WHAT MAKES A DRUG REVIEW HELPFUL? THE ROLE OF PATIENTS' HEALTH CONDITION AND MEDICAL EXPERIENCE SIGNALS

Long Xia*
Department of Management and Entrepreneurship
Love School of Business, Elon University
Koury Business Center
Elon, North Carolina 27244, USA
lxia@elon.edu

ABSTRACT

Identifying helpful drug reviews could significantly assist patients in their medication decision-making. Despite that review helpfulness has been extensively explored in the prior literature, the findings might not be applicable to drug reviews due to the considerable medical-specific characteristics. In this study, we leveraged signaling theory and developed a theoretical framework to reveal how different information signals regarding patients' health conditions influence perceived review helpfulness. We harvested a large drug review dataset covering 9,908 drugs and corresponding 147,169 reviews from WebMD and adopted a deep learning technique to extract sentence-level medical signals directly from drug reviews with promising performance. Our empirical analysis suggests that information signals related to patients' health conditions and medical experiences have significant positive impacts on reviews' perceived helpfulness. In addition, their impacts largely depend on the drug type, review volume, overall rating, and how long the reviewer has taken the drug, revealing under which conditions the effects of patients' health signals on review helpfulness are likely to be weakened or strengthened. This study provides both theoretical and methodological contributions to the research on the helpfulness of reviews, especially in the online medical review context, and provides practical implications for various stakeholders regarding medication decision-making.

Keywords: Review helpfulness; Online drug reviews; Signaling theory; Deep learning; Text mining

1. Introduction

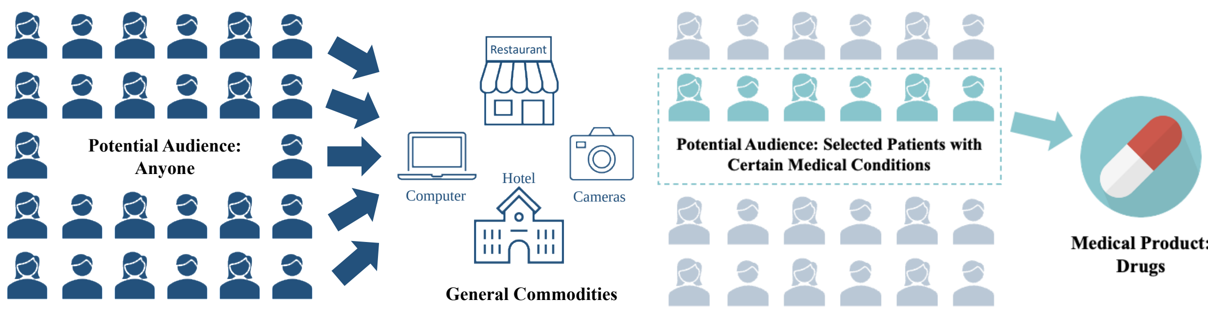
Online reviews play an increasingly important role in consumer decision-making, as they provide integral information to customers to reduce uncertainty related to products (Chevalier & Mayzlin, 2006). In the healthcare domain, we observe that various online platforms have become integral sources of valuable information for patients. Among these platforms are, for example, physician reviews (Hong et al., 2019; Wallace et al., 2014) and drug reviews (Cheerkoot-Jalim & Khedo, 2020; Liu & Chen, 2015), which assist patients in their medical decision-making. While online reviews have transformed patients from passive information receivers to active information seekers (Yan et al., 2015), the massive number of reviews has also created an information overload problem—the sheer volume of information has become a hindrance rather than a help, despite its potential benefits (Bawden & Robinson, 2009).

This information overload problem makes sifting through helpful and relevant reviews difficult, both managerially and practically (Mudambi & Schuff, 2010). A plethora of studies have explored the key determinants of review helpfulness, and a consistent finding is that both review characteristics (what the review is about) and reviewer characteristics (who contributed to the review) play an integral role in determining perceived review helpfulness (e.g., Li et al., 2013; Malik & Hussain, 2018; Siering et al., 2018). However, despite the fact that review helpfulness has been considered extensively in previous literature on various products, such as hotels (e.g., Mauro et al., 2021; Qazi et al., 2016), restaurants (e.g., Xia, 2023; Zhou & Guo, 2017), electronics (e.g., Ren & Hong, 2019; Zhang & Tran, 2010), movies (e.g., Baek et al., 2015), and books (e.g., Ren & Hong, 2019; Wang et al., 2020), those findings may not be applicable to medical products, such as pharmaceutical drugs, due to the content specificity.

First, while for general commodities (e.g., movies, books, and electronics), consumers are often the sole party responsible for making purchase choices, the medical domain is highly regulated, and a person's choice of drug is mainly determined by their healthcare providers, who possess professional expertise and knowledge (Yan & Tan, 2017). Conversely, patients often lack the specific medical expertise necessary to assess the quality and performance of drugs (Yan & Tan, 2017). This information asymmetry problem is thus more serious in the case of drugs than for

* Corresponding author

other general commodities and puts patients at a particular disadvantage (Yang et al., 2020). Second, unlike general commodities, whose consumers might be anyone (Figure 1a), consumers seeking medications belong to highly specific groups: those with certain diseases for which the drug is an indicated treatment (Figure 1b). The matter becomes even more complicated when we consider that even for patients with the same disease, drug reactions (e.g., side effects or efficacy) may vary dramatically for patients in different stages of the disease, with different medical conditions, or with different lifestyle and disease management behaviors (Chee et al., 2011; Divaris, 2017; Hsueh et al., 2016). Therefore, we propose that information regarding patients' health conditions and medical history plays an integral role in the perceptions of review helpfulness in the context of drugs. Finally, drugs differ from other products due to their associated health risks. Drug choice often determines a patient's quality of life, especially for patients with chronic diseases (Fox & Duggan, 2013), and it can even be a matter of life and death (Liu & Chen, 2015).



(a) (b)

Figure 1: Different Consumer Bases for General Commodities and Medical Products: (a) General commodities, anyone could be a potential consumer; (b) Drugs, selected patients with particular health conditions (e.g., diseases).

A direct consequence of most patients' lack of medical expertise and the devastating consequences of taking the wrong drug is that, besides the information received from medical providers, patients also actively search out additional information about drugs from their peers and other patients (Shah et al., 2019; Yan & Tan, 2017). In fact, prior studies have repeatedly shown that, compared to the information provided by doctors, who mainly receive information from drug labels or pharmaceutical companies, the information in drug reviews written by patients tends to focus on first-hand experiences and real-world performance, which might be more useful to potential readers (Wicks et al., 2010; Yan & Tan, 2017; Yan et al., 2019). For example, Yan & Tan (2017) revealed that patients actively use others' personal medical experiences to learn about drug quality, treatment experiences, and negative consequences (e.g., side effects); they often compare their treatment progress with that of other patients in similar circumstances. Given the greater information needs in the healthcare domain, identifying helpful reviews and presenting relevant information to patients is integral to supporting their medical decision-making and disease management. However, despite a plethora of studies that explore how patients interact with online healthcare communities and search for medical information (Yan & Tan, 2014, 2017; Yan et al., 2019), what makes a drug review helpful has received little research attention. To our knowledge, one of the few exceptions that has explored the determinants of drug review helpfulness was conducted by Zheng et al. (2021), who demonstrated that the medically related variables extracted from reviews, such as medical word sentiments and the ratio of medical words, influence helpfulness perception. However, they did not explore the in-depth textual information embedded in the review content, such as specific aspects of the drugs and the health conditions discussed. Given the various drug complications and treatment outcomes experienced by different patients, information about reviewers' health conditions and their medical experiences has the potential to help patients find peers with similar medical conditions, which will make it easier for them to find useful information (Yan et al., 2015; Yan & Tan, 2014).

We therefore suggest that although review helpfulness has been extensively explored in previous literature, given the medically specific characteristics of drugs, the findings of those studies might not be applicable to drug reviews (*Research Gap 1*). Moreover, given that drug performance is contingent on patients' health conditions, it is of both theoretical and practical significance to consider the impact of in-depth medical signals regarding patients' health conditions and medical experience (hereafter referred to collectively as medical signals) present in drug reviews, which have been overlooked by studies on helpfulness perception (*Research Gap 2*). With these two main research gaps identified, this study intends to answer three research questions. First, can we develop a text mining model to extract specific topics, particularly signals related to patients' health conditions and medical experience, that are often included in drug reviews (**RQ1**)? Second, what are the impacts of these medical signals on perceived review helpfulness (**RQ2**)? Third, how do their impacts vary with respect to different drug and reviewer characteristics (**RQ3**)?

To answer these questions, we leveraged signaling theory (Spence, 2002) to develop a theoretical framework to reveal how different signals related to patients' medical conditions influence the helpfulness perception of drug reviews. We harvested a large collection of drug reviews and adopted a deep learning–based natural language processing (NLP) technique to extract in-depth sentence-level medical signals with promising performance. Our empirical modeling suggests that health information and medical experience signals not only show strong predictive power for review helpfulness (direct effects) but are also moderated significantly by drug characteristics (drug type, volume, and valence) and reviewer characteristics (length of drug treatment) on review helpfulness perception (moderating effects). This study offers several contributions to the literature.

First, we contribute to the literature on online healthcare communities and review helpfulness by demonstrating that in-depth medical signals related to patients' health information and medical experiences largely impact helpfulness perception. We suggested that drugs differ from other general commodities due to their medically specific characteristics, which further enhance information asymmetry. Signaling theory, which explains how various signals might influence human behaviors in the presence of asymmetric information, is an excellent means of integrating in-depth medical signals into the design of a theoretical framework for studying review helpfulness. To the best of our knowledge, we are the first to holistically explore patients' information needs and provide quantitative evidence of the impacts of information about patients' health conditions and experiences on perceived drug review helpfulness. Despite signaling theory having been widely adopted in prior review helpfulness research, we believe it could play a more significant role in healthcare research, given the patient's lack of medical expertise and the associated elevated risks. Therefore, this study contributes to signaling theory's application in the medical domain, and we encourage future researchers to explore its utility in healthcare research to a greater extent.

Second, from a practical perspective, our findings stand to benefit the decision-making of multiple stakeholders. Given that online healthcare communities play an increasingly important role in patients' disease management (Yan et al., 2015; Yan & Tan, 2014), understanding what makes shared health information and experiences helpful is a key step toward helping reviewers and online platforms provide the information most relevant to meeting patients' information needs. For reviewers to write helpful reviews, we suggest that, in addition to various medical signals, they should also consider the signaling environment (e.g., drug type, review volume, and valence). For platforms to facilitate helpful information sharing among patients, we suggest understanding how drug reviews are perceived by patients to develop review guidelines for reviewers and enhance existing review-posting interface designs. Overall, identifying the features that constitute the most relevant and helpful drug reviews offers tremendous benefits to patients and may help them better engage in their health education and disease management.

Finally, from a methodological perspective, we developed an advanced deep learning model to extract in-depth medical signals from drug reviews. Previous studies have predominantly adopted aggregated measures (e.g., length and sentiment) and valence (e.g., ratings) to represent review content and different information aspects (Chevalier & Mayzlin, 2006; Duan et al., 2008). However, we argue that this is not sufficient in the context of drug reviews. Various factors, particularly patients' health conditions and their medical experiences, might significantly influence how patients react to a single drug. In this study, we created a sentence-level labeled dataset and adopted a deep learning approach to classify each sentence into one of four topics discussed by reviewers: 1) drug effectiveness, 2) side effects, 3) medical conditions, and 4) lifestyle and disease management behaviors. The successful adoption of the advanced text-mining technique in this study enabled us to extract deep and latent medical signals from drug reviews and uncover a series of findings that were not revealed by prior studies. Therefore, our study has the potential to stimulate the broader usage of advanced deep learning and NLP techniques in other business domains to enhance the depth and robustness of the textual content analysis.

The current study is organized as follows. Section 2 reviews the related literature and identifies research gaps. Section 3 presents the theoretical framework and hypothesis development. Section 4 describes the research method, including the data collection, variable operationalizations, and model specifications. Section 5 presents and interprets the estimation results. Section 6 discusses the research contributions, implications, and future research directions.

2. Literature Review

2.1. Prior Review Helpfulness Research

The significance of online reviews has long been recognized as playing an integral role in both consumers' online and offline purchasing decisions by providing consumers valuable information about product or service quality (Adams et al., 2017; Law et al., 2017; Liang et al., 2021; Zhu & Zhang, 2010). However, the vast number of online reviews has created an information overload problem that makes it practically impossible for consumers to read all reviews relevant to their situation (Brynjolfsson & Smith, 2000). Consumers often find it difficult to locate useful information and are overwhelmed by the number of alternatives, which can have negative consequences for both users and providers (Bawden & Robinson, 2009). Previous research has pointed out that consumers often only need a small set of reviews to obtain sufficient decision-making information (Mudambi & Schuff, 2010). Accordingly, researchers have investigated and identified various factors that impact perceived helpfulness.

As shown in Table 1, we conducted a comprehensive literature review and found that existing studies have emphasized the two predominate determinants of review helpfulness: 1) review characteristics, including review depth, readability, ratings, polarity, and emotions, and comprehensiveness, and 2) reviewer characteristics, including self-disclosed information (e.g., identity and location), expertise (e.g., activeness and experiences), reputation (e.g., expert or elite badge), and social influence (e.g., number of friends or followers). A consistent finding across this research is that, in addition to content-based factors, such as textual content embedded in the reviews, source-based factors, such as reviewer identity, also matter.

Table 1: Major Determinants of Perceived Review Helpfulness in Existing Studies

Factor category	Specific factor	Definitions and operationalizations	Representative references
Reviewer-related factors	Self-information disclosure	Disclosure of personal information: name, photo, location, reviewer identity	Xia, 2023, Liu & Hu, 2021, Sun et al., 2019, Siering et al., 2018, Karimi & Wang, 2017, Hong et al., 2017, Baek et al., 2012, Racherla & Friske, 2012
	Activeness / Experience	Number of reviews published by a reviewer or how long a reviewer has been a member	Xia, 2023, Moro & Esmerado, 2020, Bilal et al., 2021, Sun et al., 2019, Choi & Leon, 2020, Racherla & Friske, 2012, Cheng & Ho, 2015, Li & Huang, 2020, Liang et al., 2019
	Reputation	Whether a reviewer has an expert/elite badge or reviewer ranking	Xia, 2023, Moro & Esmerado, 2020, Bilal et al., 2021, Zhou & Guo, 2017, Sun et al., 2019, Choi & Leon, 2020, Huang et al., 2015, Siering et al., 2018, Baek et al., 2012, Liang et al., 2019
	Friends / Followers	Number of followers/friends	Kwon et al., 2021, Bilal et al., 2021, Zhou & Guo, 2017, Zhu et al., 2014, Cheng & Ho, 2015
Review-related factors	Rating	Review rating, usually ranging from one to five stars, in linear and/or quadratic terms	Xia, 2023, Liu & Hu, 2021, Moro & Esmerado, 2020, Zhou & Guo, 2017, Zhu et al., 2014, Mudambi & Schuff, 2010, Huang et al., 2015, Hong et al., 2017, Yin et al., 2014, Y. Wang et al., 2019, Chou et al., 2022, Kwon et al., 2021, Liang et al., 2019, Baek et al., 2015, Liang et al., 2021
	Volume	Number of reviews/ratings received by a product/service	Zhu et al., 2014, Karimi & Wang, 2017, Choi & Leon, 2020, Y. Wang et al., 2019, Liang et al., 2019
	Age	Days elapsed after the review was posted	Xia, 2023, Liu & Hu, 2021, Kwon et al., 2021, Zhou & Guo, 2017, Sun et al., 2019, Chatterjee, 2020, Zhu et al., 2014, Y. Wang et al., 2019, Chou et al., 2022, Salehan & Kim, 2016
	Depth	Review length: total word count of a review	Liu & Hu, 2021, Moro & Esmerado, 2020, Kwon et al., 2021, Zhou & Guo, 2017, Chatterjee, 2020, Karimi & Wang, 2017, Choi & Leon, 2020, Mudambi & Schuff, 2010, Huang et al., 2015, Ren & Hong, 2019, Yin et al., 2014, Chou et al., 2022, Baek et al., 2012, Salehan & Kim, 2016, Li & Huang, 2020, Mauro et al., 2021, Liang et al., 2021
	Readability	Ease of understanding reviews, measured by readability indices	Xia, 2023, Liu & Hu, 2021, Zhu et al., 2014, Fresneda & Gefen, 2019, Hong et al., 2017, Yin et al., 2014, Y. Wang et al., 2019, Chou et al., 2022, Liang et al., 2019, Liang et al., 2021
	Aspects / Comprehensiveness	Number of unique aspects covered in a review	Xia, 2023, Liu & Hu, 2021, Sun et al., 2019, Hong et al., 2017, Chou et al., 2022, Sun et al., 2019, Racherla & Friske, 2012
	Polarity / Emotions	Sentiment or emotions embedded in a review	Xia, 2023, Chen & Farn, 2020, Kwon et al., 2021, Chatterjee, 2020, Siering et al., 2018, Yin et al., 2014, Chou et al., 2022, Baek et al., 2012, Salehan & Kim, 2016, Mauro et al., 2021

2.2. Online Healthcare Community and Reviews

Online healthcare communities have become integral resources where patients seek information and support from other patients (Yan & Tan, 2014). It is estimated that "72 percent of U.S. adult Internet users search for health-related information online, 26 percent have read others' comments on their experiences, and 18 percent have gone online to look for others with similar health concerns" (Yan & Tan, 2017, p. 12). Consequently, a rich body of literature has explored how online healthcare communities provide essential support for patients' disease management and medical decision-making. Specifically, Yan & Tan (2014) revealed that patients tend to learn from others, and social support among patients can help them improve their health conditions and disease management. Yan et al. (2015) highlighted the uniqueness of the medical context and argued that the traits of patients' medical conditions largely influence their social interactions and connections in online communities. Thus, interacting with those who share similar conditions is important to patients.

As a major source of health information, online healthcare reviews, which are written by patients and contain their firsthand experiences, provide essential information to other patients in their medical decision-making (Shah et al., 2021; Yan & Tan, 2017). For instance, using hospital reviews as the research context, Ranard et al. (2016) found that Yelp hospital reviews can supplement valuable information regarding patients' perceptions of healthcare quality. In the context of doctor reviews, Segal et al. (2012) revealed that online physician reviews could be potentially adopted to identify surgeons with higher-quality care. Using patients' treatment reviews, Yan & Tan (2017) examined how patient consensus on treatment experiences affected how other patients perceived the effectiveness of their treatments. Particularly, in contrast to the quality information on drug labels, which are obtained from clinical trials and often incomplete (Yan & Tan, 2017), drug reviews contain the pragmatic and personal experiences from the patients' perspective and could help other patients assess a potential drug more accurately (Adusumalli et al., 2015).

However, despite the essential role of medical reviews in patients' decision-making, the vast amount of reviews creates an information overload problem. Hence, identifying helpful reviews, which could provide patients with relevant quality information and personal experiences, is both managerially and practically important. Although review helpfulness has been widely studied in various domains, especially in e-commerce, hospitality, and tourism, less attention has been paid to helpfulness perception in the context of medical reviews. Among the few exceptions are Alodadi & Zhou (2016), who examined a series of factors, including linguistic and semantic features, that can be used to predict the helpfulness of physician reviews. Yan et al. (2020) also assessed physicians' reviews and showed that review-related signals, such as depth and readability, and service-related signals, such as quality and popularity, impacted helpfulness perceptions. We found two studies that explored the helpfulness of drug reviews: Chou et al. (2022) used drug reviews as one of three datasets in their empirical modeling, mainly for comparison purposes; Zheng et al. (2021) explored the impacts of medical domain features, including sentiments and the ratio of medical words, both of which demonstrated significant power in predicting drug review helpfulness. However, despite the research efforts, there is a lack of systematic study of the in-depth medical signals embedded in online healthcare reviews and how these medical signals would influence the review helpfulness perceptions.

2.3. Deep Learning in Text Mining

The above discussion indicates that online reviews contain rich signals and provide valuable information to help consumers with their decision-making. However, it is difficult to quantify complex information from unstructured texts, especially given the unparalleled amount of data available. As a result, text mining, the process of transforming unstructured text into structured and meaningful information, is gaining popularity, and various text mining techniques have been developed (Adamopoulos et al., 2018). Traditional machine learning approaches, which leverage handcrafted features, have been widely used to uncover useful information from online reviews, including sentiment analysis (Chatterjee, 2020; S. Yang et al., 2020), emotion detection (Chen & Farn, 2020; Ren & Hong, 2019), and writing/linguistic style analysis (Mitra & Jenamani, 2021; Yang et al., 2021). However, the key limitation of traditional machine learning is that handcrafted features require tedious labor and extensive domain knowledge. Thus, it is time-consuming and intellectually challenging to build a set of high-quality features from the data (LeCun et al., 2015).

Recent advances in artificial intelligence (AI), especially in deep learning, have resulted in NLP systems with impressive performance in solving a variety of text mining tasks without having to extract features manually (LeCun et al., 2015). Deep learning belongs to neural networks and is different from traditional machine learning in that features can be learned automatically from raw data (Guo et al., 2018). Deep learning has made astonishing progress in various healthcare applications, including health informatics, biomedicine, and medical image analysis (Esteva et al., 2019; Miotto et al., 2018). While many concerns need to be addressed before their wider application in medicine, such as interpretability, performance, and data privacy, it has been widely accepted that the healthcare industry could benefit immensely from deep learning due to the complexity of medical problems and extensive domain knowledge requirements, both of which can be mitigated by the self-learning capabilities of deep learning (Miotto et al., 2018).

Concerning drug reviews, prior studies have mainly adopted deep learning techniques for the detection and extraction of adverse drug events to support pharmacovigilance (Șerban et al., 2019; Xia, 2022). To the best of our knowledge, no previous work has applied deep learning techniques to extract in-depth medical signals from drug reviews to assess helpfulness perception. We argue that, given the unique characteristics of drugs, particularly the information asymmetry caused by patients' lack of medical expertise, potential readers will depend more on in-depth information embedded in the review text. Such high information dependence has been further magnified by the fact that drug performance varies between patients. However, we have yet to understand what specific information embedded in drug reviews influences helpfulness perceptions. This study therefore proposes the use of deep learning–based NLP techniques to extract the information discussed by patients (at the sentence level) and assign the information to the following topics: 1) drug effectiveness, 2) side effects, 3) medical conditions, and 4) lifestyle and disease management behaviors. We compared our proposed advanced text mining model with a range of baseline algorithms and confirmed its superior performance in extracting in-depth medical signals from drug reviews.

3. Theoretical Background and Hypothesis Development

Our efforts to understand what specific information about patients' health conditions and medical experiences influence perceived review helpfulness are based on signaling theory, which has been used in studies on e-commerce and consumers' online behaviors (e.g., Connelly et al., 2011; Yang et al., 2019), as well as review helpfulness (e.g., Filieri et al., 2021; Siering et al., 2018; Yang et al., 2019). A fundamental assumption of signaling theory is the presence of asymmetric information, and signaling theory suggests that the exchange of information signals is valuable for reducing information asymmetry and influencing people's behaviors (Connelly et al., 2011; Spence, 1974). Specifically, four key elements need to be considered: sender, receiver, signals, and signaling environment (Spence, 2002). In the online review context, the senders of signals are often the reviewers who have experience with the product, while the receivers are the potential users who do not yet have experience with the product and are looking for relevant information (Siering et al., 2018). Given that there is information asymmetry between senders and receivers, signals are designed to communicate and carry information from those with more information (senders) to those with less information (receivers) (Connelly et al., 2011). Moreover, signaling theory posits that the environment may influence how signals are processed (Siering et al., 2018). Because signaling must happen in a specific context, the signaling environment (e.g., the degree of uncertainty) is crucial in determining which signals to employ and assessing the effectiveness of the signals (Connelly et al., 2011; Shah et al., 2022).

While a range of prior studies have adopted signaling theory in review helpfulness research (e.g., Siering et al., 2018; Yang et al., 2019), we believe that this theory is particularly useful in the medical context, which is highly information asymmetric and involves elevated risks (Shah et al., 2021). Although information asymmetry and risks also exist in general commodities, they are less severe than in the medical industry, where patients usually lack essential medical expertise. Notably, signaling theory provides a solid theoretical foundation for a systematic study of patients' information needs and how various signals impact people's behaviors. In the drug review context, the signal sender is a reviewer who has experience taking the drug, while the signal receiver is (likely) a patient who has limited or no experience taking the drug. From the signaling perspective, patients (receivers) seek relevant information (signals) from reviewers (senders) regarding the quality of a drug. The core of signaling theory is to explain the roles of various types of signals—in our context, medical signals—in providing valuable information to reduce information asymmetry (Spence, 2002). It is worth noting that signaling theory also emphasizes how the impacts of signals differ depending on the context (Li et al., 2019), which can include different drug and reviewer characteristics. Therefore, we propose a conceptual framework that builds on signaling theory to examine how various in-depth medical signals extracted from drug reviews may impact helpfulness perception (see Figure 2). In addition to direct impacts, we accounted for several drug and reviewer characteristics and examined their moderating effects to obtain a comprehensive understanding of how the impacts of medical signals vary according to signaling environments. In the following subsections, we elaborate on the hypothesis development.

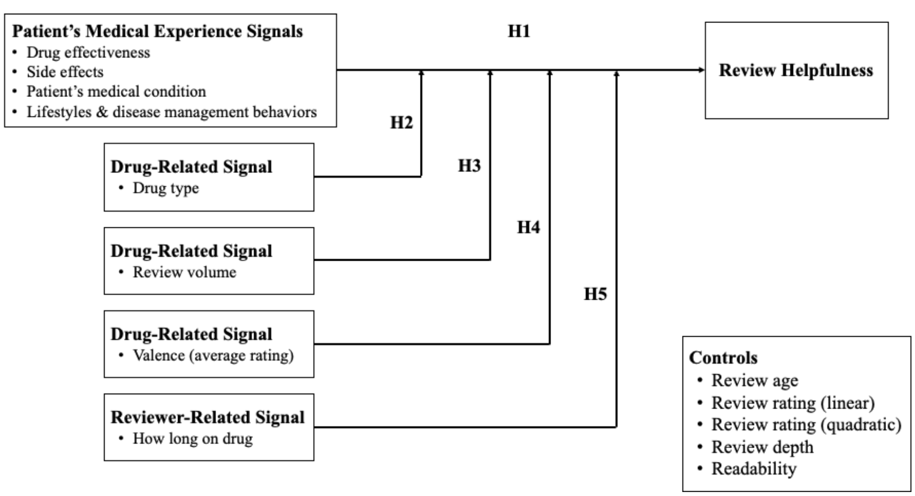


Figure 2: Research Model of Drug Review Helpfulness

3.1. Patients' Health Conditions and Medical Experiences (Direct Effects)

A core aspect of signaling theory is that, while different types of signals can be included, the presence of appropriate signals in medical reviews is vital to how they are perceived by readers because different signals can express different information and eventually lead to different perceptions (Connelly et al., 2011; Yang et al., 2019). Prior studies have indicated that patients share their knowledge and experiences by discussing health-related topics on online healthcare platforms (Gao et al., 2015; Yan & Tan, 2014; Yan et al., 2019). In particular, Joshi & Abdelfattah (2021) outlined three topics that tend to be included in drug reviews: medical conditions (i.e., reason for taking the drug), drug effectiveness, and side effects. Other studies have shown that patients' lifestyle and disease management behaviors play a key role in ensuring drug effectiveness (Chee et al., 2011; Divaris, 2017; Hsueh et al., 2016): drug performance varies dramatically for people who report different lifestyle and disease management behaviors. Accordingly, four signals—(1) drug effectiveness, (2) side effects, (3) medical conditions, and (4) lifestyle and disease management behaviors—have been included in our proposed framework.

Previous literature suggests that a major motivation for patients to seek out reviews based on others' experiences is to assess treatment effectiveness, adjust their expectations about a certain treatment, and compare their treatment progress with that of others (Yan & Tan, 2017). It is worth noting that the same treatment could have varying degrees of effectiveness among patients, resulting in different perceptions of the same drug (Wicks et al., 2010). Consequently, drug effectiveness is an essential assessment of the drug and is often considered as an integral aspect of reviews, which helps to improve other patients' understanding of how their drugs are working (Yan & Tan, 2017). Additionally, the prior study has supported a strong correlation between drug satisfaction and perceived drug effectiveness, suggesting that the perceived effectiveness could influence patients' medication adherence and compliance behaviors, which are essential in disease management, especially for chronic conditions (Leonard et al., 2020). Therefore, among many quality signals, the drug effectiveness signal plays an integral role in patients' perception of a particular drug as it relates directly to the positive consequences of the drug (Yan et al., 2019). Thus, we hypothesize the following:

H1a. *The inclusion of information signals related to <u>drug effectiveness</u> is <u>positively</u> correlated with perceived review helpfulness.*

In addition to drug effectiveness, another important piece of information is the drug's potential side effects, which further distinguish drug reviews from those of general commodities (Yan et al., 2019). A patient's choice of drug often determines their quality of life and can even be a matter of life and death (Liu & Chen, 2015). Awareness of negative experiences can improve patients' expectations of and preparation for potential risks (Yan et al., 2019). Accordingly, information signals about side effects may facilitate a more comprehensive evaluation of drug quality. It is impractical for patients to assess the actual impacts of drugs based solely on information from drug labels or even from their healthcare providers (Wicks et al., 2010), and patients may look for information signals about drug side effects from drug reviews that contain personal experiences written from the perspectives of their peers (i.e., patients with the same or similar conditions). Hence, we hypothesize the following:

H1b. *The inclusion of information signals related to <u>drug side effects</u> is <u>positively</u> correlated with perceived review helpfulness.*

As discussed above, a further unique characteristic of drugs that differentiates them from other general commodities is that the consumer base is more specialized—they are patients with certain diseases that can be treated by that drug. While this highly specific consumer base naturally groups patients with similar diseases together, it is worth highlighting that patients might be at different stages of the disease, experiencing different symptoms or different degrees of severity, or have different circumstances related to their health, such as other comorbidities and unique personal and family medical histories (Gage, 2013; O'Grady et al., 2008; Shah et al., 2021). Several studies have revealed that the treatment outcomes for a single drug differ among patients with different health conditions (Chee et al., 2011; Divaris, 2017; Hsueh et al., 2016). Therefore, it is critical for patients to consult with peers who have similar conditions so that the experiences being used to evaluate drugs are relevant and applicable (Yan et al., 2015; Yan & Tan, 2014). Given that the relative efficacy of drugs is highly dependent on the health conditions of the patients who write reviews, we hypothesize that the signals of their health conditions will help other patients identify reviewers with similar medical conditions, find information relevant to their situations more easily, and perceive these reviews as being more helpful:

H1c. *The inclusion of information signals related to the <u>health conditions</u> of patients is <u>positively</u> correlated with perceived review helpfulness.*

In addition to health conditions, previous research has highlighted that patients' lifestyles not only influence the risks of developing a specific disease, but also impact drug performance (Raghupathi & Raghupathi, 2014). Furthermore, proper drug management—following medical guidance and taking drugs at the right dose and at the right time—plays a key role in managing patients' health conditions and illnesses (Horne & Weinman, 1999; Yan et al., 2019). Yan & Tan (2017) revealed that a major reason for dissatisfaction with treatment outcomes is the failure to follow medical guidance. Given that lifestyles and disease management behaviors play a key role in influencing drug performance, we theorize that the inclusion of related signals could provide a more complete picture of drug quality and result in improved helpfulness perception. Hence, we hypothesize the following:

H1d. *The inclusion of information signals related to the lifestyles and disease management behaviors of patients is positively correlated with perceived review helpfulness.*

3.2. Moderating Effects

While the essence of signaling theory lies in the analysis of the roles of various types of signals (Spence, 2002), another core aspect of signaling theory is the signaling environment, or the situations in which these signals are transmitted (Siering et al., 2018; Spence, 2002). In particular, the signaling environment influences the perceived usefulness of certain signals because it can restrict the ability of signal receivers to accurately evaluate the signals (Connelly et al., 2011). It is worth noting that, on the whole, the signaling environment is under-researched (Connelly et al., 2011); thus, we intend to extend signaling theory by exploring how various signaling environments, including drug and reviewer characteristics, influence the impacts of patients' health conditions and medical experience signals on helpfulness perception.

3.2.1. Patients' Health Signals and Drug Types

Given the information asymmetry caused by patients' lack of medical expertise, patients often seek information from peers who share similar conditions to inform their own drug and disease management and promote better health outcomes. The diseases could fall into one of two categories: short-term or long-term (chronic) diseases (Fox, 2011). Compared to patients with short-term diseases, patients with chronic diseases (e.g., hypertension, heart diseases, diabetes, and cancer) often rely on long-term medical treatment to improve their quality of life, and self-management is especially important to them (Fox, 2011). However, it is even difficult for medical professionals to assess drug efficacy quickly, let alone for patients with limited medical expertise, which can result in uncertainty about a drug's long-term performance (Iyengar et al., 2011). As a result, patients with chronic diseases tend to take responsibility for their own health behaviors and become part of the decision-making process by seeking out online healthcare communities to obtain information about health conditions and opinions on medication performance and side effects (Mao et al., 2013). Furthermore, patients with chronic diseases are likelier to seek out informal sources of peer-to-peer assistance than to consult with healthcare providers (Fox, 2011). We suggest that different diseases create different information needs for patients. Concerning drugs used to treat chronic diseases, patients often seek more types of information signals and are likelier to pay attention to detailed content. Thus, we hypothesize the following:

H2a. *For drugs used to treat chronic diseases, the influence of drug effectiveness signals on perceived review helpfulness will be stronger.*

H2b. *For drugs used to treat chronic disease, the influence of drug side-effect signals on perceived review helpfulness will be stronger.*

H2c. *For drugs used to treat chronic disease, the influence of health condition signals on perceived review helpfulness will be stronger.*

H2d. *For drugs used to treat chronic diseases, the influence of lifestyle and disease management behavior signals on perceived review helpfulness will be stronger.*

3.2.2. Patients' Health Signals and Review Volume

Prior studies often suggested that a high review volume indicates the positive performance of a product, as a large number of reviews contributed by many people reduce their sense of risk and uncertainty about the quality of a product (Chevalier & Mayzlin, 2006). According to signaling theory, transmitting signals may minimize uncertainty, and the signal's effectiveness largely depends on the existing degree of uncertainty: in a high-uncertainty environment, signals have a greater impact than in a low-uncertainty situation (Siering et al., 2018). In other words, as review volume increases, which implies high product quality and reduced uncertainty, the impact of patients' medical signals is expected to weaken. However, given the unique characteristics of drugs, we believe that a high review volume does not necessarily correlate with high quality. For general commodities, purchase decisions are mainly based on consumers' preferences. Hence, it is reasonable to infer that review volume is related to product quality. For drugs, however, selection decisions are often not voluntary; instead, they largely depend on recommendations from medical experts (Yan & Tan, 2017). Hence, review volumes are not necessarily indicative of a drug's quality; rather, they may indicate simply that a drug has been used by many patients. Yan & Tan (2017) suggested that an increased volume of medical reviews would increase the variety of opinions from different patients and result in higher uncertainty. Therefore, we hypothesize that the impact of patients' health conditions and medical experience signals is stronger rather than weaker:

H3a. *The higher the review volume of a drug, the stronger the influence of drug effectiveness signals on perceived review helpfulness.*

H3b. *The higher the review volume of a drug, the stronger the influence of drug side effect signals on perceived review helpfulness.*

H3c. *The higher the review volume of a drug, the stronger the influence of the health condition signals of patients on perceived review helpfulness.*

H3d. *The higher the review volume of a drug, the stronger the influence of the lifestyle and disease management behavior signals of patients on perceived review helpfulness.*

3.2.3. Patients' Health Signals and Valence

Valence is a user-generated quality signal that refers to the overall evaluation of one product (Filieri, 2015). Thus, the valence of a drug is an important signal for patients because such information is essentially a crowdsourced opinion about a product's quality (Choi & Leon, 2020). However, previous studies have generated inconsistent results regarding the impact of valence on review helpfulness. For instance, while Pan & Zhang (2011) identified a significant positive relationship between valence and review helpfulness, Racherla & Friske (2012) found a negative relationship. Prior studies conducted in a medical context have revealed that patients strongly favor reducing losses over boosting gains (Chen et al., 2020; Shah et al., 2021), which suggests that patients are more prone to perceive damages than benefits. Thus, we propose that interaction effects exist between patients' health conditions, medical experience signals, and drug valence. Specifically, when valence (average rating) is low, potential readers expect to learn about the negative effects of the drug to reduce potential losses. In contrast, when valence is high, potential readers expect to learn about the effectiveness of the drug and under what conditions the drug achieves a certain level of performance. In light of this discussion, we hypothesize the following:

H4a. *The higher the drug valence, the stronger the influence of drug effectiveness signals on perceived review helpfulness.*

H4b. *The higher the drug valence, the weaker the influence of drug side effect signals on perceived review helpfulness.*

H4c. *The higher the drug valence, the stronger the influence of health condition signals on perceived review helpfulness.*

H4d. *The higher the drug valence, the stronger the influence of the lifestyle and disease management behavior signals of patients on perceived review helpfulness.*

3.2.4. Patients' Health Signals and How Long the Reviewer Has Taken a Drug

While some reviewers post reviews soon after their experience with a product or service, others tend to post reviews much later (Huang et al., 2016). If consumers have been using a product or service for a long time, they will theoretically accumulate more experience and develop a comprehensive evaluation of it (Chatterjee, 2020). This is especially true for experience goods (e.g., hotels and restaurants), as their quality can only be evaluated after the products and services have been purchased and consumed (Mudambi & Schuff, 2010). A unique characteristic of drugs is that many quality effects—effectiveness and side effects—often take time to surface (Chee et al., 2011; Katkade et al., 2018). Conversely, consumers can provide accurate information about experience goods shortly after consumption. For drugs, however, quality can be difficult to ascertain after initial consumption (Chee et al., 2011; Katkade et al., 2018). Therefore, if a patient has been taking a drug for only a short period of time, potential readers may not perceive them as being able to provide accurate quality assessments, and patients are likely to prefer information provided by experienced reviewers. As a note, patients' health conditions and their lifestyle and disease management behaviors are not believed to interact with how long they have taken the drug. Hence, we hypothesize:

H5a. *The longer a reviewer has taken a drug, the stronger the influence of drug effectiveness signals on perceived review helpfulness.*

H5b. *The longer a reviewer has taken a drug, the stronger the influence of drug side effect signals on perceived review helpfulness.*

H5c. *There is no interaction between how long a reviewer has taken a drug and the health condition signals in terms of the effect on perceived review helpfulness.*

H5d. *There is no interaction between how long a reviewer has taken a drug and lifestyle and disease management behavior signals in terms of the effect on perceived review helpfulness.*

4. Research Methodology

4.1. Data Collection

To validate our research model and test the corresponding hypotheses, we harvested a large drug review dataset comprising 9,908 unique drugs and 147,169 corresponding reviews posted between October 2007 and September 2021 from WebMD (accessed September 10, 2021), a leading healthcare publisher in the United States. For each drug, we collected the general description as well as its indications (e.g., the diseases and conditions for which the drug is recommended as a treatment). For each review, we collected a series of variables, including helpful votes, ratings, review content, when the review was posted, and how long the reviewer had been taking the drug. For consistency with many previous studies (e.g., Chatterjee, 2020; Filieri et al., 2021; Wang & Karimi, 2019; Zhou & Guo, 2017), our model included all reviews.

4.2. Variable Operationalizations

Table 2 summarizes a list of variables, including dependent variables (DV), independent variables (IV), their operationalizations, and descriptive statistics. As discussed in Section 2.1 (also see Table 1), previous literature has consistently shown that both review and reviewer factors are strong predictors of review helpfulness. Additionally, several studies have revealed that product characteristics, such as overall rating (e.g., Liu & Hu, 2021; Moro & Esmerado, 2020), popularity (e.g., Filieri et al., 2021; Lei et al., 2021), and product types (e.g., Mudambi & Schuff, 2010; Siering et al., 2018), influence consumers' perceptions of review helpfulness. Therefore, the variables presented in Table 2 include the following groups of variables: reviews, reviewers, and drugs. It is worth highlighting that, unlike other review platforms (e.g., Yelp, Amazon), most drug review platforms do not ask reviewers to create profile pages. Given the unique medical context—a patient might only ever use one drug and thus contribute only one review—this is reasonable. Consequently, most reviewer attributes, such as reviewer expertise, reputation, and social influence, which have been found to play key roles in review helpfulness perception, are not readily available on drug review platforms. Nevertheless, we were able to include one reviewer variable: how long the reviewer had been taking the drug (*time_on_drug*; see Table 2).

The correlation matrix presented in Table 3 shows that there was no problem with multicollinearity in our dataset (absolute values of correlations were less than or equal to 0.37). In addition, we inspected the variance inflation factors (VIFs). All VIFs were below 10 (1.06–5.43); therefore, it is unlikely that our results were affected by multicollinearity (Hair et al., 2010). While most variables were consistent with the definitions and operationalizations adopted in previous studies (see Table 1), we adopted deep learning to build a sentence-level classification model to capture the specific topic of each sentence. In the following section, we elaborate on the detailed steps.

Table 2: Variable Representations, Operationalizations, and Descriptive Statistics

Variable type	Variable level	Variable	Operationalization	Mean	Std.	Max.	Min.
DV	Individual Review	Helpfulness	# of helpfulness votes received by a review	7.16	9.26	157	0
IV	Individual Review	Medical_condition	# of sentences in a review discussing **patient's medical conditions**	0.57	0.84	10	0
		Effectiveness	# of sentences in a review discussing **drug's effectiveness**	0.27	0.54	7	0
		Side_effects	# of sentences in a review discussing **drug's side effects**	0.62	0.88	10	0
		Lifestyle_management	# of sentences in a review discussing **patient's lifestyle and disease management behaviors**	0.25	0.54	8	0
	Reviewer	Time_on_drug	How long the reviewer has taken the drug: 1 (< 1 month), 2 (1–6 months), 3 (6–12 months), 4 (1–2 years), 5 (2–5 years), 6 (> 10 years)	2.36	1.65	6	0
	Drug	Review_volume	# of reviews (in hundreds) a drug has received	9.54	11.70	42.67	0.01
		Valence	Average rating of a drug	3.53	0.48	5	1
		Drug_type	Whether the drug is used to treat chronic disease. Operationalized as a dummy variable, with 1 indicating the drug is used to treat chronic disease and 0 otherwise	0.43	0.49	1	0
Controls	Individual Review	Review_age	How long (in months) a review has been posted	118.55	38.31	169	0
		Rating	Star rating (1-5) associated with each review	3.53	1.31	5	1
		$Rating^2$	Quadratic term of star rating	14.18	8.54	25	1
		Depth	Number of words in the review	64.20	60.30	775	1
		Readability	Operationalized by Coleman-Liau index	29.77	24.66	319.02	0.4

Table 3: Correlations between Variables

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Helpfulness	1												
2. Medical_condition	0.17	1											
3. Effectiveness	0.08	0.27	1										
4. Side_Effects	0.17	0.35	0.06	1									
5. Lifestyle_management	0.09	0.34	0.05	0.10	1								
6. Time_on_drug	0.05	0.06	0.05	0.01	0.03	1							
7. Review_volume	0.02	0.04	0.02	0.04	-0.01	0.09	1						
8. Valence	-0.06	-0.01	0.07	-0.04	-0.01	0.15	0.14	1					
9. Drug_type	0.07	0.06	0.07	0.02	-0.01	0.10	0.28	0.16	1				
10. Review_age	0.06	0.01	0.01	-0.02	-0.01	0.05	0.22	0.14	0.09	1			
11. Rating	0.05	0.01	0.13	-0.02	-0.01	0.21	0.05	0.37	0.07	0.09	1		
12. Depth	0.21	0.37	0.32	0.34	0.23	0.02	-0.05	-0.02	-0.03	-0.08	-0.01	1	
13. Readability	0.21	0.36	0.31	0.33	0.24	0.02	-0.05	-0.02	-0.03	-0.09	-0.01	0.37	1

4.2.1. Signals Related to Patients' Medical Experiences

Our main research interest was the role played by various patients' diagnoses, health conditions, and medical experience signals in helpfulness perceptions. While some recent studies have adopted various text mining techniques to extract information from text content, deep learning has enabled us to explore more latent and informative content with promising performance (e.g., in terms of accuracy, precision, recall, and f-measure) and to dig deeper into the qualitative characteristics of review content (LeCun et al., 2015). We used this methodology to conduct an in-depth content analysis to reveal, at the sentence level, the specific medical information signals that patients discuss in their reviews. Specifically, we built an annotated dataset and labeled each sentence manually. Based on the theoretical framework presented in Figure 2, we used five labels: 1, medical conditions; 2, effectiveness; 3, side effects; 4, lifestyle and disease management behaviors; and 5, others. The last label was applied when the content did not relate to any of the first four topics. Table 4 presents sample sentences for each label category.

Table 4: Sample Sentences for each Topic Label

Topic label	Sample sentences
Medical conditions	• *I have suffered with chronic depression for most of my life.* • *I take this medication in combination with baclofen for unspecified tremors and anxiety of my medical condition.*
Effectiveness	• *It wasn't until my psychiatrist added Abilify that I experienced almost complete relief of my depression.* • *Have been taking the drug for over 5 years and it works like a charm.*
Side effect	• *For me, this medication has caused an increase in my anxiety, depression and suicidal thoughts.* • *First 2 weeks into it were terrible - chest pains, racing heartbeat, neck stiffness and body pains.*
Lifestyles and disease management behaviors	• *Female, 64, great health as a rule, non-smoker, non-drinker.* • *My dose is only 5mg/day which I take every 8am.*
Others	• *The new TD medications are way more than I can afford $$$$$.* • *Honestly no idea why they prescribe this.*

We then hired three annotators with medical backgrounds to assign one of the five labels to each sentence. Specifically, we randomly selected 4,000 reviews from our data collection and separated them into sentences, each of which was labeled by two annotators. If there was a disagreement, we followed the common practice of "adjudicating by an expert" (Davani et al., 2022; Waseem & Hovy, 2016; Xia et al., 2017): the third annotator would consider both opinions and make the final decision. After data annotation, we randomly assigned 70% of the reviews to the training set, 15% to the validation set, and 15% to the testing set (see Table 5).

Table 5: Sentence-level Labeled Topic Distributions

Datasets	# of reviews	# of sentences	Topic label	# of each label	Percentage
Training set	2,800	13,989	Medical conditions	3,190	22.8%
			Effectiveness	1,486	10.6%
			Side effect	3,408	24.4%
			Lifestyles and disease management behaviors	1,548	11.1%
			Others	4,357	31.1%
Validation set	600	3,081	Medical conditions	758	24.6%
			Effectiveness	348	11.3%
			Side effect	856	27.8%
			Lifestyles and disease management behaviors	292	9.5%
			Others	827	26.8%
Testing set	600	2,984	Medical conditions	688	23.0%
			Effectiveness	340	11.4%
			Side effect	746	25.0%
			Lifestyles and disease management behaviors	322	10.8%
			Others	888	29.8%

We then built a text classification model to determine the topic label for each sentence. As discussed in Section 2.3, deep learning, which belongs to the representation-learning method, has achieved astonishing performance in solving a range of NLP tasks without having to extract features manually. Specifically, Bidirectional Encoder Representations from Transformers (BERT), which applies transformers, has achieved state-of-the-art performance in many NLP tasks (Devlin et al., 2018). Hence, we adopted a BERT-based text classification model for our labeled dataset (see Table 6 for the model architecture, hyperparameters, and evaluation results). To our knowledge, we are the first to build a labeled dataset that covers multiple aspects of the patient health information embedded in drug reviews. For this reason, there is no existing state-of-the-art model to which we can compare. Nevertheless, as others have done (e.g., Colón-Ruiz & Segura-Bedmar, 2020; Li et al., 2021), we compared our model against several common benchmark models, including the traditional TF-IDF method, bi-directional Long Short-Term Memory (bi-LSTM), and convolutional neural network (CNN). Our proposed model achieved promising performance in classifying each sentence into a topic group, outperforming all three baseline models with large margins in all metrics (see Table 7). The trained model was then applied to the complete dataset to extract the topic that each sentence discussed.

Table 6: Model Architecture, Hyperparameters, and Evaluation Results

Deep learning architecture	Topic label	Overall accuracy	Recall	Precision	F1
BERT model: BERT base, 12 Transformer blocks, 12 self-attention heads **Hidden size**: 768 **Learning rate:** 5e-5 **Batch size:** 32 **Epochs:** 20 **Dimension:** 200 **bi-LSTM**: 100 hidden size, 0.5 dropout, 32 batch size, and 0.01 learning rate	Medical conditions	0.91	0.96	0.90	0.93
	Effectiveness		0.86	0.96	0.91
	Side effect		0.92	0.96	0.94
	Lifestyles and disease management behaviors		0.90	0.88	0.90
	Others		0.86	0.93	0.89
Overall			**0.90**	**0.93**	**0.91**

Table 7: Performance Comparisons of Topic Classification of the Proposed Model and Three Common Benchmarks

Model	Overall accuracy	Average recall	Average precision	Average F1
Our BERT-based model	**0.91**	**0.90**	**0.93**	**0.91**
TF-IDF	0.53	0.52	0.61	0.56
Bi-LSTM	0.67	0.71	0.66	0.69
CNN	0.72	0.77	0.70	0.73

4.2.2. Drug Types

Drugs were separated into two groups based on whether they were indicated for treating chronic diseases. Based on the drug indications we collected from WebMD, we checked whether the drugs were indicated for the treatment of chronic diseases or conditions, such as hypertension, heart disease, diabetes, or cancer. A full list of chronic diseases and conditions was obtained from an article published by Taylor et al. (2015), who summarized a comprehensive list of chronic diseases and conditions. Based on this operationalization, 1,292 drugs and 63,009 corresponding reviews were placed in the chronic disease group, with the remaining 8,616 drugs and 84,160 corresponding reviews placed in the nonchronic disease group.

4.2.3. Control Variables

To improve the validity of our results, we controlled a number of variables that could impact perceived review helpfulness to account for additional review-specific effects, as confirmed by previous studies (see Table 1), including review age, rating (both linear and quadratic terms), depth, and readability. Table 2 summarizes the operationalizations, most of which are consistent with the operationalizations widely adopted in previous studies (see Table 1).

4.3. Empirical Model

Our dependent variable (review helpfulness) was a nonnegative count variable and was overdispersed; the variance (85.67) largely exceeded its mean (7.16). Consistent with a number of prior studies (e.g., Chatterjee, 2020; Kwon et al., 2021; X. Wang et al., 2019; Wang et al., 2020; Zhou & Guo, 2017), negative binomial regression, which can correct for overdispersion and account for omitted variable bias, was a better choice than the Poisson model, which assumes that the variance and the mean of the count variable are equal. Consequently, we employed a negative binomial regression with robust standard errors to estimate the following model, which included interaction effects:

$$\begin{aligned} Helpfulness = {} & a + \beta(Patients'\,Health\ Conditions) + \gamma(Drug\_type) + \phi(Review\_volume) + \delta(Valence) \\ & + \theta(Time\_on\_drug) + \pi(Controls) + \upsilon\,(Patients'\ Health\ Conditions \times Drug\_type) \\ & + \rho\,(Patients'\ Health\ Conditions \times Review\_volume) \\ & + \tau(Patients'\,Health\ Conditions \times Valence) \\ & + \sigma(Patients'\,Health\ Conditions \times Time\_on\_drug) + \varepsilon \end{aligned}$$

where a is the constant term; β represents the coefficients that correspond to the four patient health condition and medical experience signals extracted from the reviews; γ, ϕ, δ, and θ represent the coefficients corresponding to drug type, review volume, valence, and how long the reviewer has taken the drug, respectively; π represents the coefficients for the control variables; υ, ρ, and τ; σ represent the coefficients for all interaction effects; and ε denotes the error term.

5. Results and Discussions

Using a negative binomial regression estimation, a hierarchical regression analysis was conducted. The variables were entered into the regression model according to the assumed causal order. The control variables were entered into the first block, followed by the variables for the main effects (***H1a–H1d***) in the second block, and those for moderating effects (***H2a–H2d, H3a–H3d, H4a–H4d, H5a–H5d***) in the third block. This resulted in three regression models (see Table 8).

5.1. Results for Main Effects

Focusing on the main effects regarding the patients' medical signals embedded in reviews, Model 2 (see Table 8) revealed that the discussion of *drug effectiveness* ($\beta = 0.030$, $p < 0.001$), *drug side effects* ($\beta = 0.081$, $p < 0.001$), *patients' medical conditions* ($\beta = 0.062$, $p < 0.001$), and *patients' lifestyle and disease management behaviors* ($\beta = 0.018$, $p < 0.001$) are all positively correlated with perceived review helpfulness. Hence, ***H1a***–***H1d*** were supported.

5.2. Results for Moderating Effects

The moderating effects were investigated, and the results are summarized in Model 3 (see Table 8). Regarding the interaction effects between the patients' medical experience signals and drug type, the results suggest that compared to nonchronic diseases, the positive effect of including *drug effectiveness* ($\beta = 0.048$, $p < 0.001$), *drug side effects* ($\beta = 0.085$, $p < 0.001$), *patients' medical conditions* ($\beta = 0.055$, $p < 0.001$), and *patients' lifestyle and disease management behaviors* ($\beta = 0.051$, $p < 0.001$) on review helpfulness will all be stronger, supporting ***H2a–H2d***.

Regarding the interaction effects between the patients' medical experience signals and review volume, the results suggest that, with an increase in *review volume*, the positive effect of including *drug side effects* ($\beta = 0.002$, $p < 0.001$), *patients' medical conditions* ($\beta = 0.002$, $p < 0.001$), and *patients' lifestyle and disease management behaviors* ($\beta = 0.003$, $p < 0.001$) on review helpfulness will all be stronger, supporting ***H3b***, ***H3c***, and ***H3d***. We did not observe a significant interaction between the *drug effectiveness* signal and *review volume* ($\beta = 0.001$, $p = 0.139$), which suggests that drug effectiveness information is not the main signal for which potential readers are looking when they read reviews of popular drugs. Hence, ***H3a*** was not supported.

Table 8: Estimation Results of Negative Binomial Regression Analyses

Signal group	Related hypotheses	Independent variables	Dependent variables: *Helpfulness*		
			Model (1)	**Model (2)**	**Model (3)**
Control Variables	***NA***	Review_age	0.003(8.1e-5)***	0.003(8.2e-5)***	0.003(8.2e-5)***
		Rating	-0.316(0.013)***	-0.295(0.013)***	-0.289(0.013)***
		$Rating^2$	0.055(0.002)***	0.057(0.002)***	0.056(0.002)***
		Depth	0.004(4e-4)***	0.004(4e-4)***	0.004(4e-4)***
		Readability	-0.005(0.001)***	-3.7e-4(0.001)	-2.9e-4(0.001)
Main effects					
Patients' Medical Experience Signals	***H1 (H1a–H1d)***	Effectiveness		0.030 (0.006)***	-0.206(0.047)***
		Side_effect		0.081(0.004)***	0.295(0.033)***
		Medical_condition		0.062(0.005)***	0.252(0.04)***
		Lifestyle_management		0.018(0.007)***	0.161(0.052)***
Drug-Related Signals	***NA***	Drug_type		0.164 (0.005)***	0.126(0.009)***
	NA	Review_volume		0.001(3e-4)***	0.001(3e-4)***
	NA	Valence		-0.304(0.007)***	-0.263(0.009)***
Patient-Related Signal	***NA***	Time_on_drug		0.030(0.002)***	0.036(0.003)***
Interaction effects					
Patients' Medical Experience Signals * Drug-Related Signals	***H2 (H2a–H2d)***	Effectiveness*Drug_type			0.048(0.012)***
		Side_effect* Drug_type			0.085(0.009)***
		Medical_condition*Drug_type			0.055(0.010)***
		Lifestyle_management*Drug_type			0.051(0.011)***
	H3 (H3a–H3d)	Effectiveness* Review_volume			0.001(5e-4)
		Side_effect* Review_volume			0.002(4e-4)***
		Medical_condition* Review_volume			0.002(4e-4)***
		Lifestyle_management* Review_volume			0.003(6e-4)***
	H4 (H4a–H4d)	Effectiveness*Valence			0.058(0.013)***
		Side_effect*Valence			-0.078(0.010)***
		Medical_condition*Valence			0.043(0.012)***
		Lifestyle_management*Valence			-0.042(0.015)**
Patients' Medical Experience Signals * Patient-Related Signal	***H5 (H5a–H5d)***	Effectiveness*Time_on_drug			0.009(0.003)**
		Side_effect*Time_on_drug			0.012(0.003)***
		Medical_condition*Time_on_drug			-0.003(0.003)
		Lifestyle_management*Time_on_ drug			-0.004(0.004)
		N	147,169	147,169	147,169
		Log Pseudolikelihood	-440351.98	-438,452.18	-438,270.57
		Wald Chi-Square	55,866.17	54,860.19	54,759.81
		P Value Chi-Square	<0.0001	<0.0001	<0.0001

Note: Standard errors in parentheses. *p < 0.05, **p < 0.01, ***p < 0.001

Regarding the interaction effects between the patients' *medical experience* signal and drug valence, the results suggest that, with an increase in *drug valence*, the positive effect of including *side effect* signals ($\beta = -0.078$, $p < 0.001$) and *patients' lifestyle and disease management behavior* signals ($\beta = -0.042$, $p < 0.01$) would be weaker, while the positive effects of including *drug effectiveness* ($\beta = 0.058$, $p < 0.001$) and *patients' medical conditions* ($\beta = 0.043$, $p < 0.001$) would be stronger. Hence, ***H4a***, ***H4b***, and ***H4c*** were all supported, while ***H4d*** was not.

In addition, as discussed in Section 3.2.4, while consumers can provide accurate information shortly after product consumption, drug quality information can take time to surface. ***H5a*** and ***H5b*** posit that the positive impacts of drug effectiveness and side-effect signals will be stronger when reviewers have been taking the drug for a longer period of time. The results in Model 3 indeed suggest significant positive interactions between *drug effectiveness* and *how long the reviewer has taken the drug* (β = 0.009, p < 0.01) as well as between *drug side effects* and *how long the reviewer has taken the drug* (β = 0.012, p < 0.001). In addition, consistent with our expectations, the *length of experience* with a drug had no significant interactions with the signals involving *patients' medical conditions* (β = -0.003, p = 0.919) and *lifestyle and disease management behaviors* (β = -0.004, p = 0.267). Hence, ***H5a***–***H5d*** are supported.

5.3. Discussions of Findings

We uncovered several interesting results that have not yet been revealed by prior studies. Specifically, by taking advantage of the deep learning–based text mining approach, we extracted sentence-level medical signals from online drug reviews with satisfactory performance. These signals include patients' medical conditions, drug effectiveness, side effects, and lifestyle and disease management behaviors. Our findings indicate that, while all four signals demonstrate positive impacts on the helpfulness perception of drug reviews, their impacts are moderated differently by drug and reviewer characteristics. Specifically, contrary to our hypothesis (***H3a***), we did not observe a significant interaction between drug effectiveness and review volume. One possible explanation is that, while having more reviews does not necessarily indicate overall drug quality, the review volume suggests that the effectiveness of the drugs has been commonly recognized, especially by medical experts who prescribe them. Hence, drug effectiveness information is not the main signal in which potential readers are interested. In addition, our results suggest that with an increase in drug valence, the positive effect of including patients' lifestyle and disease management behavior signals on review helpfulness will be weaker, which contradicts our hypothesis (***H4d***). This can be explained by the unique characteristics of drugs: unsatisfactory experiences could result from patients' failure to adhere to a prescription regimen (Clifford et al., 2006; Yan & Tan, 2017). Consequently, when the average rating is low, potential readers will likely expect detailed information about patients' personal lifestyle and disease management behaviors. In addition, our results suggest that the longer reviewers have been taking the drug, the higher the probability that their reviews will be perceived as helpful. This seems reasonable, as consumers perceive reviews from reviewers who have used a product for a longer period of time as more knowledgeable. This is especially true for drugs, as many drug effects take time to surface.

In addition to the main variables of interest, we revealed some interesting findings from the control variables. First, the results of Model 2 (Table 8) confirmed that both *review age* (β = 0.003, p < 0.001) and *depth* (β = 0.004, p < 0.001) were positively and significantly related to review helpfulness, which suggests that reviews that were posted earlier have a higher probability of getting more votes and that users are likelier to vote for more in-depth reviews. These results are consistent with many prior studies whose authors reported the positive impacts of both variables (e.g., Mudambi & Schuff, 2010; Salehan & Kim, 2016; Y. Wang et al., 2019). Regarding *ratings*, our results show that the linear term has a significant negative effect (β = -0.295, p < 0.001) and the quadratic term has a significant positive effect (β = 0.057, p < 0.001) on review helpfulness. Hence, we found a U-shaped relationship, implying that extreme reviews (highly positive or negative ratings) would be more helpful than moderate reviews (neural ratings). Previous research has revealed mixed findings regarding the role of extreme reviews. Many studies have stated that extreme reviews are considered more helpful (e.g., Kwon et al., 2021; Liang et al., 2019; Yin et al., 2014), while some studies have suggested a negative impact because consumers often perceive extreme reviews as untrustworthy (e.g., Chatterjee, 2020; Salehan & Kim, 2016). Meanwhile, a number of studies have indicated that reviews conveying more negative content tend to receive more helpfulness votes, a phenomenon often referred to as negativity bias (Baumeister et al., 2001; Lee et al., 2017). One possible explanation for these mixed findings is that the impacts vary with respect to different products. In the context of drug reviews, potential readers may pay more attention to drug effectiveness and side effects, representing the positive and negative aspects (extreme aspects) of drugs.

For *readability*, interestingly, we found that it has no significant impact on perceived review helpfulness (β = -3.7e-4, p = 0.706), conflicting with many prior studies that have demonstrated that the higher the readability, the likelier a review will be perceived as helpful (e.g., Fang et al., 2016; Ghose & Ipeirotis, 2010; Kuan et al., 2015; Lei et al., 2021; Malik & Hussain, 2018). To ensure that our finding was robust, in addition to the Coleman-Liau index, we tried different readability measurements, including the Gunning Fog Index (Gunning, 1969), the SMOG Index (Mc Laughlin, 1969), and Flesch Reading Ease (Kincaid et al., 1975), and obtained consistent results. We believe this phenomenon can be explained by the fact that drug reviews often contain rich medical terms and concepts, both of which can make the reviews less readable. As discussed previously, drugs have a very specific audience: patients with similar health conditions, who are often familiar with terms and concepts pertaining to their medical conditions. Therefore, although including medical content may decrease readability, such information has no negative impact on how patients perceive review helpfulness. This finding is supported by several prior studies, in which the authors illustrated that the readability of the review does not influence diagnosticity or persuasion power (Fresneda & Gefen, 2019; Hong et al., 2017; Zhu et al., 2014).

6. Contributions and Future Work

This study has shed light on how the detailed medical signals embedded in online drug reviews affect their helpfulness perceptions, according to potential readers. We suggest that, given the unique characteristics of drugs, particularly the information asymmetry caused by patients' lack of medical domain knowledge, potential readers will likely depend more on textual information embedded in the review contents. Thus, it is important for us to explore the in-depth medical signals from drug reviews so that other patients have a clear idea of the reviewers' health conditions and experiences. Specifically, we used the signaling theory as the theoretical underpinning for our proposed research model to holistically address patients' information demands. To test the proposed hypotheses, we adopted a deep learning–based text mining approach to extract medical signals from drug reviews obtained from WebMD, a popular drug review website, and employed negative binomial regression to study their impacts on perceived review helpfulness. Based on the results of our experiments and empirical modeling, we answered three research questions.

Specifically, for **RQ1**, we showed that our proposed BERT-based deep learning model achieved superior performance when extracting in-depth medical signals without the need for manual feature engineering steps. To answer **RQ2**, by empirically analyzing the extracted in-depth medical signals, we found that signals related to patients' health conditions and medical experiences have significant impacts on review helpfulness. To answer **RQ3**, we further explored the moderating effects. Our results revealed that the impacts of medical signals on review helpfulness perceptions largely depend on signaling environments such as drug characteristics (i.e., drug type, volume, and valence) and reviewer characteristics (i.e., how long the reviewer has been on the drug). Overall, our findings suggest that, without considering the role of in-depth information signals embedded in reviews, explaining review helpfulness can be biased. This study makes both theoretical and methodological contributions to support future review helpfulness research, especially in the medical context, and offers practical implications for various stakeholders in medical decision-making.

6.1. Theoretical Implications

First, we contribute to a better understanding of consumer perceptions of review helpfulness in the context of drug reviews. Given that an unprecedented amount of online data often causes information overload problems, being able to identify helpful information is a great advantage, especially for medical products, as medical decisions often dramatically impact patients' quality of life. Although many prior studies have explored the perception of review helpfulness from various perspectives, our discussion of the special characteristics of medical products demonstrates that reviewers' unique health conditions and medical experiences are of great significance to patients' perceptions of reviews. Specifically, based on the online healthcare community literature, we identified four main topics that patients often discuss in their online posts: medical conditions, drug effectiveness, side effects, and lifestyle and disease management behaviors. To the best of our knowledge, we are the first to consider these in-depth medical signals and explore their impacts on the perception of drug review helpfulness. A comprehensive analysis of the detailed topic signals in the medical context provides a theoretical foundation to assess patients' decision-making processes. Although review helpfulness has been extensively explored, our findings suggest that there is still much to uncover within the medical context.

Second, we proposed a theoretical framework grounded in signaling theory to explain how the detailed medical signals embedded in drug reviews affect perceived review helpfulness. Our discussion demonstrates that, unlike general commodities, a lack of necessary medical domain knowledge puts patients at an information disadvantage, which significantly enhances information asymmetry and the potential risks of choosing the wrong drug. As a result, signaling theory, which assumes information asymmetry and suggests how various signals can be communicated to reduce uncertainties, is especially appropriate for analyzing patients' information needs in the medical domain. Specifically, our findings indicate that not only do patients' medical signals demonstrate positive impacts on the helpfulness perception of drug reviews, but their impacts are moderated differently by drug and reviewer characteristics. Given that the signaling environment is an under-researched aspect of signaling theory (Connelly et al., 2011), our findings also contribute to signaling theory by shedding light on the signaling environments in which the impacts of patients' medical signals on review helpfulness are likely to be weakened or strengthened. Although signaling theory has been widely adopted by prior review helpfulness research, we believe this theory can play a more integral role in the medical context, and we hope that this study will encourage future researchers to further explore its utility in healthcare research.

6.2. Managerial and Practical Implications

Our findings also offer important managerial and practical implications that can benefit multiple stakeholders, particularly reviewers and online drug review platforms.

For *reviewers*, a major motivation for writing reviews is to provide helpful information to future users. Our focus on in-depth medical signals provides direct insights that can help reviewers understand what specific medical information (e.g., medical conditions, drug effectiveness, side effects, and disease management behaviors) should be included in their reviews to make drug reviews more helpful to other patients. In addition, reviewers should take the signaling environment, such as drug type, review volume, and valence, into consideration. For example, for drugs with high valence, reviewers should include more information about drug effectiveness and their own medical conditions; for drugs with low valence, reviewers should emphasize the potential side effects, as readers will likely expect the negative aspects of the drug, and their lifestyle and disease management behaviors, as readers may expect these factors would contribute to unsatisfactory performance. Hence, our findings offer practical insights for reviewers on how to write helpful reviews effectively and efficiently to benefit other patients in their medication choices; these findings also pertain primarily to healthcare platforms. While reviewers on other platforms (e.g., Yelp, Amazon) often write multiple reviews and accumulate writing experience, reviewers on drug review platforms may not have access to multiple learning opportunities if they only have experience with just one drug and, thus, contribute only one review.

In addition, our findings can enhance how healthcare platforms understand the roles that different signals play in drug review perceptions. We believe that healthcare platforms can act on this information in two ways. First, our results can help them identify the reviews likely to be perceived as helpful and incorporate patients' health condition signals to identify and recommend helpful reviews to meet the information needs of potential readers. Second, online platforms can go one step further by developing a more dynamic and intuitive review interface. More concretely, most review platforms offer only a single text box for adding textual content (see Figure 3). We suggest that separate text boxes be created for different essential aspects of the products. More importantly, what text boxes are presented to reviewers can be dynamically adjusted based on product and reviewer characteristics. Based on our findings, for popular drugs with high ratings, the guidelines may simply ask reviewers to consider disclosing more information about their medical conditions when writing reviews (Figure 4a). If a reviewer has been taking a drug for a long time, including more information about drug effectiveness and potential side effects could make their reviews more helpful (Figure 4b). Thus, our study offers online platforms useful insights to make them more user-friendly for both users and reviewers.

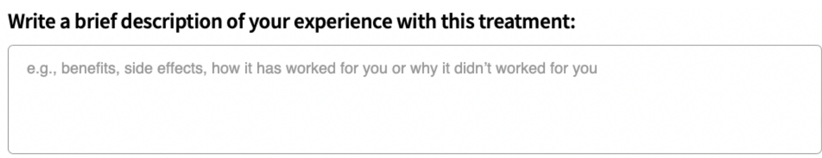
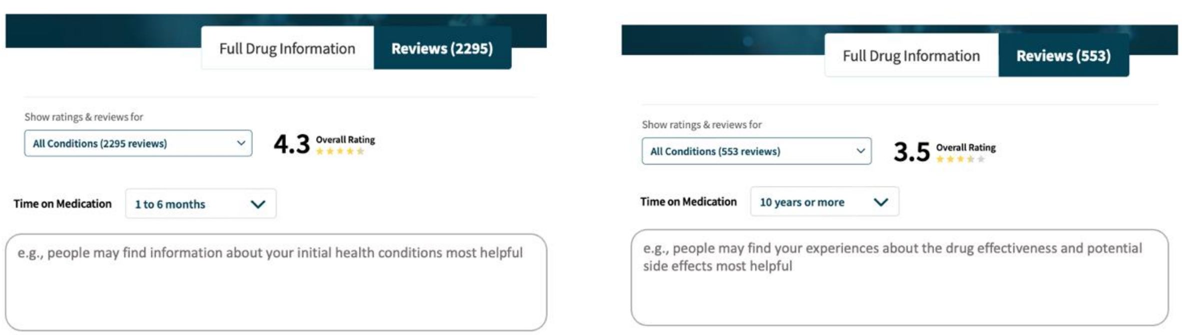


Figure 3: Most Review Platforms Offer Only a Single Textbox for Adding Textual Content (Obtained from WebMD).



(a) (b)

Figure 4: Review Platforms can Use Our Findings to Improve Their Review Interface to Dynamically Adjust What Textboxes Are Presented to Reviewers at the Time of Writing: (a) For popular drugs with high valence, we recommend including information about patients' medical conditions; (b) For reviewers who have been taking a drug for a very long time, we recommend including information about drug effectiveness and potential side effects.

6.3. Methodological Implications

From a methodological perspective, this study adopted an advanced deep learning–based text mining technique to extract deep and latent signals regarding patients' health conditions and medical experiences. We built a deep learning–powered, sentence-level text classification model to classify each sentence into one of four topics that patients often discuss in online healthcare communities. While adopting text mining techniques to extract qualitative signals from reviews has become a trend in recent years, most extant studies have focused mainly on review sentiment, emotions, and writing styles at the whole-review level. Very few studies have explored detailed information signals (e.g., topics) at the individual-sentence level. Our successful adoption of deep learning for extracting in-depth signals at a granular level enabled us to uncover a series of findings not discovered in prior studies and illustrated the usefulness of such methods as alternatives to traditional text mining methods. This study thus contributes methodologically by showing the potential of deep learning techniques for extracting latent but informative attributes directly from unstructured review texts.

In addition, it is worth noting that while we have applied a deep learning technique to study the impacts of specific topics on the perceived helpfulness of reviews, our proposed methodology can be extended to other research contexts. For instance, in the context of restaurant reviews in the era of the COVID-19 pandemic, researchers could use our method to develop a deep learning–based text mining model to classify each sentence in a review into one of the predefined topics (e.g., food quality, dining environment, service quality, safety measures) and study their granular impacts on consumers' decision-making during the pandemic. Hence, while our research mainly expands the text mining literature by employing deep learning–based NLP methodology in studying medical reviews, we encourage future researchers to explore a wide range of contexts in which advanced deep learning techniques could be applicable.

6.4. Limitations and Future Research

Naturally, our study is not exempt from limitations, many of which offer directions for future research. First, the use of secondary data necessitates certain limitations that may suggest alternative interpretations. Most notably, it is challenging to establish a causal relationship using naturally occurring reviews, even though we have controlled for a variety of variables that are believed to affect review helpfulness. Future researchers might consider adopting controlled experiments to infer causal relationships. Second, our analysis was conducted exclusively with drug reviews from WebMD. We exhaustively included all the drugs and corresponding reviews we collected from the website to improve the robustness of our analyses. However, it is still possible that not all of our findings can be generalized to other platforms. Future researchers may adopt our research framework to assess drug reviews from other platforms to improve generalizability. Third, although our results suggest that reviewers' medical conditions and drug types play integral roles in perceived helpfulness, the impacts of specific diseases are largely unknown. It is our belief that patients with arthritis will have significantly different information needs than patients with cancer, although both are chronic diseases. Future research could further explicate the impact of specific diseases on review perceptions.

Acknowledgment

The author thanks the editor-in-chief, associate editor, and two anonymous reviewers for their constructive suggestions throughout the review process. The author was supported in part by the New Faculty Summer Research and Faculty Reassigned Time at Elon University.

REFERENCES

Adamopoulos, P., Ghose, A., & Todri, V. (2018). The impact of user personality traits on word of mouth: Text-mining social media platforms. *Information Systems Research*, *29*(3), 612-640.

Adams, D. Z., Gruss, R., & Abrahams, A. S. (2017). Automated discovery of safety and efficacy concerns for joint & muscle pain relief treatments from online reviews. *International Journal of Medical Informatics, 100*, 108-120.

Adusumalli, S., Lee, H., Hoi, Q., Koo, S.-L., Tan, I. B., & Ng, P. C. (2015). Assessment of web-based consumer reviews as a resource for drug performance. *Journal of Medical Internet Research, 17*(8), e211.

Alodadi, N., & Zhou, L. (2016). Predicting the helpfulness of online physician reviews. In *2016 IEEE International Conference on Healthcare Informatics (ICHI)* (pp. 1-6).

Baek, H., Ahn, J., & Choi, Y. (2012). Helpfulness of online consumer reviews: Readers' objectives and review cues. *International Journal of Electronic Commerce, 17*(2), 99-126.

Baek, H., Lee, S., Oh, S., & Ahn, J. (2015). Normative social influence and online review helpfulness: Polynomial modeling and response surface analysis. *Journal of Electronic Commerce Research, 16*(4), 290-306.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5*(4), 323-370.

Bawden, D., & Robinson, L. (2009). The dark side of information: Overload, anxiety and other paradoxes and pathologies. *Journal of Information Science, 35*(2), 180-191.

Bilal, M., Marjani, M., Hashem, I. A. T., Malik, N., Lali, M. I. U., & Gani, A. (2021). Profiling reviewers' social network strength and predicting the "helpfulness" of online customer reviews. *Electronic Commerce Research and Applications, 45*, 101026.

Brynjolfsson, E., & Smith, M. D. (2000). Frictionless commerce? A comparison of Internet and conventional retailers. *Management Science, 46*(4), 563-585.

Chatterjee, S. (2020). Drivers of helpfulness of online hotel reviews: A sentiment and emotion mining approach. *International Journal of Hospitality Management, 85*, 102356.

Chee, B. W., Berlin, R., & Schatz, B. (2011). Predicting adverse drug events from personal health messages. In *Proceedings of the American Medical Informatics Association (AMIA) Annual Symposium* (pp. 217–226).

Cheerkoot-Jalim, S., & Khedo, K. K. (2020). A systematic review of text mining approaches applied to various application areas in the biomedical domain. *Journal of Knowledge Management*, *25*(3), 642-668.

Chen, L., Baird, A., & Straub, D. (2020). A linguistic signaling model of social support exchange in online health communities. *Decision Support Systems, 130*, 113233.

Chen, M. J., & Farn, C. K. (2020). Examining the influence of emotional expressions in online consumer reviews on perceived helpfulness. *Information Processing & Management, 57*(6), 102266.

Cheng, Y. H., & Ho, H. Y. (2015). Social influence's impact on reader perceptions of online reviews. *Journal of Business Research, 68*(4), 883-887.

Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research*, *43*(3), 345-354.

Choi, H. S., & Leon, S. (2020). An empirical investigation of online review helpfulness: A big data perspective. *Decision Support Systems*, *139*, 113403.

Chou, Y. C., Chuang, H. H. C., & Liang, T. P. (2022). Elaboration likelihood model, endogenous quality indicators, and online review helpfulness. *Decision Support Systems, 153*, 113683.

Clifford, S., Barber, N., Elliott, R., Hartley, E., & Horne, R. (2006). Patient-centred advice is effective in improving adherence to medicines. *Pharmacy World and Science*, *28*(3), 165-170.

Colón-Ruiz, C., & Segura-Bedmar, I. (2020). Comparing deep learning architectures for sentiment analysis on drug reviews. *Journal of Biomedical Informatics, 110*, 103539.

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of Management*, *37*(1), 39-67.

Davani, A. M., Díaz, M., & Prabhakaran, V. (2022). Dealing with disagreements: Looking beyond the majority vote in subjective annotations. *Transactions of the Association for Computational Linguistics, 10*, 92-110.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Divaris, K. (2017). Fundamentals of precision medicine. *Compendium of Continuing Education in Dentistry, 38*(8 Suppl), 30-32.

Duan, W., Gu, B., & Whinston, A. B. (2008). The dynamics of online word-of-mouth and product sales—An empirical investigation of the movie industry. *Journal of Retailing, 84*(2), 233-242.

Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., Cui, C., Corrado, G., Thrun, S., & Dean, J. (2019). A guide to deep learning in healthcare. *Nature Medicine, 25*(1), 24-29.

Fang, B., Ye, Q., Kucukusta, D., & Law, R. (2016). Analysis of the perceived value of online tourism reviews: Influence of readability and reviewer characteristics. *Tourism Management, 52*, 498-506.
Filieri, R. (2015). What makes online reviews helpful? A diagnosticity-adoption framework to explain informational and normative influences in e-WOM. *Journal of Business Research, 68*(6), 1261-1270.
Filieri, R., Raguseo, E., & Vitari, C. (2021). Extremely negative ratings and online consumer review helpfulness: The moderating role of product quality signals. *Journal of Travel Research, 60*(4), 699-717.
Fox, S. (2011). *The social life of health information, 2011*. California Healthcare Foundation.
Fox, S., & Duggan, M. (2013). *Tracking for health.* Pew Research Center.
Fresneda, J. E., & Gefen, D. (2019). A semantic measure of online review helpfulness and the importance of message entropy. *Decision Support Systems, 125*, 113117.
Gage, E. A. (2013). Social networks of experientially similar others: Formation, activation, and consequences of network ties on the health care experience. *Social Science & Medicine, 95*, 43-51.
Gao, G., Greenwood, B. N., Agarwal, R., & McCullough, J. S. (2015). Vocal minority and silent majority. *MIS Quarterly, 39*(3), 565-590.
Ghose, A., & Ipeirotis, P. G. (2010). Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics. *IEEE Transactions on Knowledge and Data Engineering, 23*(10), 1498-1512.
Gunning, R. (1969). The fog index after twenty years. *Journal of Business Communication, 6*(2), 3-13.
Guo, J., Zhang, W., Fan, W., & Li, W. (2018). Combining geographical and social influences with deep learning for personalized point-of-interest recommendation. *Journal of Management Information Systems, 35*(4), 1121-1153.
Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis (7th ed.).* Prentice Hall.
Hong, H., Xu, D., Wang, G. A., & Fan, W. (2017). Understanding the determinants of online review helpfulness: A meta-analytic investigation. *Decision Support Systems, 102*, 1-11.
Hong, Y. A., Liang, C., Radcliff, T. A., Wigfall, L. T., & Street, R. L. (2019). What do patients say about doctors online? A systematic review of studies on patient online reviews. *Journal of Medical Internet Research, 21*(4), e12521.
Horne, R., & Weinman, J. (1999). Patients' beliefs about prescribed medicines and their role in adherence to treatment in chronic physical illness. *Journal of Psychosomatic Research, 47*(6), 555-567.
Hsueh, P. Y. S., Chang, H., & Ramakrishnan, S. (2016). Next generation wellness: A technology model for personalizing healthcare. In *Healthcare Information Management Systems* (pp. 355-374). Springer.
Huang, A. H., Chen, K., Yen, D. C., & Tran, T. P. (2015). A study of factors that contribute to online review helpfulness. *Computers in Human Behavior, 48*, 17-27.
Huang, N., Burtch, G., Hong, Y., & Polman, E. (2016). Effects of multiple psychological distances on construal and consumer evaluation: A field study of online reviews. *Journal of Consumer Psychology, 26*(4), 474-482.
Iyengar, R., Van den Bulte, C., & Valente, T. W. (2011). Opinion leadership and social contagion in new product diffusion. *Marketing Science, 30*(2), 195-212.
Joshi, S., & Abdelfattah, E. (2021). Multi-class text classification using machine learning models for online drug reviews. In *2021 IEEE World AI IoT Congress (AIIoT)* (pp. 262-267).
Karimi, S., & Wang, F. (2017). Online review helpfulness: Impact of reviewer profile image. *Decision Support Systems, 96*, 39-48.
Katkade, V. B., Sanders, K. N., & Zou, K. H. (2018). Real world data: An opportunity to supplement existing evidence for the use of long-established medicines in health care decision making. *Journal of Multidisciplinary Healthcare, 11*, 295.
Kincaid, J. P., Fishburne Jr, R. P., Rogers, R. L., & Chissom, B. S. (1975, February 1). *Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel*. Retrieved from: https://apps.dtic.mil/sti/citations/ADA006655.
Kuan, K. K., Hui, K.-L., Prasarnphanich, P., & Lai, H.-Y. (2015). What makes a review voted? An empirical investigation of review voting in online review systems. *Journal of the Association for Information Systems, 16*(1), 48-71.
Kwon, W., Lee, M., Back, K.-J., & Lee, K. Y. (2021). Assessing restaurant review helpfulness through big data: Dual-process and social influence theory. *Journal of Hospitality and Tourism Technology, 12*(2), 177-195.
Law, D., Gruss, R., & Abrahams, A. S. (2017). Automated defect discovery for dishwasher appliances from online consumer reviews. *Expert Systems with Applications, 67*, 84-94.
LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436-444.
Lee, M., Jeong, M., & Lee, J. (2017). Roles of negative emotions in customers' perceived helpfulness of hotel reviews on a user-generated review website: A text mining approach. *International Journal of Contemporary Hospitality Management, 29*(2), 762-783.

Lei, Z., Yin, D., & Zhang, H. (2021). Focus within or on others: The impact of reviewers' attentional focus on review helpfulness. *Information Systems Research, 32*(3), 801-819.

Leonard, K. V., Robertson, C., Bhowmick, A., & Herbert, L. B. (2020). Perceived treatment satisfaction and effectiveness facilitators among patients with chronic health conditions: A self-reported survey. *Interactive Journal of Medical Research, 9*(1), e13029.

Li, J., Tang, J., Yen, D. C., & Liu, X. (2019). Disease risk and its moderating effect on the e-consultation market offline and online signals. *Information Technology & People, 32*(4), 1065-1084.

Li, M., Huang, L., Tan, C.-H., & Wei, K.-K. (2013). Helpfulness of online product reviews as seen by consumers: Source and content features. *International Journal of Electronic Commerce, 17*(4), 101-136.

Li, M., & Huang, P. (2020). Assessing the product review helpfulness: Affective-cognitive evaluation and the moderating effect of feedback mechanism. *Information & Management, 57*(7), 103359.

Li, X., Cui, M., Li, J., Bai, R., Lu, Z., & Aickelin, U. (2021). A hybrid medical text classification framework: Integrating attentive rule construction and neural network. *Neurocomputing, 443*, 345-355.

Liang, S., Schuckert, M., & Law, R. (2019). How to improve the stated helpfulness of hotel reviews? A multilevel approach. *International Journal of Contemporary Hospitality Management*, *31*(2), 953-977.

Liang, T. P., Cheng, J. C., Saini, V., & Hsu, J. S. C. (2021). Is being helpful good enough for online reviews? Exploring the role of information credibility and data source through meta-analysis. *Journal of Electronic Commerce Research, 22*(4), 336-362.

Liu, X., & Chen, H. (2015). A research framework for pharmacovigilance in health social media: Identification and evaluation of patient adverse drug event reports. *Journal of Biomedical Informatics, 58*, 268-279.

Liu, Y., & Hu, H.-f. (2021). Online review helpfulness: The moderating effects of review comprehensiveness. *International Journal of Contemporary Hospitality Management, 33*(2), 534-556.

Malik, M., & Hussain, A. (2018). An analysis of review content and reviewer variables that contribute to review helpfulness. *Information Processing & Management, 54*(1), 88-104.

Mao, J. J., Chung, A., Benton, A., Hill, S., Ungar, L., Leonard, C. E., Hennessy, S., & Holmes, J. H. (2013). Online discussion of drug side effects and discontinuation among breast cancer survivors. *Pharmacoepidemiology and Drug Safety, 22*(3), 256-262.

Mauro, N., Ardissono, L., & Petrone, G. (2021). User and item-aware estimation of review helpfulness. *Information Processing & Management, 58*(1), 102434.

Mc Laughlin, G. H. (1969). SMOG grading-a new readability formula. *Journal of Reading, 12*(8), 639-646.

Miotto, R., Wang, F., Wang, S., Jiang, X., & Dudley, J. T. (2018). Deep learning for healthcare: Review, opportunities and challenges. *Briefings in Bioinformatics, 19*(6), 1236-1246.

Mitra, S., & Jenamani, M. (2021). Helpfulness of online consumer reviews: A multi-perspective approach. *Information Processing & Management, 58*(3), 102538.

Moro, S., & Esmerado, J. (2020). An integrated model to explain online review helpfulness in hospitality. *Journal of Hospitality and Tourism Technology, 12*(2), 239-253.

Mudambi, S. M., & Schuff, D. (2010). Research note: What makes a helpful online review? A study of customer reviews on Amazon.com. *MIS Quarterly*, *34*(1), 185-200.

O'Grady, L. A., Witteman, H., & Wathen, C. N. (2008). The experiential health information processing model: Supporting collaborative web-based patient education. *BMC Medical Informatics and Decision Making, 8*(1), 1-8.

Pan, Y., & Zhang, J. Q. (2011). Born unequal: A study of the helpfulness of user-generated product reviews. *Journal of Retailing, 87*(4), 598-612.

Qazi, A., Syed, K. B. S., Raj, R. G., Cambria, E., Tahir, M., & Alghazzawi, D. (2016). A concept-level approach to the analysis of online review helpfulness. *Computers in Human Behavior, 58*, 75-81.

Racherla, P., & Friske, W. (2012). Perceived 'usefulness' of online consumer reviews: An exploratory investigation across three services categories. *Electronic Commerce Research and Applications, 11*(6), 548-559.

Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: promise and potential. *Health Information Science and Systems, 2*(1), 1-10.

Ranard, B. L., Werner, R. M., Antanavicius, T., Schwartz, H. A., Smith, R. J., Meisel, Z. F., Asch, D. A., Ungar, L. H., & Merchant, R. M. (2016). Yelp reviews of hospital care can supplement and inform traditional surveys of the patient experience of care. *Health Affairs, 35*(4), 697-705.

Ren, G., & Hong, T. (2019). Examining the relationship between specific negative emotions and the perceived helpfulness of online reviews. *Information Processing & Management, 56*(4), 1425-1438.

Salehan, M., & Kim, D. J. (2016). Predicting the performance of online consumer reviews: A sentiment mining approach to big data analytics. *Decision Support Systems, 81*, 30-40.

Segal, J., Sacopulos, M., Sheets, V., Thurston, I., Brooks, K., & Puccia, R. (2012). Online doctor reviews: Do they track surgeon volume, a proxy for quality of care? *Journal of Medical Internet Research, 14*(2), e50.

Șerban, O., Thapen, N., Maginnis, B., Hankin, C., & Foot, V. (2019). Real-time processing of social media with SENTINEL: A syndromic surveillance system incorporating deep learning for health classification. *Information Processing & Management, 56*(3), 1166-1184.

Shah, A. M., Ali, M., Qayyum, A., Begum, A., Han, H., Ariza-Montes, A., & Araya-Castillo, L. (2021). Exploring the impact of linguistic signals transmission on patients' health consultation choice: Web mining of online reviews. *International Journal of Environmental Research and Public Health, 18*(19), 9969.

Shah, A. M., Muhammad, W., & Lee, K. (2022). Examining the determinants of patient perception of physician review helpfulness across different disease severities: A machine learning approach. *Computational Intelligence and Neuroscience, 2022*.

Shah, A. M., Yan, X., Shah, S. A. A., Shah, S. J., & Mamirkulova, G. (2019). Exploring the impact of online information signals in leveraging the economic returns of physicians. *Journal of Biomedical Informatics, 98*, 103272.

Siering, M., Muntermann, J., & Rajagopalan, B. (2018). Explaining and predicting online review helpfulness: The role of content and reviewer-related signals. *Decision Support Systems, 108,* 1-12.

Spence, M. (1974). *Market signaling: Informational transfer in hiring and related screening processes*. Harvard University Press.

Spence, M. (2002). Signaling in retrospect and the informational structure of markets. *American Economic Review, 92*(3), 434-459.

Sun, X., Han, M., & Feng, J. (2019). Helpfulness of online reviews: Examining review informativeness and classification thresholds by search products and experience products. *Decision Support Systems, 124*, 113099.

Taylor, S. J., Pinnock, H., Epiphaniou, E., Pearce, G., Parke, H. L., Schwappach, A., Purushotham, N., Jacob, S., Griffiths, C. J., & Greenhalgh, T. (2015). A rapid synthesis of the evidence on interventions supporting self-management for people with long-term conditions: PRISMS–practical systematic review of self-management support for long-term conditions. *Health Services and Delivery Research, 2*, 54.

Wallace, B. C., Paul, M. J., Sarkar, U., Trikalinos, T. A., & Dredze, M. (2014). A large-scale quantitative analysis of latent factors and sentiment in online doctor reviews. *Journal of the American Medical Informatics Association, 21*(6), 1098-1103.

Wang, F., & Karimi, S. (2019). This product works well (for me): The impact of first-person singular pronouns on online review helpfulness. *Journal of Business Research, 104,* 283-294.

Wang, X., Tang, L. R., & Kim, E. (2019). More than words: Do emotional content and linguistic style matching matter on restaurant review helpfulness? *International Journal of Hospitality Management, 77,* 438-447.

Wang, Y., Wang, J., & Yao, T. (2019). What makes a helpful online review? A meta-analysis of review characteristics. *Electronic Commerce Research, 19*(2), 257-284.

Wang, Y., Wang, J., Yao, T., & Li, M. (2020). What makes peer review helpfulness evaluation in online review communities? An empirical research based on persuasion effect. *Online Information Review, 44*(6), 1267-1286.

Waseem, Z., & Hovy, D. (2016). Hateful symbols or hateful people? Predictive features for hate speech detection on twitter. In *Proceedings of the NAACL Student Research Workshop* (pp. 88-93). Association for Computational Linguistics.

Wicks, P., Massagli, M., Frost, J., Brownstein, C., Okun, S., Vaughan, T., Bradley, R., & Heywood, J. (2010). Sharing health data for better outcomes on PatientsLikeMe. *Journal of Medical Internet Research, 12*(2), e19.

Xia, L. (2022). Historical profile will tell? A deep learning-based multi-level embedding framework for adverse drug event detection and extraction. *Decision Support Systems, 160*, 113832.

Xia, L. (2023). The impacts of geographic and social influences on review helpfulness perceptions: A social contagion perspective. *Tourism Management, 95*, 104687.

Xia, L., Wang, G. A., & Fan, W. (2017). A deep learning based named entity recognition approach for adverse drug events identification and extraction in health social media. In *International Conference on Smart Health* (pp. 237-248). Springer.

Yan, L., Peng, J., & Tan, Y. (2015). Network dynamics: How can we find patients like us? *Information Systems Research, 26*(3), 496-512.

Yan, L., & Tan, Y. (2014). Feeling blue? Go online: An empirical study of social support among patients. *Information Systems Research, 25*(4), 690-709.

Yan, L., & Tan, Y. (2017). The consensus effect in online health-care communities. *Journal of Management Information Systems, 34*(1), 11-39.

Yan, L., Yan, X., Tan, Y., & Sun, S. X. (2019). Shared minds: how patients use collaborative information sharing via social media platforms. *Production and Operations Management*, *28*(1), 9-26.

Yan, X., Khan, S., & Shah, S. J. (2020). Exploring the impact of review and service-related signals on online physician review helpfulness: A multi-methods approach. In *Pacific Asia Conference on Information Systems (PACIS)*.

Yang, H., Du, H. S., & Shang, W. (2020). Understanding the influence of professional status and service feedback on patients' doctor choice in online healthcare markets. *Internet Research*, *31*(4), 1236-1261.

Yang, S., Yao, J., & Qazi, A. (2020). Does the review deserve more helpfulness when its title resembles the content? Locating helpful reviews by text mining. *Information Processing & Management*, *57*(2), 102179.

Yang, S., Zhou, C., & Chen, Y. (2021). Do topic consistency and linguistic style similarity affect online review helpfulness? An elaboration likelihood model perspective. *Information Processing & Management*, *58*(3), 102521.

Yang, S., Zhou, Y., Yao, J., Chen, Y., & Wei, J. (2019). Understanding online review helpfulness in omnichannel retailing. *Industrial Management & Data Systems*, *119*(8), 1565-1580.

Yin, D., Bond, S. D., & Zhang, H. (2014). Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews. *MIS Quarterly*, *38*(2), 539-560.

Zhang, R., & Tran, T. (2010). Helpful or unhelpful: A linear approach for ranking product reviews. *Journal of Electronic Commerce Research*, *11*(3), 220-230.

Zheng, J., Wen, P., Ji, X., Lyu, X., & Yang, Y. (2021). Helpfulness prediction of online drug reviews. In 2*021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE)*. IEEE.

Zhou, S., & Guo, B. (2017). The order effect on online review helpfulness: A social influence perspective. *Decision Support Systems*, *93*, 77-87.

Zhu, F., & Zhang, X. (2010). Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *Journal of Marketing*, *74*(2), 133-148.

Zhu, L., Yin, G., & He, W. (2014). Is this opinion leader's review useful? Peripheral cues for online review helpfulness. *Journal of Electronic Commerce Research*, *15*(4), 267-280.