on the PRWs either as structured or unstructured information. Further, the factor loading of this item is ranked the highest in Tangibles (0.814). It is easy to present as structured information and is highly relevant given the ongoing pandemic. Physician profile pages are created to test the effect of providing information on the two factors.

Two levels are considered for each factor - *Available* and *Not Available*. Information on each of the two items is either added to a physician profile or listed as "N/A". The study can be characterised as a 2 (The doctor gives an adequate amount of time to me: Available versus Not Available) x 2 (Cleanliness of the hospital/clinic: Available vs Not Available) full factorial design. The subject of interest here is physician booking on an online platform. We thus create a simulated scenario where participants are exposed to four physicians' profiles, each of which provides one of the four conditions of the 2 × 2 experiment. The experiment design is shown in Table 7. The information on the first item is presented as the average time spent with the patient and hygiene star ratings for the clinic/hospital.

The experiment is conducted such that each participant goes through all four treatment conditions. This is done via a Latin Square design, which has the advantage of eliminating the demand effect. The sequence of the treatments as given in Table 7 is randomised among participants. The Latin square design and randomisation of sequences are given in Table 8. Each treatment group is mapped to the profiles of a set of four physicians. A set of 2 male and 2 female physicians are selected to eliminate gender bias.

5.1. Pre-tests and pilot study

A pre-test is conducted to understand how to add the information to the physicians' profiles. Two sets of profiles are created based on the Group 1 sequence in Table 8 – one where all physicians are rated together after the profiles have been shown, and one where physicians are sequentially rated after each profile has been shown. This pre-test is conducted using 20 participants, 10 of whom are shown each set of profiles. The results indicated that participants must be shown the profiles sequentially to avoid confusion. The values required to operationalise the levels of both the factors are finalised based on the feedback from the participants (Average time spent with the patient – 15 min and Hygiene rating – 4 stars). Participant responses for average time spent indicate a length of 10–15 min as ideal. The mean, median and mode of the ideal consultation time reported by the respondents are

Table 7
Experimental Design.

		The doctor gave adequate time to me	
		Not Available	Available
Cleanliness of the hospital/clinic	Not Available	Condition 4 (C4)	Condition 2 (C2)
	Available	Condition 3 (C3)	Condition 1 (C1)

Table 8
Latin Square Design.

	Group 1	Group 2	Group 3	Group 4
Physician 1	C1	C2	C3	C4
Physician 2	C2	C1	C4	C3
Physician 3	C3	C4	C1	C2
Physician 4	C4	C3	C2	C1

17.17, 15 and 15, respectively. A consultation time of 15 min is selected based on the above results. The position of the added information is also adjusted based on feedback. The items are renamed to "Average consultation time" and "Hygiene rating" for ease of understanding of respondents. A pilot study is conducted using 37 participants to understand whether the Latin Square design is correctly interpreted with the modified profiles. We find that the improved positioning of additional information shows a positive impact on the pilot study results.

5.2. Main experiment

The main experiment is conducted with post-graduate students at a reputed business school in India. Most of the participants are regular healthcare consumers and conversant with PRWs. We collected data from 153 participants, with approximately 51% and 42% participants in the range of 18–24 and 25–34, respectively. 68% of the participants have booked a physician's appointment in the last six months, and around 55% have previously booked a physician's appointment online. The experiment is conducted on a virtual video conferencing platform where all participants log in and are briefed on the study. The participants are informed that the purpose of the experiment is to understand the factors that contribute to the selection of a physician online. After the briefing, a link is sent to the participants to fill out a scenario-based survey. The respondents are randomly assigned to one of the four sequences mentioned in Table 8. The context of the experiment is set as follows – "*Assume you are suffering from a fever for the past 3 days. You now want a medical diagnosis for your condition. To this end, you wish to visit a General Physician for treatment and medication. You navigate to an online doctor's appointment portal, where you are presented with the profiles of four doctors. Please go through the profiles of each of the four doctors carefully. You will have to answer some questions based on the information present in the profiles in the later section of the survey.*"

The four physician profiles are simulated by changing the data on actual physicians' webpages observed on the PRWs. One of the sample profiles used in the experiment is presented in Fig. A.2. Participants are asked to rate the likelihood of booking an appointment with the physician on a scale of 1 ("not at all likely") to 7 ("very likely"). Participants are also asked questions that capture their familiarity with the process – whether they have previously booked a doctor's appointment online (Yes or No). Demographic variables like age and gender are captured at the end of the survey.

5.3. Analysis and results

5.3.1. Manipulation checks

The respondents answer questions on a separate page as part of the manipulation check once all four profiles are rated. The questions are 1) whether they can recall the average consultation time of the physician and 2) whether they can recall the hygiene rating of the physician. 92.8% of respondents remember both correctly, indicating that the manipulations work as intended. Questions regarding the choice of clinic and physicians' experience are added before and after the manipulation checks to prevent the participants from accurately guessing the study's hypothesis. The responses to these questions are not considered in our analysis. An open-ended question is asked regarding how they rated the best physician as a post-study questionnaire. Only 6.54% of respondents could guess the actual objective of the study. The study can thus be considered viable.

5.3.2. Experimental results

The data is grouped by ratings of physicians according to each condition (as given in Tables 6 and 7). The data is tested for normality, where the Shapiro-Wilk test is significant, which indicates that the data is non-normal. The non-normality is verified by analysing the residuals of a preliminary ANOVA test. Levene's test is also significant, indicating heteroscedasticity in the data. We follow the procedure outlined by [50], which shows a two-step transformation for achieving normality and homoscedasticity in such cases. After carrying out the transformations, the tests are re-executed, showing that heteroscedasticity has been addressed, but the data is still non-normal. As a result, non-parametric testing is carried out using the Independent-Samples Kruskal-Wallis test. The ratings are converted to ranks and grouped by condition. The test is found to be significant ($H = 73.045$, $df = 3$, ASymp. Sig. $= 0.000$). This indicates that the distributions of the different physicians' ratings (ranks) are not the same across the categories of conditions. Further pairwise comparisons of conditions are shown using the Dunn-Bonferroni post-hoc test. The mean ranks of groups and pairwise comparison results are shown in Tables 9 and 10, respectively.

Providing information on the average time spent with the patient and the hygiene rating positively impacts physician selection. No significant positive impact was found for providing information only on average time spent. In addition, responses suggest that hygiene rating has more impact than average time spent. The higher impact of hygiene can be attributed to the Covid-19 pandemic, which has made consumers more aware of good hygiene practices [51].

The relatively lower importance given to average time spent may be attributed to the age group of the respondents. Most of the respondents in the study are young students. Extant healthcare literature indicates that time spent with the doctor tends to be less important for young patients than for elderly patients [52]. The results from the experiment demonstrate that the experience of selecting physicians online can be enriched if PRWs provide information on the items identified as not available, thus answering the third research question.

6. Discussion and conclusion

In this paper, the information gap that exists on PRWs is identified. The analysis of empirical data reveals the information considered important by patients. Furthermore, analysis of the unstructured information (patient reviews) using a topic modelling technique reveals the topics considered important by patients post-consultation, where we find eight distinct topics. Twelve out of the twenty-five items are partially present or not present on the PRWs in either structured or unstructured form. This finding demonstrates that there is indeed a gap between the information that patients seek from PRWs and the information that is available. The controlled experiment conducted with two items that constitute part of the information gap demonstrates that providing relevant information to patients significantly improves the physician selection process.

Contributions to Academia: This paper makes several significant contributions to the extant literature. First, our study adds to the online healthcare communities (OHC) literature by identifying the information gap that exists on the PRWs. Although the extant IS healthcare literature

Table 9
Mean ranks obtained per condition via Kruskal-Wallis test.

Conditions	N	Mean Rank
C1 (both available)	153	384.57
C2 (only consultation time available)	153	269.67
C3 (only hygiene rating available)	153	338.19
C4 (neither available)	153	233.58
Total	612	

Table 10
Pairwise comparisons of conditions using Post-hoc Tests.

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig. (Bonferroni)
C4-C2	36.0915	19.42	1.86	0.063	0.379
C4-C3	104.6176	19.42	5.39	0.000	0.000
C4-C1	150.9902	19.42	7.78	0.000	0.000
C2-C3	−68.5261	19.42	−3.53	0.000	0.003
C2-C1	114.8987	19.42	5.92	0.000	0.000
C3-C1	46.37255	19.42	2.39	0.017	0.102

has studied patient reviews and their usage, credibility, and impact on the selection of physicians, to the best of our knowledge, this work is the first one to study the adequacy of information on PRWs. We believe that studying information quality on PRWs has great potential for future research. While our study focuses on physicians in general, future research can contribute to the OHC literature by studying the influence of information quality on the selection of specialist physicians. Second, our work extends the literature on IS success and usage in the domain of OHCs. The study demonstrates that providing relevant information benefits the patients and establishes the link between informational quality and patient experience on PRW. Third, our study extends the HEALTHQUAL framework [16] to the domain of online physician selection. Fourth, this study adds to the extant literature by categorising the broad strands of information available as unstructured information (patient reviews) on PRWs. This approach can be applied to other OHCs to better understand the discussions that take place both among patients and between patients and physicians.

Contributions to Practice: Our work provides several action points for the stakeholders of PRW businesses. First, business managers can take the necessary steps to address the gaps in information as identified by our study. For example, sales and marketing managers responsible for inbound web traffic can initiate modifications to their websites to address the lack of relevant information for patients. Some of these modifications can be incorporated as structured information. Second, the experimental study has shown the positive impact of adding structured information related to the average time spent by a physician and the cleanliness of the clinic/hospital. For some information points, such as waiting time, our research suggests that although it is present as unstructured information (patients tend to write about negative experiences regarding waiting time), PRW managers may consider presenting the average waiting time as structured information. As the number of bookings for a physician for a specific time period increases, the PRW can determine the expected wait time and display the same. Since patients could then compare waiting times across periods, they could decide on a convenient slot. Third, other information gaps, such as information on whether the physician prescribes only necessary diagnostic tests/medication, may be incorporated only as unstructured information. In this regard, these websites will have to devise means of nudging the patients to write about these aspects of their experiences with the physicians. Fourth, the effect of the Covid-19 pandemic is quite evident as patients react positively to the addition of hygiene factors to existing PRW profiles. Business development managers may take note of this and add such relevant information to the sites, which might encourage more patients to seek services from PRWs. Fifth, physicians and hospital administrators can use our research findings to provide information on their profiles to address patients' information needs. For example, our research demonstrates that a physician's expertise is an important point for patients and needs special attention. Physicians may take extra care in writing about their expertise regarding qualifications and specialisations on their profiles.

This work examines PRWs as an information system that caters to the information needs of patients looking to consult a physician online. As the Covid-19 pandemic continues to impact healthcare services worldwide, traffic on these websites has increased manifold, and patients recognise these websites to be good sources of information on physicians. Teleconsultation appointments facilitated through these websites have also recorded remarkable growth rates in the last two years. In light of current developments, it is important to check if the PRWs can satisfy the information needs of the patients. To the best of our knowledge, this work is the first to combine text reviews of patients with an empirical study of patients' opinions, as well as an experimental verification. This interdisciplinary approach allows us to focus on the informational aspect of PRWs and identify an important lacuna, namely, the lack of adequate information.

6.1. Limitations and future research

The research has a few limitations, which will hopefully provide directions for future research. First, patients may gather information from multiple sources before booking an appointment. For example, in the hospitality industry, it is a common practice to explore multiple websites during the information search phase to find a hotel [53]. Cross-referencing physicians across multiple websites is a big challenge considering the scale of these datasets and the lack of uniformity in data fields.

Second, the LDA model used for text analysis may not be able to uncover meaningful, well-interpretable topics from the reviews beyond ten topics. While our findings are borne out by the topics uncovered, this is a limitation of the model used in our study. Topic modelling in healthcare is still in its nascent stage compared to other research domains. Researchers may investigate techniques that are better suited for healthcare data in the future.

Third, two of the items that constitute the information gap are selected and presented as structured data to conduct the experimental study. An elaborate experiment must be carried out to gauge the effect of addressing the information gap in its entirety. This entails designing an experiment where respondents are exposed to both structured and unstructured data. It would be an interesting research challenge to identify the information items that should be presented as structured information and the ones that are more suitable for unstructured information.

Fourth, due to the limitation of the review data, we have not been able to consider patient heterogeneity while extracting the topics. The information on the severity of the disease, chronic disease, etc., is not present in the review data, and we have not been able to investigate the information needs of patients based on those attributes. Similarly, it is not possible to ascertain from the information available on the PRW whether patients have consulted a certain physician for ailments other than the specialty provided on the profile page of the physician. Therefore, information needs for different physician specialisations have not been studied in this work. It would be interesting to explore these facets of information needs in future studies.

Author contribution

All authors contributed equally towards this paper.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendix A. Tables and figures

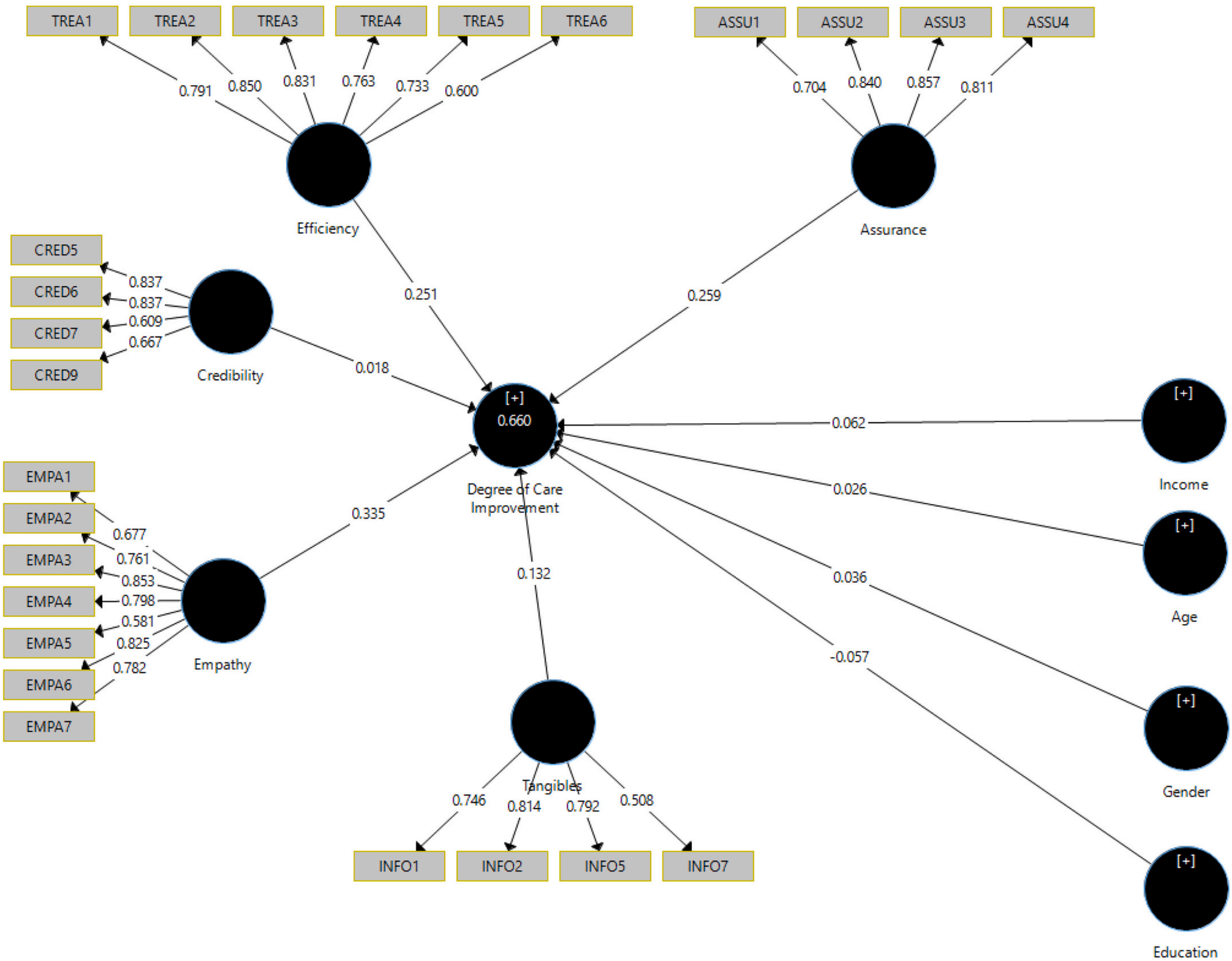

Fig. A.1. SEM Model.

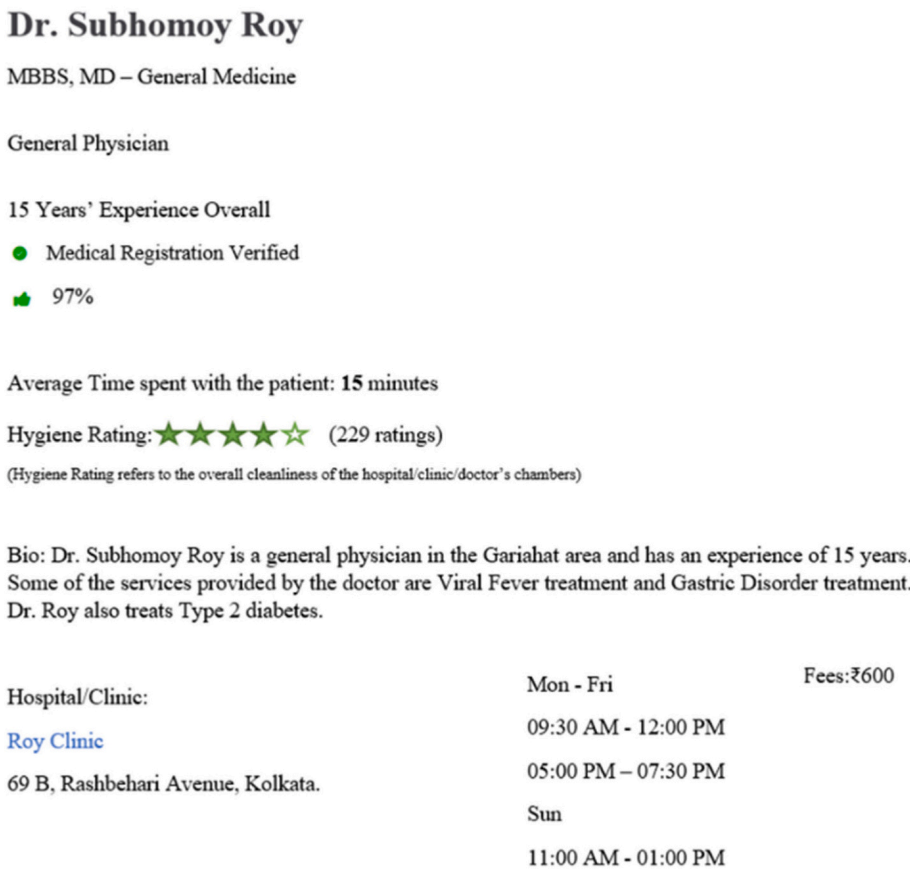

Fig. A.2. Sample Physician Profile.

Table A.1
Model Statistics, CFA Results and Ranking of Important Factors.

Constructs[a]	Lambda	Mean	SD	Rank
F1. Credibility of the doctor	*Ch. Alpha: 0.721*			
The doctor's ratings from the website(s)	0.837	4.85	1.69	22
The doctor's reviews from the website(s)	0.837	4.95	1.68	11
Years of experience of doctor	0.609	5.70	1.38	16
The doctor's achievements: – quality certification, accreditations, awards received, membership of medical societies	0.667	4.81	1.59	25
*Doctor's age	0.328	4.23	1.90	–
*Gender of the doctor	0.119	2.39	1.85	–
*Language of communication (with the patient)	0.384	3.98	1.89	–
F2. Doctor's interaction with the patient – Empathy	*Ch. Alpha: 0.874*			
The doctor behaved/behaves politely with me	0.677	6.21	1.17	24
The doctor gave/gives an adequate amount of time to me	0.761	6.32	1.03	5
The doctor listened/listens to my symptoms	0.853	6.58	0.79	18
The doctor explained/explains my medical problem to me	0.798	6.59	0.83	8
The doctor was understanding of my financial or time constraints/limitations	0.581	5.79	1.25	1
The doctor spent/spends adequate time to understand my problem	0.825	6.41	0.94	12
The doctor diagnosed/diagnoses my medical problem satisfactorily	0.782	6.59	0.80	9
F3. Information about the doctor and hospital – Tangibles	*Ch. Alpha: 0.694*			
The expertise of the doctor	0.746	6.47	0.85	3
Overall cleanliness of the hospital/clinic	0.814	6.45	0.84	13
Neatness of the doctor	0.508	5.27	1.38	21
Waiting time at the doctor's office	0.746	6.47	0.85	14
*Reputation of doctor's medical college	0.497	4.94	1.67	–
*Availability of sufficient choices of booking slots	0.668	5.32	1.40	–
*Doctor's office is nearby	0.492	5.06	1.44	–
*Availability of preferred time slots	0.644	5.39	1.35	–
F4. Assurance regarding the medical staff – Assurance	*Ch. Alpha: 0.816*			
Medical care was/is performed carefully	0.704	6.42	0.84	6
The hospital staff was/is always honest and genuine	0.840	6.29	0.96	17
The hospital staff respected/respects my feelings	0.857	6.08	1.13	2
The hospital staff was/is friendly and cheerful	0.811	5.86	1.24	10
F5. Treatment of the doctor – Efficiency	*Ch. Alpha: 0.856*			

(*continued on next page*)

Table A.1 (*continued*)

Constructs[a]	Lambda	Mean	SD	Rank
The doctor provided/provides necessary medications to me	0.791	6.29	1.02	23
The doctor provided/provides necessary diagnostic tests to me	0.850	6.21	1.12	4
The doctor provided/provides necessary procedures to me	0.831	6.27	0.99	20
The cost of medication was/is reasonable	0.763	6.04	1.17	15
The doctor's fee was/is reasonable	0.733	5.95	1.17	19
The doctor was/is available for consultation in an emergency	0.600	6.27	1.13	7
F6. The outcome of seeing the doctor - Degree of Care Improvement	*Ch. Alpha: 0.845*			
The medical staff treated/treats me to the best of their ability	0.738	6.08	1.10	
My medical recovery was/is fast	0.716	6.14	1.01	
I was (have been) informed of related diseases and their dangers	0.721	6.27	1.05	
My medical condition improved/improves after treatment	0.802	6.53	0.80	N/A[b]
Medical staff were/are helpful to me	0.820	6.19	0.96	
My treatment was/is value-for-money	0.703	6.19	1.16	

[a] Items in italics have been added to the constructs, and items marked with a star (*) have been removed.
[b] Dependent variable for regression.

References

[1] S. Li, R.J. Lee-Won, J. McKnight, Effects of online physician reviews and physician gender on perceptions of physician skills and primary care physician (PCP) selection, Health Commun. 34 (2019) 1250–1258, https://doi.org/10.1080/10410236.2018.1475192.

[2] K. Okike, N.R. Uhr, S.Y.M. Shin, K.C. Xie, C.Y. Kim, T.T. Funahashi, M.H. Kanter, A comparison of online physician ratings and internal patient-submitted ratings from a large healthcare system, J. Gen. Intern. Med. 34 (2019) 2575–2579, https://doi.org/10.1007/s11606-019-05265-3.

[3] F. Rothenfluh, E. Germeni, P.J. Schulz, Consumer decision-making based on review websites: are there differences between choosing a hotel and choosing a physician? J. Med. Internet Res. 18 (2016) https://doi.org/10.2196/jmir.5580.

[4] S. McLennan, D. Strech, A. Meyer, H. Kahrass, Public awareness and use of German physician ratings websites: cross-sectional survey of four north German cities, J. Med. Internet Res. 19 (2017), https://doi.org/10.2196/jmir.7581.

[5] X. Han, B. Li, T. Zhang, J. Qu, Factors associated with the actual behavior and intention of rating physicians on physician rating websites: cross-sectional study, J. Med. Internet Res. 22 (2020), e14417, https://doi.org/10.2196/14417.

[6] S. Tripathi, 500% Spike in Online Medical Consultation under Lockdown: Practo Report | Business Standard News. https://www.business-standard.com/article/current-affairs/500-spike-in-online-medical-consultation-under-lockdown-practo-report-120062201323_1.html, 2020 (accessed December 15, 2021).

[7] A. Rodriguez, The Pandemic Ushered in "a New Era of Medicine": These Telehealth Trends Are Likely here to Stay, USA TODAY. https://www.usatoday.com/story/news/health/2021/03/21/telehealth-trends-covid-19-mental-health-primary-care-coronavirus/4699979001/, 2021.

[8] D. Arnott, S. Gao, Behavioral economics for decision support systems researchers, Decis. Support. Syst. 122 (2019), 113063, https://doi.org/10.1016/j.dss.2019.05.003.

[9] M. Etco, S. Sénécal, P.M. Léger, M. Fredette, The influence of online search behavior on consumers' decision-making heuristics, J. Comput. Inf. Syst. 57 (2017) 344–352, https://doi.org/10.1080/08874417.2016.1232987.

[10] N. Amblee, R. Ullah, W. Kim, Do product reviews really reduce search costs? J. Organ. Comput. Electron. Commer. 27 (2017) 199–217, https://doi.org/10.1080/10919392.2017.1332142.

[11] M. Lee, S. Youn, Electronic word of mouth (eWOM): how eWOM platforms influence consumer product judgement, Int. J. Advert. 28 (2009) 473–499, https://doi.org/10.2501/S0265048709200709.

[12] K.Z.K. Zhang, S.J. Zhao, C.M.K. Cheung, M.K.O. Lee, Examining the influence of online reviews on consumers' decision-making: a heuristic-systematic model, Decis. Support. Syst. 67 (2014) 78–89, https://doi.org/10.1016/j.dss.2014.08.005.

[13] S. Grabner-Kräuter, M.K.J. Waiguny, Insights into the impact of online physician reviews on patients' decision making: randomized experiment, J. Med. Internet Res. 17 (2015), e93, https://doi.org/10.2196/jmir.3991.

[14] V. Venkatesh, S. Brown, Y. Sullivan, Guidelines for conducting mixed-methods research: an extension and illustration, J. Assoc. Inf. Syst. 17 (2016) 435–494, https://doi.org/10.17705/1jais.00433.

[15] G. Bentum-Micah, Z. Ma, W. Wang, S.A. Atuahene, V. Bondzie-Micah, Perceived service quality, a key to improved patient satisfaction and loyalty in healthcare delivery: the Servqual dimension approach, J. Health Med. Sci. 3 (2020), https://doi.org/10.31014/aior.1994.03.02.114.

[16] D.H. Lee, HEALTHQUAL: a multi-item scale for assessing healthcare service quality, Serv. Bus. 11 (2017) 491–516, https://doi.org/10.1007/s11628-016-0317-2.

[17] F. AlOmari, Measuring gaps in healthcare quality using SERVQUAL model: challenges and opportunities in developing countries, Meas. Bus. Excell. (2020), https://doi.org/10.1108/MBE-11-2019-0104.

[18] B. Leach, M. Gradison, P. Morgan, C. Everett, M.J. Dill, J.S. de Oliveira, Patient preference in primary care provider type, Healthcare. 6 (2018) 13–16, https://doi.org/10.1016/j.hjdsi.2017.01.001.

[19] A. Parasuraman, V.A. Zeithaml, L. Berry, SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality 1988 (64) (1988) 12–40.

[20] R. Vandamme, J. Leunis, Development of a multiple - item scale for measuring hospital service quality, Int. J. Serv. Ind. Manag. 4 (1993) 30–49, https://doi.org/10.1108/09564239310041661.

[21] F.-A. Luca, C.A. Ioan, C. Sasu, The importance of the professional personal brand. The doctors' personal brand, Proc. Econ. Finance. 20 (2015) 350–357, https://doi.org/10.1016/s2212-5671(15)00083-0.

[22] Y. Xu, M. Armony, A. Ghose, The interplay between online reviews and physician demand: an empirical investigation, Manag. Sci. (2021), https://doi.org/10.1287/mnsc.2020.3879.

[23] G.G. Gao, B.N. Greenwood, R. Agarwal, J.S. McCullough, VOCAL minority and silent majority: how do online ratings reflect population perceptions of quality, MIS Q. Manag. Inf. Syst. 39 (2015) 565–589, https://doi.org/10.25300/MISQ/2015/39.3.03.

[24] S.F. Lu, H. Rui, Can we trust online physician ratings? Evidence from cardiac surgeons in Florida, Manag. Sci. 64 (2018) 2557–2573, https://doi.org/10.1287/mnsc.2017.2741.

[25] D.H. Saifee, Z. Zheng, I.R. Bardhan, A. Lahiri, Are Online Reviews of Physicians Reliable Indicators of Clinical Outcomes?, Workshop Inf. Syst. Econ. Seoul South Korea, 2017.

[26] S. Lantzy, D. Anderson, Can consumers use online reviews to avoid unsuitable doctors? Evidence from RateMDs.com and the federation of state medical boards, Decis. Sci. 51 (2020) 962–984, https://doi.org/10.1111/deci.12398.

[27] Y. Chen, S. Lee, User-generated physician ratings and their effects on patients' physician choices: evidence from yelp, SSRN Electron. J. (2021), https://doi.org/10.2139/ssrn.3796740.

[28] C.P. Furner, R.A. Zinko, The influence of information overload on the development of trust and purchase intention based on online product reviews in a mobile vs. web environment: an empirical investigation, Electron. Mark. 27 (2017) 211–224, https://doi.org/10.1007/s12525-016-0233-2.

[29] K. Tates, M.L. Antheunis, S. Kanters, T.E. Nieboer, M.B. Gerritse, The effect of screen-to-screen versus face-to-face consultation on doctor-patient communication: an experimental study with simulated patients, J. Med. Internet Res. 19 (2017), e421, https://doi.org/10.2196/jmir.8033.

[30] S.C. Martino, R. Grob, S. Davis, A.M. Parker, M.L. Finucane, J.L. Cerully, L. Rybowski, D. Shaller, M. Schlesinger, Choosing doctors wisely: can assisted choice enhance patients' selection of clinicians? Med. Care Res. Rev. 76 (2019) 572–596, https://doi.org/10.1177/1077558717743822.

[31] A.D. Shukla, G. Gao, R. Agarwal, How digital word-of-mouth affects consumer decision making: evidence from doctor appointment booking, Manag. Sci. 67 (2021) 1546–1568, https://doi.org/10.1287/mnsc.2020.3604.

[32] L. Bouayad, M.C. Tremblay, H. Jain, C. Gaudioso, Intrinsic motivation to share health information: Design guidelines and features for patient sharing platforms, in: S. Hofmann, O. Müller, M. Rossi (Eds.), Des. Digit. Transform. Co-Creat. Serv. Citiz. Ind, Springer International Publishing, Cham, 2020, pp. 68–74.

[33] B.H. Bornstein, D. Marcus, W. Cassidy, Choosing a doctor: an exploratory study of factors influencing patients' choice of a primary care doctor, J. Eval. Clin. Pract. 6 (2000) 255–262, https://doi.org/10.1046/j.1365-2753.2000.00256.x.

[34] C.G. Lis, M. Rodeghier, D. Gupta, The relationship between perceived service quality and patient willingness to recommend at a national oncology hospital network, BMC Health Serv. Res. 11 (2011), https://doi.org/10.1186/1472-6963-11-46.

[35] J. Hair, R. Anderson, B. Babin, W. Black, Multivariate Data Analysis: A Global Perspective: Pearson Upper Saddle River, 2010.

[36] M. Riani, A.C. Atkinson, A. Cerioli, Finding an unknown number of multivariate outliers, J. R. Stat. Soc. Ser. B Stat Methodol. 71 (2009) 447–466, https://doi.org/10.1111/j.1467-9868.2008.00692.x.

[37] A. Wieland, C.F. Durach, J. Kembro, H. Treiblmaier, Statistical and judgmental criteria for scale purification, Supply Chain Manag. 22 (2017) 321–328, https://doi.org/10.1108/SCM-07-2016-0230.

[38] J. Henseler, M. Sarstedt, Goodness-of-fit indices for partial least squares path modeling, Comput. Stat. 28 (2013) 565–580, https://doi.org/10.1007/s00180-012-0317-1.

[39] D.M. Blei, A.Y. Ng, M.I. Jordan, Latent dirichlet allocation, J. Mach. Learn. Res. 3 (2003) 993–1022.
[40] H. Geva, G. Oestreicher-Singer, M. Saar-Tsechansky, Using retweets when shaping our online persona: topic modeling approach, MIS Q. 43 (2019) 501–524.
[41] D.G. Ko, F. Mai, Z. Shan, D. Zhang, Operational efficiency and patient-centered health care: a view from online physician reviews, J. Oper. Manag. 65 (2019) 353–379, https://doi.org/10.1002/joom.1028.
[42] A.M. Shah, X. Yan, S. Tariq, M. Ali, What patients like or dislike in physicians: analyzing drivers of patient satisfaction and dissatisfaction using a digital topic modeling approach, Inf. Process. Manag. 58 (2021), 102516, https://doi.org/10.1016/j.ipm.2021.102516.
[43] D. Mimno, H.M. Wallach, E. Talley, M. Leenders, A. McCallum, Optimizing semantic coherence in topic models, in: Proc. Conf. Empir. Methods Nat. Lang. Process., Association for Computational Linguistics, USA, 2011, pp. 262–272.
[44] H.M. Wallach, I. Murray, R. Salakhutdinov, D. Mimno, Evaluation methods for topic models, in: Proc. 26th Annu. Int. Conf. Mach. Learn. - ICML 09, ACM Press, Montreal, Quebec, Canada, 2009, pp. 1–8, https://doi.org/10.1145/1553374.1553515.
[45] J. Bischof, E.M. Airoldi, Summarizing topical content with word frequency and exclusivity, in: Proc. 29th Int. Conf. Mach. Learn. ICML-12, 2012, pp. 201–208.
[46] D.M. Blei, J.D. Lafferty, Topic models, in: Text Min, Chapman and Hall/CRC, 2009, pp. 101–124.
[47] I. Sutherland, K. Kiatkawsin, Determinants of guest experience in Airbnb: a topic modeling approach using LDA, Sustainability. 12 (2020) 3402, https://doi.org/10.3390/su12083402.
[48] H. Fen Hu, A.S. Krishen, When is enough, enough? Investigating product reviews and information overload from a consumer empowerment perspective, J. Bus. Res. 100 (2019) 27–37, https://doi.org/10.1016/j.jbusres.2019.03.011.
[49] S. Bhattacharyya, I. Bose, S-commerce: influence of Facebook likes on purchases and recommendations on a linked e-commerce site, Decis. Support. Syst. 138 (2020), 113383, https://doi.org/10.1016/j.dss.2020.113383.
[50] G.F. Templeton, A two-step approach for transforming continuous variables to normal: implications and recommendations for IS research, Commun. Assoc. Inf. Syst. 28 (2011) 41–58, https://doi.org/10.17705/1cais.02804.
[51] Y.H. Lin, C.H. Liu, Y.C. Chiu, Google searches for the keywords of "wash hands" predict the speed of national spread of COVID-19 outbreak among 21 countries, Brain Behav. Immun. 87 (2020) 30–32, https://doi.org/10.1016/j.bbi.2020.04.020.
[52] T.I. Lemon, R.H. Smith, Consultation content not consultation length improves patient satisfaction, J. Fam. Med. Prim. Care. 3 (2014) 333.
[53] L. McCarthy, D. Stock, R. Verma, How Travelers Use Online and Social Media Channels to Make Hotel-choice Decisions 10, 2010. https://ecommons.cornell.edu/handle/1813/71099.

Soham Ghosh is a doctoral candidate of Management Information Systems at the Indian Institute of Management Calcutta. He has done his Bachelor of Engineering in Instrumentation and Electronics from Jadavpur University, Kolkata. His research interests are in business analytics, health IT, and online healthcare communities (OHCs). His work has previously appeared in the proceedings of conferences like ICIS 2021.

Soumyakanti Chakraborty is an Associate Professor in the area of Management Information Systems at the Indian Institute of Management Calcutta. Dr. Soumyakanti Chakraborty received his doctorate from Indian Institute of Management Calcutta. Dr. Chakraborty's research interests are in the areas of platform economics, combinatorial auctions, and healthcare analytics. His research articles have appeared in journals like EJOR, Decision Support Systems, OMEGA, Information Systems Frontiers, etc.

Narain Gupta is professor at Management Development Institute (MDI), Gurgaon, Gurugram, India (Among top 10 ranking business school in India since a decade in national rankings). He has published in multiple A* and A category journal such as DSS, JKM, JCP, C&IE, IJPR, IJPE, IBR, IMR, MIP, EPW, JORS, and multiple business cases in Richard Ivey. He is a PhD from IIM Ahmedabad (EQUIS Accredited), Rank one business school in India. He has worked in A P Moller Maersk shipping and the consulting companies for five years. He has been a visiting fellow of Aarhus University, Denmark for his course on Business Analytics.

Sumanta Basu is an Associate Professor in the Operations Management Department at the Indian Institute of Management Calcutta. He received his doctorate in Production and Quantitative Methods from the Indian Institute of Management Ahmedabad. Sumanta has six years of work experience in IT consulting and in supply chain. He has three US patents on process optimization during his stint in the industry. He works in the area of revenue management, pricing of information goods, economic analysis of interactions in supply chain, optimization etc. His research articles have appeared in EJOR, OMEGA, IJPR, JRCS, INFOR and a few other scholarly journals.