Diabetic patient review helpfulness: Unpacking online drug treatment reviews by text analytics and design science approach

Yi Feng[1], Yunqiang Yin[2], Dujuan Wang[1,*], Lalitha Dhamotharan[3], Joshua Ignatius[4], Ajay Kumar[4, *]

[1] Business School, Sichuan University, Chengdu 610064, China

[2] School of Economics and Management, University of Electronic Science and Technology of China, Chengdu 610064, China

[3] Centre for Simulation, Analytics and Modelling (CSAM), University of Exeter Business School, EX4 4PU, UK

[4] Aston Business School, Aston University, B47ET, Birmingham, UK

[5] EMLYON Business School, France

Abstract: The transparency of online reviews of drug treatment in patients with diabetes supports the use of text analytics to investigate review helpfulness based on the dual-process theory and design science approach. The first purpose of our study is to explore the influences of informational elements (emotions with the degrees of different arousal, review length) and normative elements (perceived effectiveness and ease of use, and patient satisfaction) in online drug treatment reviews on review helpfulness. We also examine the moderate role of review length on the relationship between patient satisfaction and review helpfulness. The second purpose is to explore the influences of the review topics on review helpfulness. Our study reveals four essential findings. First, not all emotions significantly influence review helpfulness, and only low-arousal emotions have a significant positive influence on review helpfulness. Second, an inverted U-shaped relationship between review length and review helpfulness and a U-shaped relationship between patient satisfaction and review helpfulness are confirmed. Third, review length has a moderate influence on the inverted U-shaped relationship between patient satisfaction and review helpfulness. Finally, the review topics related to blood sugar, family medical history, dosing time and injection significantly influence review helpfulness. These findings may serve as a stepping stone for future research on review helpfulness in the healthcare context, offering guidance for patients with diabetes, design implications for platform providers, and drug improvement suggestions for pharmaceutical companies.

Keywords: Online Drug Treatment Review; Dual-Process Theory; Review Helpfulness; Design Science Approach

*: corresponding author. E-mail address: djwang@scu.edu.cn (D.Wang)

1. Introduction

A shift from paper to information technology has been undertaken in the healthcare field (Kordzadeh and Warren 2017; Baechle et al., 2020). With the rapid growth of the World Wide Web, hundreds of social media-supported health-specific platforms have also emerged, such as WebMD and Ask Patients. These platforms provide a public platform for patients to post ratings and reviews of their medication experience, which plays an important role in the management of medication for peer patients (Hautala et al. 2021; Lamano et al. 2021). Statistically, three-quarters of adults with a college education are willing to use the Internet to write online reviews and share health information. More than a quarter of Americans choose to learn about others' treatment experiences by reading online reviews on healthcare platforms (Adusumalli et al. 2015).

Typically, online reviews include star ratings for multiple different attributes and comments about the experience of using a product or service and critiques about product or service attributes. Platform providers, Product providers and users highly depend on online helpful reviews (Kuan et al. 2015; Yang et al. 2021; Ray et al. 2021). Platforms can achieve greater attention and traffic by identifying and providing helpful reviews, and product providers can understand the views of patients and develop product improvement, as well as potential readers get useful information and make more informed decisions from the reviews that are perceived as more helpful (Filieri et al., 2018; Srivastava and Kalro, 2019; Choi and Leon, 2020). However, the proliferation of patient online reviews and the wealth of available information produce information overload (Park and Hong 2018; Risselada et al. 2018; Srivastava and Kalro, 2019; Filieri et al., 2018), making it difficult for platform providers, product providers and potential readers to sieve, process and identify the helpful ones among numerous reviews. Hence, investigating the determinants of perceived review helpfulness (*RH*) of patient online reviews benefits potential patients, platforms and product providers.

Online reviews are highly unstructured, but advanced techniques, such as text mining, design science, and topic modeling, can generate quantitative insights (Kuan et al. 2015; Mudambi and Schuff 2010; Zhao et al. 2019; Yang et al. 2021; Ray et al. 2021). Existing studies have investigated the helpfulness of reviews based on informational cues and normative cues in online reviews across various fields, such as e-commerce (Kuan et al. 2015; Mudambi and Schuff 2010), hospitality (Zhao et al. 2019; Yang et al. 2021) and tourism management (Ray et al. 2021). However, the extant literature has yielded mixed findings on the influence of informational cues and normative cues in online reviews on *RH* due to the limited knowledge of a deeper understanding of the underlying mechanisms (Meek et al. 2021; Choi and Leon, 2020; Filieri et al., 2018; Chou, Chuang and Liang, 2022). Additionally, a comprehensive study encompassing informational cues and normative cues of online drug treatment reviews has not yet been conducted, resulting in the determinants of online drug treatment review helpfulness still being unclear. An extension across the research context is also essential because patients must be treated with drugs to alleviate their condition, the purpose of which is largely different from that of other customers buying everyday products (Park and Hong 2018; Adusumalli et al. 2015).

Patients with chronic diseases are the main group sharing the experience of drug treatment through healthcare platforms because chronic disease management is mostly self-management (Kordzadeh and Warren 2017). Diabetes is one of the most common chronic diseases, and the global prevalence of diabetes has steadily increased over the past 50 years and has reached pandemic levels. The prevalence of both diagnosed and undiagnosed diabetes was estimated at 9.3% (463 million people) in 2019 and is expected to rise to 10.2% (578 million) by 2030 and 10.9% (700 million) by 2045 (Aschner et al. 2021; Saeedi et al. 2019). Accordingly, in this study, we focus on online drug treatment reviews of diabetic patients. And we aim to understand the *RH* from the informational cues and normative cues and topic information of online drug treatment reviews and reconcile inconsistent findings in the literature by answering the following two research questions.

Question 1: How are the informational cues and normative cues of online drug treatment reviews associated with *RH*?

Question 2: What is the influence of dimension-specific topics on *RH*?

To answer the above research questions, we drew on design science research to study *RH* based on online drug treatment reviews of diabetic patients from the WebMD website. Specifically, the dual-process theory guided hypothesis development. And we extracted informational elements (emotions with different degrees of arousal, review length), normative elements (patient satisfaction, perceived effectiveness and ease of use) and topic information from patients' online drug treatment reviews using text mining techniques. Econometric methods were employed in the demonstration and evaluation stages to explore the determinants of *RH* and the moderating influence of review length on the inverted U-shaped relationship. Furthermore, the influence of dimension-specific topics in drug treatment reviews was also investigated. More importantly, we performed the robustness tests to demonstrate the reliability of our findings.

Our study has three major theoretical contributions to the existing literature related to *RH*. First, we reinforce the validity of the dual process theory in the assessment of *RH* in the healthcare field and advance the development of medical social big data in design science research by incorporating structured and unstructured data. Second, we reconcile inconsistent findings on the influences of informational cues and normative cues of online reviews, which provide an understanding of the underlying mechanisms to researchers and insights for future research in the field of healthcare. Finally, we employ Tobit regression, Poisson regression, negative binomial regression and latent Dirichlet allocation method to unpack online drug treatment reviews, which guarantees the robustness of our findings and contributes to the development of integrating texting mining techniques and econometric methods in online healthcare platforms.

Our work also provides valuable practical implications for three types of individuals: diabetic patients seeking drug treatment, pharmaceutical companies offering drug treatment, and platform providers giving communication. First, we found that reviews related to 'family medical history' and 'dosing time' could help potential patients. Platform providers can design new functions and items according to reviews, such as browsing reviews by topic and scoring drugs by topic. Patients can share more drug treatment experiences, and potential patients can directly browse the reviews of interest rather than aimlessly browse in reviews. Further, platform providers can promote more reviews with extreme satisfaction and long length to potential patients rather than neutral and short ones. This can help potential patients find helpful information faster among numerous online reviews and improve platform providers' operational efficiency.

We organize the rest of the paper as follows: In Section 2, we briefly review the related literature on *RH*, respectively. Section 3 presents the research design. In Section 4, we develop a series of hypotheses related to *RH*. In Section 5, we introduce the research design stage including data collection, data preprocessing and feature engineering and text regressions. All results on the text regression and robustness checks are analyzed and evaluated in Section 6. Finally, in Section 7, we discuss the paper's contributions and limitations and suggest directions for future research.

2. Literature Review

Explaining *RH* has always been a hot topic in marketing, information systems, and decision-support systems (Meek et al., 2021; Mudambi and Schuff 2010; Filieri et al., 2018; Chatterjee 2020). Many studies have tried to explain *RH* by text mining based on reviews in different fields (e.g., food reviews, hotel reviews, and e-commerce reviews). Review-related determinants explored thus far include emotions (e.g. Chatterjee 2020; Chen and Farn 2020; Meek et al. 2021), terms of text depth (e.g. Chatterjee 2020; Choi and Leon 2020; Kuan et al. 2015; Mudambi and Schuff 2010), satisfaction (e.g.; Meek et al., 2021; Mudambi and Schuff 2010; Filieri et al., 2018), longevity (e.g. Salehan and Kim 2016; Zhou and Guo 2017). However, there are also contradictory findings of how these determinants affect perceived helpfulness. Table 1 summarizes the prior literature findings and the contents worthy of further investigation of the three determinants related to this study.

Table 1. Summary of Prior Studies Investigating Review Helpfulness.

<table>
<tr><th>Determinant</th><th>Prior Literature and Findings</th><th>Content for Further Investigation</th></tr>
<tr><td>Positive and Negative Emotions</td><td>Baek et al. (2012); Chatterjee (2020); Kuan et al. (2015); Meek et al. (2021); Salehan and Kim (2016); Yin et al. (2014)</td><td rowspan="2">Previous research has only focused on the influence of positive and negative emotions on RH, which is insufficient to capture the influence of the various emotions (Chen and Farn 2020). For this reason, many studies have taken a more fine-grained look at the influences of discrete emotions, including anxiety, anger, hope, etc., on RH. However, the influence of discrete emotions with different degrees of arousal on RH has not been fully investigated, which deserves further exploration.</td></tr>
<tr><td>Discrete Emotions</td><td>Chatterjee (2020); Chen and Farn (2020); Craciun et al. (2020); Forman et al. (2008); Ren and Hong (2019); Chou, Chuang and Liang (2022)</td></tr>
</table>

Review Length	*Positive:* Baek et al. (2012); Choi and Leon (2020); Kuan et al. (2015); Liu and Park (2015); Mudambi and Schuff (2010); Salehan and Kim (2016); Yin et al. (2014); Yue and Zhang (2011) *Negative:* Chatterjee (2020) *Curvilinear:* Li and Huang (2020); Lutz, Prollochs and Neumann (2022)	Review length has been identified as having a significant influence on *RH*, but the way it is influenced remains debated. Previous studies have argued that review length has a linear and significant influence on *RH*, which includes both positive and negative influences. Additionally, Li and Huang (2020) pointed out that the influence of review length on *RH* was potentially curvilinear and verified that it. Thus, the influence of review length on *RH* deserves further investigation.
Customer Satisfaction	*Positive:* Chatterjee (2020); Liu and Park (2015); Yue and Zhang (2011) *Negative:* Choi and Leon (2020); Zhang et al. (2010) *Curvilinear:* Baek et al. (2012); Chua and Banerjee (2015); Filieri et al. (2018); Liu and Park (2015); Meek et al. (2021); Mudambi and Schuff (2010)	Satisfaction has been identified as having a significant influence on *RH*, but the way it is influenced also remains several arguments. Previous studies suggested that patient satisfaction has a linear and significant influence on *RH*, which includes both positive and negative influences. Additionally, several studies pointed out that there is a curvilinear relationship between patient satisfaction and *RH*. the influence of review length on *RH* deserves further investigation. It is worth exploring how patient satisfaction influences the helpfulness of online drug reviews in the healthcare context.

Table 2 presents the research gaps between our study and the last four years of studies related to *RH*. First, it can be observed from Table 2 that most studies focused on *RH* of online reviews for hotels, restaurants, and products, and only Chou, Chuang, and Liang (2022) investigated *RH* in the context of patients' drug treatment reviews. Although the fields are different, theories related to *RH* can be applied to check the helpfulness of patient drug treatment reviews. Second, when discussing emotional expressions in online reviews, prior studies focused on valence (positive valence and negative valence), and arousal was under-studied. However, the influence of emotions with different arousal on *RH* may be different. Additionally, the researchers still hold mixed findings on the influences of satisfaction and review length, and have little knowledge of a deeper understanding on the underlying mechanism of how satisfaction and review length influence *RH*, especially in the healthcare platform.

Furthermore, how the dimension-specific topics in the online reviews influence *RH* has not been examined, as well as robustness checks of key findings, are ignored by most studies, leading to doubts about the reliability of existing findings. Finally, although prior studies widely took a natural log transformation and used linear regression to estimate the coefficients of the model, log-linear models for count data may make a biased estimate (Mallipeddi et al., 2021). And using Tobit regression, Poisson regression and negative binomial regression to estimate the coefficients is a better way. Thus, we conducted this study on the helpfulness of patient reviews using text-mining techniques, filling the above gaps based on the drug treatment online reviews in the healthcare field.

Table 2. Research Gaps Between Our Study and the Last Four Years of Studies.

Research	Context	Emotions	Review Length	Satisfaction	Review Topic	Robust Check	Model
Chou, Chuang, and Liang (2022)	Online product reviews, drug treatment reviews	Emotions with different arousals	√ (+)	×	×	×	Linear regression
Meek (2021)	Online restaurant reviews	Positive and negative emotions	×	√ (U)	×	×	Linear regression
Chatterjee (2020)	Online hotel reviews	Emotions with different arousals	√ (+)	√ (+)	×	×	Poisson regression, Negative binomial regression
Craciun, Zhou, and Shan (2020)	Online product reviews	Discrete emotions	×	×	×	×	Linear regression, logit model
Choi and Leon (2020)	Online product reviews	×	√ (+)	√ (-)	×	×	Tobit regression and linear regression
Li and Huang (2020)	Online product reviews	×	√ (inverted-U)	×	×	√	Tobit regression
Srivastava and Kalro (2019)	Online hotel reviews, restaurant reviews	Positive and negative emotions	√ (+)	√ (U)	×	√	Tobit regression
Ren and Hong (2019)	Online product reviews	Negative emotion	×	×	×	×	Tobit regression
Wang et al. (2019)	Online restaurant reviews	Discrete emotions	×	√ (-)	×	×	Negative binomial regression
Filieri et al. (2018)	Online hotel reviews	×	√ (inverted-U)	√ (U)	×	×	Tobit regression
Our study	Online drug treatment review	Emotions with different arousals	√ (inverted-U)	√ (U)	√	√	Tobit regression, Poisson regression and negative binomial regression

3. Research Design

Design science research is a complement to the natural behavioral science method dominating information systems research, which has considerable potential to address the relevance and rigor gaps in information systems research (Arnott and Pervan 2012). The advantage of design science research is that its product can manage important information-related tasks by creating innovations that define the problem, practices, and technical capabilities, verified in research in various fields. Prescriptive knowledge, such as models and concepts, can be generated through decision science research. Thus, we use a nominal process model for design science research to guide the entire study. Specifically, the design science framework includes seven stages (Hevner et al. 2004): problem identification, objective definition, design, development, demonstration, evaluation, and communication. Figure 1 maps these steps in our study to the various stages of the process.

Two main questions were identified in the problem identification stage, as discussed in Section 1.

Next, the four objectives were determined based on problem identification, including verifying the hypotheses proposed in Section 4 (addressing in Section 6.1), finding the association of reviews with different topics with *RH* (addressing in Section 6.2), checking the robustness of findings (addressing in Section 6.3), as well as providing practical guidance for diabetes, platform providers, and pharmaceutical companies (presenting in Sections 7.1 and 7.2). The design stage performed three tasks: data collection, data pre-processing and feature engineering, and text regression. Specifically, data related to online drug treatment reviews for patients with diabetes were retrieved from WebMD, and data pre-processing was performed for the initial data. Based on the processed data, features were extracted from the pre-processed data using the NRC word-emotion association lexicon and the latent Dirichlet allocation method (LDA), which were used to guide subsequent regression construction. First, Tobit regressions were developed to analyze the influences of informational elements (emotions with different degrees of arousal, review length) and normative elements (perceived effectiveness, perceived ease of use, and patient satisfaction) on *RH*, and the moderating influence of review length on the relationship between patient satisfaction (*PS*) and *RH*. Subsequently, the extracted topics also were used to construct topic text regressions, which help us identify the association between different topics and *RH*.

In the demonstration and evaluation stages, we constructed models designed in the design stage, and analyzed the results of the models. Based on the evaluations, we verified the hypotheses proposed in Section 4 and interpreted the results obtained. Robustness checks were also performed to verify the reliability and consistency of the findings. Finally, the study findings were discussed based on all the results and existing studies, and the theoretical and practical contributions of our study are concluded in the communication stage.

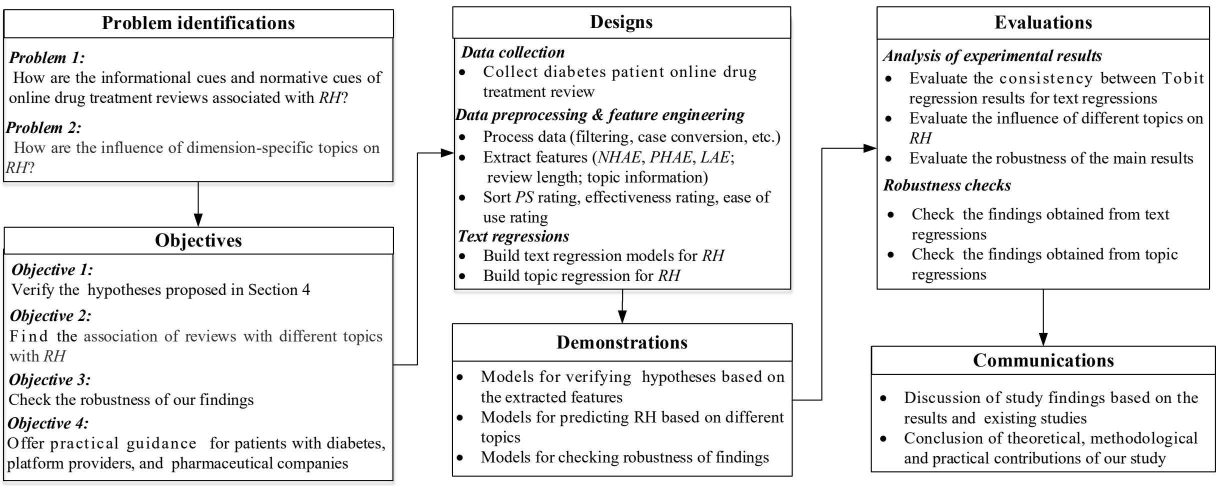


Figure 1. The Design Science Framework of Our Study.

4. Research Proposition Development

Dual-process theory, as the theoretical foundation, derives our research proposition development in the online review setting. The dual-process theory (DPT) proposed by Deutsch and Gerard (1955) is a psychological theory based on the recipient’s evaluation of the received information, where the received information is an essential source of influence, including content, recipient, source, etc. (Filieri 2015). The DPT provides strong theoretical support for the exploration of online *RH*. Specifically, DPT suggests that the characteristics of online consumer reviews can be divided into normative elements related to the context of the review and informational elements associated with the content of the review, which together influence the information assessment of potential readers and thus affects the persuasiveness of online reviews (Filieri 2015; Meek et al. 2021). Examples of informational elements in *RH* include the discussion content within the reviews (Filieri 2015; Meek et al. 2021), such as the emotions expressed by the reviewers in the reviews and the length of the reviews written by the reviewers (Meek et al. 2021). Examples of normative elements are the numerical evaluations of products rated by customers (Filieri 2015; Meek et al. 2021). Guiding by DPT, we can analyze how the informational and normative elements of drug treatment reviews can influence the persuasiveness or perceived helpfulness of the review, leading to a vote of helpfulness for potential patients.

Accordingly, this study aims to investigate the influence of informational and normative elements on the *RH* of drug treatment based on DPT. The different degrees of arousal and depth of content expressed by patients in the reviews reflect the informational cues of multidimensional textual content. Numerical assessments, including *PS* star ratings, effectiveness star ratings, and ease of use star ratings, reflect normative cues of multidimensional text content in reviews. In summary, drawing on DPT, the research model of this study on *RH* is designed, as shown in Figure 2, and all the hypotheses developed in the model are elaborated below.

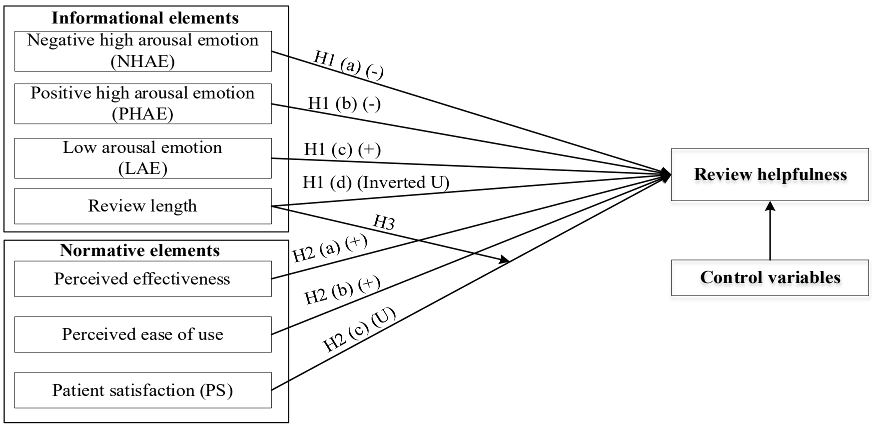


Figure 2. Theoretical Model of Review Helpfulness.

4.1 Informational Influences

The influence of emotions expressed by patients on RH: The emotions expressed by patients are multidimensional, including happiness, trust, anger, and sadness, and emotions with different valences and arousals generate other influences on the perception of potential patients. The study introduces one important aspect of emotions: arousal levels. Cavanaugh et al. (2016) divided emotions into different dimensions with different degrees of arousal, including negative high-arousal emotions (*NHAE*), positive high-arousal emotions (*PHAE*), and low-arousal emotions (*LAE*).

In terms of *NHAE*, anger and fear are categorized as *NHAE*, and *NHAE* in a review reflects a patient's extreme negative feelings during this experience. Potential readers are generally suspicious of online reviews, primarily in terms of their authenticity (Chatterjee 2020). Although the accessibility-diagnosticity model indicates that negative emotions lead to more diagnosticity (Hong et al. 2016), the outcomes of negative emotions differ substantially across the degree of arousal (Salehan and Kim 2016). A review with *NHAE* may often lead to less information available, due to the cognitive ability of a potential reader faces hindrance under high arousal (Filieri, 2016; Salehan and Kim, 2016). Thus, we propose the following hypothesis:

H1 (a): *Negative high-arousal emotions expressed in the review negatively influence review helpfulness.*

PHAE is generally measured by the trust and joy of patients. (Cavanaugh et al. 2016; Ray et al., 2021). Despite the prior research stated that *PHAE* increased the motivation of humans, potential readers are more likely to doubt that such reviews with *PHAE* are paid reviews and fake reviews that exaggerate the effectiveness of a product, and thus distrust them (Chatterjee 2020; Ray et al., 2021). Furthermore, the reviews of positive emotions with high-arousal are less credible and informative because human cognitive abilities can be misjudged under high-arousal emotions, thereby negatively affecting *RH* (Filieri 2016; Salehan and Kim 2016). Thus, we propose the following hypothesis:

H1 (b): *Positive high-arousal emotions expressed in the review negatively influence review helpfulness.*

Compared to high-arousal emotions, *LAE* is harder to access because it is difficult to overcome other cognitive processes (Salehan and Kim 2016). *LAE* reflects subtle, plausible emotional changes (Chatterjee 2020; Ray et al. 2021), which is measured by feelings of sadness and anticipation (Cavanaugh et al. 2016). *LAE* is considered neutral and takes time to awaken (Chatterjee 2020; Ray et al. 2021). Thus, vivid and slightly negative emotions increase the diagnosticity and credibility, and these emotions are defined as *LAE*. Paid reviews are less likely to use such emotions and are more inclined to use high-arousal emotions (Filieri 2016; Salehan and Kim 2016). Accordingly, the *LAE* expressed in this review can be considered more credible, and we propose the following hypothesis:

H1 (c): *The low-arousal emotions expressed in the review positively influence review helpfulness.*

The influence of review length written by patients on RH: Review length is generally measured as the number of characters or words contained in a review, which has been widely verified to correlate positively with *RH* (Choi and Leon 2020; Chatterjee 2020; Kuan et al. 2015; Mudambi and Schuff 2010). Two arguments are the basis of this assumption. First, longer reviews take up more screen space and are more visually prominent. Hence, they are less likely to be ignored than shorter reviews (Kuan et al. 2015). Second, compared with more concise reviews, longer reviews can reduce product-related uncertainty because they contain more product-related information, while shorter reviews are expected to be superficial (Chatterjee 2020; Choi and Leon 2020; Mudambi and Schuff 2010). Additionally, longer reviews are more persuasive (Schwenk 1986), which, in turn, leads to increased levels of consumer confidence, and potential customers can now make decisions based on these reviews (Tversky and Kahneman 1978).

However, information processing theory stated that information use may be limited when the amount of available information is scarce or abundant—information overload may become an issue at specific review lengths (Fink et al. 2018). Information overload is the finite limit on a person's ability to process information and the fact that performance deteriorates once these limits are exceeded (Jacoby 1977). Long product reviews mean that potential consumers have to spend numerous information processing resources to process product review content (Li and Huang 2020). Similarly, previous research has found that potential customers are more likely to feel 'dissatisfied', 'frustrated', and 'agitated' when faced with a difficult task that requires numerous information processing resources (Li and Huang 2020), which lead to a negative influence on the task. Accordingly, the length of an excessively long product review triggers a negative emotional disposition in the potential consumer's reading process, which subsequently inhibits the possibility of a thorough evaluation of the product review content. This can cause potential customers to become reluctant to read the content or abandon the task immediately (Malhotra 1982), ultimately negatively influencing the evaluation of *RH*. We argue that a nonlinear relationship exists between review length and *RH*. Short product reviews provide less information, resulting in potential consumers not being able to reduce their uncertainty about the product from the explanation of the review content to support an assessment of the *RH*. Excessively long reviews can have a negative influence on potential consumers in the assessment of *RH*, considering people's limited information processing capabilities and emotional, and cognitive decisions. Thus, we develop a hypothesis that posits a reverted U-shaped relationship between review length and *RH*.

H1 (d): *Review length has an inverted U-shaped relationship with review helpfulness.*

4.2 Normative Influences

The influence of perceived effectiveness and ease of use on RH: Humans' decisions and behaviors result from their perceptions and emotions (Zeithaml 1988). Potential patients with the same conditions can derive helpful information from perceived effectiveness and ease of use to infer the quality of the drug. Thus, in addition to emotions, a patient's perception of the drug also influences the judgment of potential patients. Effectiveness and ease of use of the drug are two important assessment indicators. The effectiveness of the drug refers to whether the drug effectively relieves the patient's condition and pain or meets the patient's needs and expectations (Lu et al. 2021). Ease of use of the drug refers to whether the patient feels easy and comfortable when using the drug (Park and Hong 2018). Potential patients tend to use filters in terms of various quantitative ratings to reduce their search costs and evaluate alternatives (Olbrich et al. 2012). As perceived effectiveness and ease-of-use ratings increase, reviews become more relevant to potential patient' information from reviews. In turn, they use information from such reviews to help them make decisions (Filieri 2015). Thus, we posit the following hypothesis:

H2 (a): *Compared with reviews with low perceived effectiveness ratings, reviews with high perceived effectiveness ratings would be more helpful for potential patients.*

H2 (b): *Compared with reviews with low perceived ease-of-use ratings, reviews with high perceived ease-of-use ratings would be more helpful for potential patients.*

The influence of patient satisfaction on RH: Satisfaction is usually presented in the form of a comprehensive score in the five-star rating system, which is a factor for readers to judge the quality of the potential drug intuitively without further cognitive effort. This means that it may influence perceived helpfulness for potential patients—whether potential patients judge whether the reviews are helpful. On the one hand, existing studies stated that when a product has high satisfaction, consumers pay less attention to individual reviews (Choi and Leon 2020; Mudambi and Schuff 2010). Reviews with high satisfaction among the numerous reviews can be challenging to trust. These reviews make people believe that organizations or other institutions posted them (Choi and Leon 2020; Mudambi and Schuff 2010). Therefore, reviews with lower satisfaction are more likely to be reasonable and credible for potential patients than those with higher satisfaction. Studies have found that low-satisfaction reviews are effective in making them more helpful. For example, people tend to rely on low-satisfaction reviews because they have a higher diagnostic value and enhanced depth (Choi and Leon 2020; Mudambi and Schuff 2010).

While existing studies have also pointed out that low-satisfaction reviews are more helpful to readers than high-satisfaction reviews, Yue and Zhang (2011) also observed a positive bias in the relationship between user satisfaction and *RH*—the main effect of satisfaction on *RH* is positive or linearly increasing. However, contrary to the findings of Yue and Zhang, a meta-analysis of review helpfulness suggested that this influence was weakly positive and not significant (Pumawirawan et al. 2015). These findings have driven scholars to consider both positive and negative biases by focusing on the extreme nature of satisfaction. The relationship between the satisfaction state and *RH* is V-shaped or U-shaped based on the conformity hypothesis. Both reviews with low and high satisfaction were more helpful than those with moderate satisfaction (Forman et al. 2008; Pavlou and Dimoka 2006). Thus, it is reasonable to assume that there is a U-shaped relationship between patient satisfaction (*PS*) and *RH*, and we hypothesize the following:

H2 (c): *Patient satisfaction has a U-shaped relationship with review helpfulness.*

The moderating role of review length in the influence of PS on RH: There is relatively little research on the moderating influence of review length on the relationship between *PS* and *RH*. The review length could be particularly relevant for the judgment of potential patients. This is because a patient provides a more multifaceted evaluation of the effects of their treatment in terms of various features and may write a longer review to convince the potential patient that their extreme assessment is accurate and credible

(Filieri et al. 2018). For the potential patient, extreme satisfaction (i.e. reviews with low *PS* or high *PS*) with fewer explanations in reviews is more likely to be considered a feature of fake reviews (Filieri 2016). They will take these reviews with a pinch of salt. Subsequently, we hypothesize that longer reviews may increase the plausibility and credibility of a review compared to shorter reviews and justify extremely positive or negative evaluations. In such cases, we expect review length to moderate the relationship between *PS* and *RH* and propose the following hypothesis:

H3: *Review length has a moderating influence on the inverted U-shaped relationship between patient satisfaction and review helpfulness. Longer reviews with extreme satisfaction will be voted as more helpful than shorter reviews with extreme satisfaction.*

5. Design Stage

Design stage aims to design the research program to address the proposed questions and the developed objectives, which consists of three steps in our study. We first seek to collect a dataset of online drug treatment reviews from online drug platforms. After acquiring the online drug review dataset, we use text mining techniques to pre-process the collected reviews, and extract features drawing on the NRC word-emotion association lexicon and the latent Dirichlet allocation method (LDA) based on the research proposition development. Finally, we construct text regression using the extracted features to evaluate the influence of information elements and normative elements on *RH*, and the influence of topics on *RH*. The details of each step are described as follows.

5.1 Data Collection

This study collected online drug treatment reviews for diabetic patients from the well-known professional health portal WebMD (www.webmd.com). WebMD accumulates numerous first-hand users' drug ratings and reviews, but the private information of all patients who post reviews on this website is confidential. Thus, WebMD achieves reliable protection of patients' sensitive private information, and several studies have carried out investigations based on the WebMD dataset (Hautala et al. 2021; Lamano et al. 2021). In our study, we obtained the drug reviews dataset of WebMD from the world's largest data science community (Kaggle website: www.kaggle.com), which covered drug treatment reviews posted on WebMD by patients with different diseases from 2008 to 2018.

Subsequently, we pre-processed data according to the following three principles. First, the patient's disease type should be diabetes, and diseases not related to diabetes were not collected in this study. Second, the reviewer type of the collected data should be patient, and the other types (caregiver, none) were not included. For each sample, information on patient sex, age, condition, perceived effectiveness, perceived ease of use, description of the side effects of the drug, *PS*, *RH*, and review needs to be completed, and samples with missing values were not collected. Third, the time span of reviews is from 2008 to 2018.

Based on the above principles, a sample of 4995 can be obtained, which includes structured data (sex, age, effectiveness, ease of use, *PS*, *RH*) and unstructured data (textual data). Subsequently, emotions, and topics are extracted from textual data through the feature engineering process. These extracted features and features from structured data have emerged to guide subsequent empirical analyses. Figure 3 displays the collected data, where the red boxes denote the collected structured data, and the blue boxes are the unstructured data. The descriptions of drugs with and without side effects are shown in Figure 3(b) for different drugs.

(b)

Figure 3. Explanation of Collected Data from WebMD.

5.2 Data Pre-processing and Feature Engineering

Textual data include excessive valuable information worth exploring and extracting, and it needs to be pre-processed before features can be removed. First, not all drugs have a description of their side effects. For patients who have taken the drug that contains the description of side effects, as shown at the top of Figure 3 (b), the variable 'side_effects' equals 1 and vice versa. Second, we exclude reviews with less than ten words for patient reviews, and all letters were changed to lowercase. Next, we remove all trivial words ('the', 'are', etc.)—as trivial words do not provide virtual meanings—words with less than two letters (e.g., 'y', 'x', etc.), and the number and stop-words, blank spaces, punctuations, and drug names, etc. Stemming reduces the words to their 'stems'. A total of 4995 data points is obtained through the above operations, and subsequent feature engineering is conducted based on them.

Review helpfulness: Many studies have described various methods for defining and measuring online reviews' helpfulness (Cao et al. 2011; Kuan et al. 2015; Meek et al. 2021). There are different ways to measure *RH* on various websites. The two most common ways are as follows. First, users were asked the question 'Was this review helpful to you?' to capture the *RH*. Second, the *RH* is measured by getting the number of likes by setting a 'like' and 'dislike' click below the review (Meek et al. 2021). For the reader, regardless of whether the review is positive or negative, a 'like' indicates that they agree with the review and find it helpful. Following Meek et al. (2021), we set the number of likes as *RH* in this study.

Emotions: Currently, the NRC word-emotion association lexicon (also called EmoLex) is widely used in various consumer review analyses (Siering et al. 2018; Gour et al., 2022). In this study, we used EmoLex for emotion assessment, which provides an assessment of eight emotions (anger, anticipation, disgust, fear, joy, sadness, surprise, and trust) and the number of positive and negative sentiment words available in the reviews. Furthermore, referring to Cavanaugh et al. (2016) and Ray et al. (2021), we used the sum of fear and anger to represent *NHAE*, the sum of trust and joy to represent *PHAE*, and the sum of sadness and anticipation to represent *LAE*.

Topic information from the topic model: Latent Dirichlet allocation (LDA), as a probabilistic topic modeling approach, is widely applied to text data in different fields, including customer reviews (Wu and Chang 2020) and electronic medical record reviews (Baechle et al. 2020). It can help researchers identify hidden semantic structures and hidden topic information in large-scale collected corpora (Wang and Xu 2018), allowing researchers to understand the reviewer's voice better. The main idea of LDA is to characterize documents as random mixtures of latent topics, where a distribution of words classifies each topic (Wu and Chang, 2020; Wang and Xu, 2018). In this study, we attempted to utilize LDA to extract potential topic information from reviews of diabetic patients to guide subsequent exploration of topics influencing *RH*. Using all words for constructing LDA models would cause a '*curse of dimensionality*'. Therefore, nouns are generally considered in constructing LDA models (Wang and Xu 2018). Thus, before constructing the LDA model, we conducted part-of-speech tagging for reviews, and only the extracted nouns were employed for LDA model construction. We treated each review as a word frequency vector and used the bag-of-words method to transform text data into numerical data compatible with data-mining algorithms. Finally, joint distribution was used to calculate the conditional distribution of the topics for a given set of words. After constructing the LDA model, we can obtain the different topics and the probability pro_i^v that reviews $i, i = 1,2, \ldots, n$ belong to topic $v, v = 1,2, \ldots, m$, where

$pro_i^1 + pro_i^2 + \cdots + pro_i^m = 1$.

To determine the optimal number of topics, we used perplexity. Perplexity measures the predictive performance of the probability distribution, and appropriate probability distributions have relatively low perplexity. It is one of the more popular methods for determining the number of topics (Wang and Xu 2018). We set the appropriate number of topics to six based on this method. The selected topics included blood sugar, family medical history, dosing time, injection, body changes, gastrointestinal symptoms, and sleep. Table 3 lists representative words for each topic. First, Topic 1 contains a bag of words about blood sugar in diabetes, including blood sugar levels, blood sugar control, and how blood sugar acts after diet. Second, Topic 2 focuses on describing the family's medical history, mainly including whether the family has hereditary diabetes and the impact on the life of the patient and family members. Topic 3 describes the length of time the patient has been taking the drug, including when it is taken, the number of doses, and how long it is continuously taken. Topic 4 focuses on how the drug is injected, including the injection site and the reactions effected by the injection, such as itching, pain, and redness. Additionally, body changes were mainly distributed in words derived from Topic 5, including nervousness, weight change, and hair loss. Finally, Topic 6 is about the side effects on the gastrointestinal tract and sleep that patient experience after taking the drug, including constipation, stomach ache, headache, and drowsiness.

Table 3. The Representative Topic Words of Six Topics.

Topic	Representative Words
Topic 1: Blood sugar	sugar, blood, level, control, up, down, drop, lbs (low blood sugar), keep, well, severe, bad
Topic 2: Family medical history	family, history, husband, mom, genetic, high, life, eat, grandpa, young, age, childhood, son
Topic 3: Dosing time	day, week, time, hour, morning, night, work, start, dose, help, know, thing, am, dinner, eat
Topic 4: Injection	injection, site, side, itching, reaction, pain, redness, pain, dosage, mcg, swell, problem, trouble
Topic 5: Body changes	nausea, body, nervousness, weight, hair loss, eye, skin, taste, heart, muscle, leg, dried, pound
Topic 6: Gastrointestinal and sleep	stomach, stomachache, diarrhea, headache, drowsiness, constipation, vomit, insomnia, appetite

After we transformed the unstructured data (textual data) to structured data using the EmoLex and LDA topic models, we merged them with the initial structured data collected from WebMD. Table 4 presents all variables and their statistical information and definitions after completing feature engineering.

5.3 Text Regressions

In the last task of the design stage, we conducted text regressions on *RH* to verify the developed propositions. We employed Tobit regression to construct a text regression model of *RH* due to the specific features of *RH* and the censored nature of the sample. In conducting the text regression of *RH*, the dependent variable is the number of helpfulness votes, the independent variables include *NHAE*, *PHAE*, *LAE*, review length, perceived effectiveness, perceived ease of use, and *PS*, and the moderate variable is review length. Furthermore, patient characteristics also may influence the *RH*, thus patient characteristics are controlled for model construction to exclude interference in the results (Adusumalli et al. 2015). Hence, we control for patients' sex, age, side effects and longevity of patient reviews when conducting the text regressions. Finally, the five text regression models are developed as follows.

Model 1: $Review\ Helpfulness_i = \beta_0 + \beta_1 Sex_i + \beta_2 Age_i + \beta_3 Side\_effects_i + \beta_4 Longevity_i + \varepsilon_i$,

Model 2: $Review\ Helpfulness_i = \beta_0 + \beta_1 Sex_i + \beta_2 Age_i + \beta_3 Side\_effects_i + \beta_4 Longevity_i + \beta_5 NHAE_i + \beta_6 PHAE_i + \beta_7 LAE_i + \beta_8 Review\ length_i + \beta_9 Review\ length_i^2 + \varepsilon_i$,

Model 3:$Review\ Helpfulness_i = \beta_0 + \beta_1 Sex_i + \beta_2 Age_i + \beta_3 Side\_effects_i + \beta_4 Longevity_i + \beta_5 Effectivenss_i + \beta_6 Ease\ of\ use_i + \beta_7 PS_i + \beta_8 PS_i^2 + \varepsilon_i$,

Model 4: $Review\ Helpfulness_i = \beta_0 + \beta_1 Sex_i + \beta_2 Age_i + \beta_3 Side\_effects_i + \beta_4 Longevity_i + \beta_5 NHAE_i + \beta_6 PHAE_i + \beta_7 LAE_i + \beta_8 Review\ length_i + \beta_9 Review\ length_i^2 + \beta_{10} Effectivenss_i + \beta_{11} Ease\ of\ use_i + \beta_{12} PS_i + \beta_{13} PS_i^2 + \varepsilon_i$,

Model 5: $Review\ Helpfulness_i = \beta_0 + \beta_1 Sex_i + \beta_2 Age_i + \beta_3 Side\_effects_i + \beta_4 Longevity_i + \beta_5 NHAE_i + \beta_6 PHAE_i + \beta_7 LAE_i + \beta_8 Effectivenss_i + \beta_9 Ease\ of\ use_i + \beta_{10} PS_i + \beta_{11} PS_i^2 + \beta_{12} Review\ length_i + \beta_{13} Review\ length_i \times PS_i + \beta_{14} Review\ length_i \times PS_i^2 + \varepsilon_i$.

where model 1 includes only control variables, model 2 provides control variables and informational elements, model 3 includes control variables and normative elements, and model 4 contains control variables, informational and normative elements. In model 2, the valance and significance level of β_8 indicate whether there is a curvilinear relationship (and the shape, if it exists) between review length and *RH*, and the inflection point can be calculated through the estimations of β_8 and β_9. In model 3, the valance and significance level of β_7 indicate whether there is a curvilinear relationship (and the shape, if it exists) between review length and *RH*, and the inflection point can be calculated through the values of β_7 and β_8. Model 5 is designed to validate the moderating role of review length in the relationship between *PS* and *RH*.

A particular topic may dominate in each review corpus, and the topic regressions of *RH* are also constructed using Tobit regression to analyze the influence of different topics on *RH*. The regression model is listed as follows.

Model 6: $Review\ Helpfulness_i = \beta_0 + \beta_1 Sex_i + \beta_2 Age_i + \beta_3 Side\_effects_i + \beta_4 Longevity_i + \beta_5 Blood\ sugar_i + \beta_6 Family\ medical\ history_i + \beta_7 Dosing\ time_i + \beta_8 Injection_i + \beta_9 Body\ changes_i + \varepsilon_i$.

Table 4. Statistical Information and Definitions of All Features/Variables.

	Variable	Min	Max	Mean	Standard Deviation	Definition
Control Variables	Sex_i	0	1	0.4	0.49	Gender of the patient who writes review i
	Age_i	1	8	5.65	1.244	Age group of the patient who writes review i
	$Side\_effects_i$	0	1	0.89	0.307	Description of whether the medication taken by the patient who wrote review i Judgement of side effects
	$Longevity_i$	13	4548	3159	1008	The number of days since a review posted
Independent Variables (Informational Elements)	$Negative\ high\ arousal\ emotions_i (NHAE_i)$	0	21	2.83	2.56	Anger and fear of the review i
	$Positive\ high\ arousal\ emotions_i\ (PHAE_i)$	0	26	2.37	2.79	Trust and joy of the review i
	$Low\ arousal\ emotions_i\ (LAE_i)$	0	37	3.67	3.22	Sadness and anticipation of the review i
	$Review\ length_i$	10	2001	314.8	230.05	Total words of the review i
Independent Variables (Normative Elements)	$Effectiveness_i$	1	5	3.4	1.451	Effectiveness rating of the patient who writes review i
	$Ease\ of\ use_i$	1	5	4.04	1.257	Ease of use rating of the patient who writes review i
	$Patient\ satisfaction_i\ (PS_i)$	1	5	2.97	1.611	Satisfaction rating of the patient who writes review i
Independent Variables (Topics)	$Blood\ sugar_i$	0.016	0.717	0.158	0.098	The probability of the topic 'blood sugar' extracted from the review i
	$Family\ medical\ history_i$	0.027	0.609	0.035	0.045	The probability of the topic 'family medical history' extracted from the review i
	$Dosing\ time_i$	0.037	0.781	0.195	0.010	The probability of the topic 'dosing time' extracted from the review i
	$Injection_i$	0.055	0.668	0.087	0.111	The probability of the topic 'injection' extracted from the review i
	$Body\ changes_i$	0.036	0.753	0.284	0.139	The probability of the topic 'body changes' extracted from the review i
	$Gastrointestinal\ and\ sleep_i$	0.018	0.672	0.242	0.164	The probability of the topic 'gastrointestinal and sleep' extracted from the review i
Dependent Variable	$Review\ Helpfulness_i$	0	127	9.72	10.019	Helpful vote count of review i

Table 5. The Results of Text Regression on RH. (Noted: Robust standard errors in parentheses, $\beta(Standard\ error)$; $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.)

		Dependent Variable: Review Helpfulness				
		Model 1	Model 2	Model 3	Model 4	Model 5
Control Variables	*Sex*	-1.085*** (0.311)	-0.526* (0.303)	-0.428* (0.305)	-0.363 (0.278)	-0.611** (0.296)
	Age	-0.407*** (0.121)	0.016 (0.118)	-0.457*** (0.119)	-0.063 (0.111)	-0.021 (0.117)
	Side_effects	-3.104*** (0.400)	-4.546*** (0.385)	-2.976*** (0.395)	-3.762*** (0.451)	-4.098*** (0.376)
	Longevity	0.001*** (0.0002)	0.001*** (0.0002)	0.001*** (0.0002)	0.001*** (0.000)	0.001*** (0.0002)
Independent Variables (Informational Elements)	*Negative high arousal emotions (NHAE)*		-0.182 (0.121)		-0.157 (0.097)	-0.215** (0.121)
	Positive high arousal emotions (PHAE)		0.160 (0.106)		0.111 (0.086)	0.140 (0.106)
	Low arousal emotions (LAE)		0.292** (0.134)		0.285*** (0.102)	0.310** (0.134)
	Review length		0.021*** (0.002)		0.018*** (0.002)	0.006*** (0.002)
	Review length^2		-8.89e-06*** (1.67e-06)		-7.59e-06*** (1.10e-06)	
Independent Variable (Normative Elements)	*Effectiveness*			1.131*** (0.176)	0.647*** (0.160)	0.774*** (0.173)
	Ease of use			0.177 (0.147)	0.077 (0.128)	0.185 (0.140)
	Patient satisfaction (PS)			-6.566*** (0.513)	-5.346*** (0.501)	-6.737*** (0.538)
	Patient satisfaction^2 *(PS*^{2}*)*			0.923*** (0.083)	0.802*** (0.079)	0.917*** (0.081)
Interaction Terms	*Patient satisfaction×Review length*					0.003*** (0.001)
	Patient satisfaction^2*×Review length*					-3.56e-07*** (1.03e-07)
	Constant	12.453*** (0.992)	4.525*** (1.097)	16.078*** (1.289)	10.164*** (1.178)	11.508*** (1.326)

6. Demonstration and Evaluation—Results

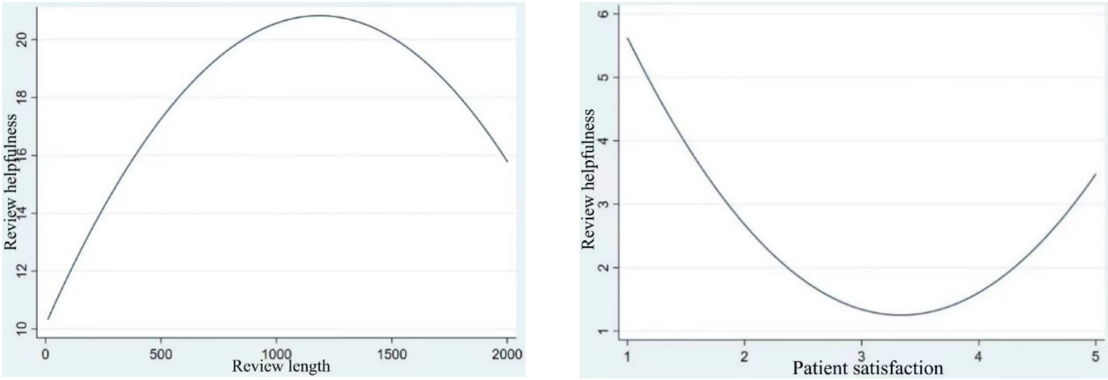
Demonstration and evaluation stage aims to analyze the results of text regression models and verify the propositions developed in Section 4, which includes three contents in our study. We first obtain the parameter estimation of information elements and normative elements, analyze whether the hypotheses established are supported, and explain the logic behind them. We then capture topics that have a significant influence on *RH* by evaluating the parameter estimation of each topic. After completing these two contents, the findings extracted from the models are summarized, but the reliability of the findings obtained from the model still needs to be further checked. Hence, we further carry out a robustness check to verify whether our findings are robust and reliable.

6.1 Results of Text Regressions

Before constructing the regression models, we examined the value of variance inflation factors (VIFs) in the models, which could help us understand any cointegration problem between the variables. The VIF values of all variables were far below the typical benchmark value of 10—there was no multicollinearity problem (Zhao et al., 2019). Thus, we performed text regressions according to the models introduced in the design stage, and Table 5 reports the results of models proposed in Subsection 5.3 using Tobit regression. Next, we analyzed and verified the hypotheses developed in Section 4.

First, we analyzed the influence of informational elements on *RH*. Through the analysis of the results in Table 5, the coefficient of *NHAE* is only significantly negative in model 5. Although it also has a negative influence on *RH* in models 2 and 4, these influences are not significant at the 5% level; thus, H1(a) is only partially supported. Further, *PHAE* does not have a significant influence on *RH* at the 5% level in models 2, 4, and 5; therefore, H1(b) is not supported. This result is surprising, and one possible explanation is that our study's specificity of products and consumer groups differs from some of the findings explored in existing studies on common products (Hong et al. 2016; Salehan and Kim 2016). H1(c) states that the *LAE* expressed by patients is more likely to be considered helpful. As Table 5 shows, the coefficients of *LAE* in models 2,4,5 are positive and significant, indicating that they are positively associated with perceived helpfulness for potential patients, consistent with H1(c).

Additionally, models 2 and 4 show that the coefficient of *review length*^{2} is negative and significant, and the coefficient of *review length* is also significant. Therefore, we calculate the inflection point of the curve between *review length* and *RH* according to the coefficient of *review length* and *review length*^{2}. Specifically, the inflection points in model 2 can be calculated by substituting the estimated coefficients of *review length* and *review length*^{2} into $\frac{\beta_8}{2\times\beta_9}$, and the value is 1181. Similarly, the value of the inflection point in model 4 is 1186 ($\frac{\beta_8}{2\times\beta_9}$). We also display this curvilinear relationship between *review length* and *RH* in Figure 4 (a). Thus, H1(d) is supported.



(a) The Relationship Between Review Length and Review Helpfulness

(b) The Relationship Between Patient Satisfaction and Review Helpfulness

Figure 4. The Influence of Review Length and Patient Satisfaction on Review Helpfulness.

Subsequently, the influence of normative elements on *RH* was analyzed. Table 5 shows that perceived effectiveness and perceived ease of use are positively associated with *RH*. Nevertheless, only the coefficient of perceived effectiveness is significant in models 3–5. Thus, H2(a) is supported, while H2(b) is not supported. For *PS*, model 3 shows that the coefficient of PS^2 (β_8) is positive and significant, and the coefficient of *PS* (β_7) is also significant. Therefore, the inflection point of the curve between *PS* and *RH* can be calculated by substituting the estimated coefficients of *PS* and PS^2 into $-\frac{\beta_7}{2\times\beta_8}$, and the value is 3.56. Similarly, the inflection point of *PS* in model 4 can also be calculated in the same way ($-\frac{\beta_{12}}{2\times\beta_{13}}$) and is 3.33. The curvilinear relationship between *PS* and *RH* is shown in Figure 4 (b). These results support our idea of a U-shaped relationship between *PS* and *RH* such that *RH* decreases as the extent of *PS* increases, thus supporting H2(c).

Finally, we focus on the moderating role of review length in the U-shaped relationship between *PS* and *RH*. As reported in model 5 of Table 5, review length effectively moderates the U-shaped relationship between *PS* and *RH* because their interaction terms are significant (*Patient satisfaction*×*Review length*: $\beta = 0.003, p < 0.01$; *Patient satisfaction*^{2}×*Review length*: $\beta = -3.56\text{e}-07, p < 0.01$). To further explore how the review length moderates this relationship, we plotted an interaction diagram, as shown in Figure 5. Specifically, when the rating of *PS* is from low to moderate, the slope between *PS* and *RH* becomes steeper under a short review length compared to a long review length—it has a weakening effect on the relationship between *PS* and *RH*. Additionally, the review length influences the U-shaped relationship between *PS* and *RH* by the movement of the turning point. As the comment length increased, the turning point of the *PS* moved to the left and upwards. The moderating role of review length on the inverted U-shaped relationship between *PS* and *RH* is verified, thus supporting H3. Table 6 summarizes the experimental results of all the hypotheses proposed in Section 4.

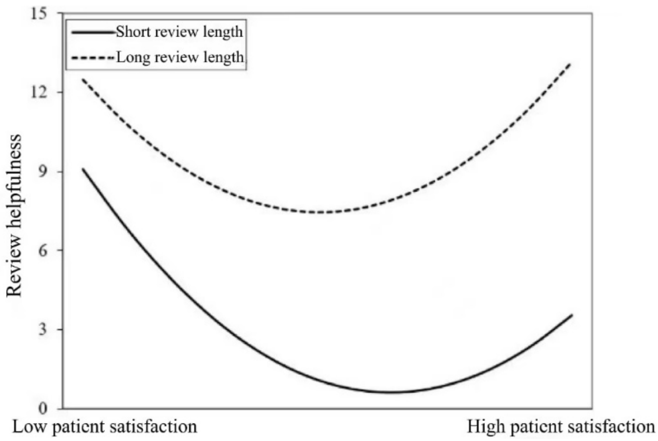


Figure 5. Moderating Effects of Review Length.

Table 6. Summary of Hypotheses Testing.

Hypothesis	Relationship	Results
H1(a)	Negative high arousal emotions→Review helpfulness	Partially supported
H1(b)	Positive high arousal emotions→Review helpfulness	Not supported
H1(c)	Low arousal emotions→Review helpfulness	Supported
H1(d)	Review length^2→Review helpfulness	Supported
H2(a)	Perceived effectiveness →Review helpfulness	Supported
H2(b)	Perceived ease of use →Review helpfulness	Not supported
H2(c)	Patient satisfaction^2→Review helpfulness	Supported
H3	Review length ×Patient satisfaction^2→Review helpfulness	Supported

6.2 Topic Analysis for Review Helpfulness

To capture the association of the topics with *RH*, we constructed regression models for six topics using Tobit regression, where the last topic was removed from the regression models referring to Wen et al. (2019). This is because its probability can be calculated according to the probability of other topics (complete collinearity). Table 7 reports the experimental results, where the second column does not consider the robust estimator of variance, whereas the third column does. For *RH*, it can be observed that ‘blood sugar’, ‘family medical history’, ‘dosing time’, and ‘body changes’ all have a significant positive coefficient. Conversely, only the ‘injection’ has a significant negative regression coefficient. This suggests that the reviews belonging to the topics ‘blood sugar’, ‘family medical history’, ‘dosing time’, and ‘body changes’ would help potential patients. Conversely, ‘injection’ is negatively associated with *RH*. We also report the robustness results for the influence of topics on *RH* in the following section, which is consistent with the above analysis, thus further proving the reliability of our findings.

Table 7. The Results of Topic Regression.

		Dependent Variable: Review Helpfulness		
		Model 1	Model 6	Model 6 (Robust)
Control Variables	*Sex*	-1.085*** (0.312)	-0.970*** (0.303)	-0.970*** (0.297)

	Age	-0.407*** (0.124)	-0.344*** (0.120)	-0.344*** (0.118)
	Side_effects	-3.104*** (0.491)	-0.494 (0.512)	-0.494 (0.446)
	Longevity	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)
Independent Variables (Topics)	*Blood sugar*		11.159*** (1.668)	11.159*** (1.680)
	Family medical history		21.761*** (3.271)	21.761*** (4.332)
	Dosing time		17.199*** (1.596)	17.199*** (1.726)
	Injection		-19.046*** (1.471)	-19.046*** (1.350)
	Body changes		0.273 (1.191)	0.273 (1.196)
	Constant		5.569*** (1.369)	5.569*** (1.340)

Noted: Robust standard errors in parentheses; $\beta(Standard\ error)$; * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

6.3 Robustness Checks

Two robustness tests were performed to validate the hypotheses further. First, we rigorously tested the inverted U-shaped relationship between review length and *RH* and the U-shaped relationship between *PS* and *RH*. Additionally, the u-test was conducted to verify the proposed curvilinear relationships by referring to Lind and Mehlum (2010), and Table 8 summarizes the results, which support our ideas. For the inverted U-shaped relationship between review length and *RH*, *RH* rises as the extent of review length increases, however, *RH* drops when the value of review length is beyond the infection point. The results of the u-test further reveal that the slope is significantly positive (slope: 0.018) on the lower bound and significantly negative (slope: -0.012) on the upper bound. Similarly, *RH* drops with a slope of 3.743 as the extent of *PS* increases; however, *RH* rises with a slope of 2.671 when the value of *PS* is beyond the infection point. Accordingly, the robustness test results strengthen the validity of the findings.

Table 8. The Results of the U-Test.

		Lower Bound	Upper Bound
Review Length	Slope	0.01784	-0.01237
	t-value	11.9953	-3.8208
	p-value	0	0
	Overall test of presence of an inverted U shape: *t*-value =3.82; *p*-value=0.00		
	95% Fieller interval for extreme point: [1022.20,1456.40]		
Patient Satisfaction	Slope	-3.7429	2.6712
	t-value	-10.6125	7.6867
	p-value	0	0
	Overall test of presence of a U shape: *t*-value =7.69; *p*-value=0.00		
	95% Fieller interval for extreme point: [3.1523; 3.5409]		

We also retest and validate the robustness of the hypotheses testing results by applying alternative analytical techniques, including negative binomial regression and Poisson regression. *RH* is positive and integer; therefore, these two regressions are also appropriate. Tables 9-10 report the results of models 4, 5, and 6 under these two regressions of robustness checks respectively. Further, the estimation results, as shown in Tables 9-10, are consistent with our main results and findings. Consequently, we conclude that our findings are robust.

Table 9. The Robustness Check Results of Text Regression.

		Dependent Variable: Review Helpfulness			
		Negative Binomial Regression		Poisson Regression	
		Model 4	Model 5	Model 4	Model 5
Control Variables	*Sex*	-0.042 (0.030)	-0.049 (0.030)	-0.041 (0.029)	-0.055^{*} (0.029)
	Age	-0.003 (0.013)	-0.005 (0.013)	-0.008 (0.011)	-0.010 (0.011)
	Side_effects	-0.347^{***} (0.033)	-0.328^{***} (0.033)	-0.338^{***} (0.032)	-0.324^{***} (0.032)
	Longevity	0.00007^{***} (0.00002)	0.00007^{***} (0.00001)	0.00007^{***} (0.00002)	0.0007^{***} (0.0002)
Independent Variables (Informational Elements)	*Negative high arousal emotions (NHAE)*	-0.020^{**} (0.010)	-0.018^{*} (0.010)	-0.015 (0.009)	-0.013 (0.010)
	Positive high arousal emotions (PHAE)	0.005 (0.008)	0.004 (0.009)	0.008 (0.008)	0.007 (0.008)
	Low arousal emotions (LAE)	0.032^{***} (0.010)	0.030^{***} (0.011)	0.025^{**} (0.010)	0.023^{**} (0.010)
	Review length	0.002^{***} (0.000)	0.0004^{***} (0.0001)	0.002^{***} (0.0002)	0.0003^{**} (0.0001)
	Review length^2	$-9.67e-07^{***}$ (1.19e-07)		$-8.36e-07^{***}$ (1.29e-07)	
Independent Variables (Normative Elements)	*Effectiveness*	0.088^{***} (0.017)	0.090^{***} (0.017)	0.066^{***} (0.016)	0.072^{***} (0.015)
	Ease of use	0.010 (0.014)	0.015 (0.014)	0.009 (0.013)	0.013 (0.013)
	Patient satisfaction (PS)	-0.604^{***} (0.051)	-0.730^{***} (0.054)	-0.570^{***} (0.048)	-0.704^{***} (0.052)
	Patient satisfaction^2 *(PS*^{2}*)*	0.089^{***} (0.008)	0.096^{***} (0.008)	0.086^{***} (0.008)	0.094^{***} (0.008)
Interaction Terms	*Patient satisfaction×Review length*		0.0004^{***} (0.00007)		0.0003^{***} (0.0006)
	Patient satisfaction^2*×Review length*		$-4.58e-08^{***}$ (6.76e-09)		$-3.94e-08^{***}$ (9.04e-09)
	Constant	2.211^{***} (0.139)	2.585^{***} (0.136)	2.291^{***} (0.130)	2.684^{***} (0.125)

Noted: Robust standard errors in parentheses; $\beta(Standard\ error)$; * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 10. The Robustness Checks Results of Topic Regression.

		Dependent Variable: Review Helpfulness	
		Negative Binomial Regression	Poisson Regression
		Model 6	Model 6
Control Variables	*Sex*	-0.093^{***} (0.030)	-0.087^{***} (0.029)
	Age	-0.038^{***} (0.012)	-0.038^{***} (0.011)
	Side_effects	-0.034 (0.040)	-0.013 (0.038)
	Longevity	0.000^{***} (0.000)	0.000^{***} (0.000)
Independent Variables	*Blood sugar*	1.022^{***} (0.155)	1.054^{***} (0.156)

(Topics)	*Family medical history*	1.980*** (0.359)	1.976*** (0.326)
	Dosing time	1.588*** (0.152)	1.567*** (0.148)
	Injection	-2.180*** (0.167)	-1.982*** (0.156)
	Body changes	0.057 (0.122)	0.099 (0.119)
	Constant	1.986*** (0.136)	1.957*** (0.131)

Noted: Robust standard errors in parentheses; $\beta(Standard\ error)$; * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

7. Communication—Discussions, Conclusions, Contributions, and Limitations

7.1 Discussions

Academically, previous studies have examined *RH* in e-commerce (Kuan et al. 2015; Mudambi and Schuff 2010; Meek et al., 2021), hospitality (Zhao et al. 2019; Yang et al. 2021), and tourism fields (Ray et al. 2021), but poor studies have explored the determinants of *RH* in the healthcare field. Patients write online reviews to share their drug treatment experience, these reviews convey extensive information to potential patients, platform providers and product providers, which is worth examining in depth. In our study, using online drug treatment reviews as a carrier, we aimed to unpack the influence of the informational elements and normative elements in online drug treatment reviews on review helpfulness (*RH*) by applying the design science approach and dual-process theory (DPT). We also investigated the influence of dimension-specific topics in online drug treatment reviews on *RH*. These explorations can enhance our understanding of *RH* in healthcare.

First, many studies have pointed out that emotions significantly influence *RH*. Hong et al. (2016) pointed out that reviews that convey negative emotions tend to be more helpful to potential patients. However, they did not investigate the influence of negative emotions with different degrees of arousal on *RH*. Our study found that not all emotions with varying levels of arousal influenced *RH*. Although extreme emotions, whether positive or negative, are generally considered untrustworthy in online e-commerce reviews (Chatterjee 2020), the findings show that only negative high arousal emotions (*NHAE*) have a significant negative influence on *RH*. Conversely, patients' low arousal emotions (*LAE*) in the reviews significantly positively influence *RH*, which means that the reviews containing sadness and anticipation can help potential patients. This result might also imply that *LAE* expresses slight dissatisfaction and more expectation of product improvement than complaints about the drugs. In turn, potential patients perceive the reviewer as exerting more awareness and effort, and our findings are consistent with those of current consumer behavior studies (Chen and Farn 2020; Chatterjee 2020).

Second, most empirical studies on online reviews over the past decade have observed that review length positively influences various performance dimensions, including *RH* and sales (Chevalier and Mayzlin 2006; Choi and Leon 2020; Mudambi and Schuff 2010). A naturally following question is why previous studies have not investigated whether the relationship is curvilinear. Fink et al. (2018) pointed out that longer product reviews do not necessarily lead to higher sales. Our study answers this question from an *RH* perspective and verifies the inverted U-shaped relationship between review length and *RH*. This finding s in line with Filieri et al. (2018) and Li and Huang (2020), which serve as further evidence of the nonlinear and non-monotonous relationship between information availability and decision-making in the context of *RH*.

In the normative elements, perceived effectiveness is supported. It has a significant positive influence on *RH*, while perceived ease of use is not supported. This means that reviews with higher effectiveness ratings can help potential patients make better and faster decisions when they evaluate alternatives following non-compensatory purchase decision rules (Johnson and Meyer 1984). Additionally, prior literature has demonstrated a U-shaped relationship between customer satisfaction and *RH* in other fields (Kalro,2019; Meek, 2021; Filieri et al., 2018). Interestingly, we verify an asymmetric U-shaped relationship between *PS* and *RH*, which provides new evidence of this relationship in healthcare. The reason behind the asymmetric U-shaped relationship between *PS* and *RH* is that reviews with extreme *PS* (overly low satisfaction or overly high satisfaction) have a higher diagnostic value and higher depth compared with those with moderate satisfaction. Accordingly, potential patients could receive more helpful information from low- and high-satisfaction reviews.

Concerning the moderating effects, our study suggests that review length moderates the U-shaped relationship between *PS* and *RH*. We confirmed that a long review with extreme *PS* is more likely to be voted as helpful by patients than a short review accompanied by extreme *PS*. When patient satisfaction increases from low to the inflection point, the negative influence of *PS* on *RH* is weakened with increasing review length. When patient satisfaction increases from inflection point to high, the positive influence of *PS* on *RH* is strengthened with increasing review length. Additionally, a longer review causes a leftward movement of the curve's inflection point between *PS* and *RH*. The principle behind this finding can be explained by the fact that longer reviews may add more arguments that substantiate the patient's extreme satisfaction, leading potential patients to perceive the review as more rational and reasonable, which ultimately affects the persuasiveness and potential patients' perception of helpfulness (Filieri et al. 2018; Filieri 2016). Finally, reviews related to blood sugar, family medical history, dosing time, body changes, and injections can also help potential patients. However, topic injection negatively influences the judgment of potential patients regarding their helpfulness.

7.2 Conclusion and Contribution

Although the important factors influencing *RH* in the context of e-commerce have been well documented (Meek et al. 2021; Choi and Leon, 2020; Filieri et al., 2018; Chou, Chuang and Liang, 2022), *RH* on online drug platforms is understudied, and existing literature only focuses on presenting inconsistent findings on important factors affecting *RH*. As the number of online drug treatment reviews for diabetic patients accumulates, they urgently need to be exploited and mined to provide valuable insights for pharmaceutical companies and healthcare providers. The purpose of our study is to understand the *RH* from the informational cues and normative cues and topic information of online drug treatment reviews and reconcile inconsistent findings in the existing literature.

Based on online drug treatment reviews for diabetic patients, using econometric methods and texting mining techniques as the methodological lens, and design science approach and DPT as the theoretical lens, this study advances our understanding for the determinants of *RH* in online drug treatment reviews. Our findings reveal the influences of informational elements (different arousal degrees of emotions, review length) and normative elements (perceived effectiveness and ease of use, and patient satisfaction) with *RH*. First, not all emotions significantly influence review helpfulness, and only low-arousal emotions have a significant positive influence on *RH*. Second, the empirical findings offer evidence that an inverted U-shaped relationship exists between review length and *RH* and that an asymmetric U-shaped relationship exists between *PS* and *RH*. Third, the moderating role of review length on the inverted U-shaped relationship between patient satisfaction and review helpfulness is verified. Finally, the review topics related to blood sugar, family medical history, dosing time and injection significantly influence *RH*. To the best of our knowledge, this work is the first to systematically investigate *RH* by combining the framework of the design science approach and DPT from multiple perspectives in the healthcare context, and we illustrate the theoretical and practical implications of this study as follows.

7.2.1 Theoretical Implications

This study contributes three major theoretical insights to the prior literature. First, it reinforces the validity of the DPT in the assessment of *RH* and the design science approach in the field of healthcare. Patients write online reviews to express their emotions regarding drug treatment and provide information about their perceptions in terms of effectiveness, ease of use, and satisfaction. These expressions convey information to potential patients. When potential patients read online drug treatment reviews, they apply both informational and normative influences to assist them in making decisions regarding whether the reviews are helpful (Filieri 2015; Deutsch and Gerard, 1955). Existing studies have relied on only the valence of emotions (positive and negative) to investigate the influence of informational elements (i.e. emotions) on *RH* (Meek et al. 2021; Srivastava and Kalro, 2019; Ren and Hong, 2019). Unlike the above research, we further measured the emotions according to the different degrees of arousal and explored the influence of *NHAE*, *PHAE*, and *LAE* on *RH* because emotions with varying degrees of arousal may have a completely different influence on *RH* according to the division of emotions (Cavanaugh et al. 2016). Our findings indicate that only *LAE*, such as sadness and anticipation, help potential patients, and reviews with *NHAE* lead to lower *RH*. These findings are also an extension of the emotion literature in healthcare and a valuable contribution to this research domain.

Second, rich studies have investigated the influence of customer satisfaction and review length on *RH*, and various inconsistent arguments have been put forward (Choi and Leon 2020; Meek et al. 2021;

Mudambi and Schuff 2010; Salehan and Kim 2016). Our study supports the existence of a U-shaped relationship between *PS* and *RH*, which is consistent with the findings of Mudambi and Schuff (2010), Filieri et al. (2018) and Meek et al. (2021) in e-commerce. Furthermore, we re-examined the influence of review length on *RH*. Prior studies have suggested that review length positively influences *RH*, as longer reviews can contain more information and are more diagnostic (Mudambi and Schuff 2010; Salehan and Kim 2016; Kuan et al. 2015; Choi and Leon 2020; Liu and Park 2015). Conversely, our study provides two pieces of evidence of an inverted U-shaped relationship between review length and *RH*. First, longer reviews require higher levels of information processing. People's limited ability to process information beyond the point of optimal cognitive load may reduce the use of information in potential patients' decision-making (Fink et al. 2018; Jacoby 1977). Second, excessively long reviews require potential patients to expend considerable resources, such as time and effort, to understand and evaluate the review content, which can trigger negative emotional dispositions in potential patients and thus negatively influence *RH* (Li and Huang 2020). Additionally, we demonstrate the moderating effect of review length on the U-shaped relationship between *PS* and *RH*. In summary, our study provides a salient contribution to consolidating the findings of the impact of review length and *PS* on *RH* in most prior studies, which can extend the understanding of *RH* of online healthcare platforms to researchers and insights for future research in the field of healthcare.

Finally, we advance the development of integrating texting mining methods and econometric methods in online healthcare platforms, and exact the steps of data analysis with the guidance of the design science approach by exploring this work. Despite there are lots of existing studies that investigate *RH* in the context of E-commerce platforms leveraging texting mining methods and econometric methods (Choi and Leon, 2020; Li and Huang, 2020), causal explanations for the *RH* of online drug treatment reviews on online healthcare platforms using texting mining methods and econometric methods are understudied. We combine the texting mining methods (NRC word-emotion association lexicon and latent Dirichlet allocation method) and econometric methods (Tobit regression, Poisson regression, negative binomial regression) to investigate the *RH* of online drug treatment reviews, which guarantees the robustness of our findings. Moreover, we employ the design science approach to segment the data analysis into seven stages and identify the operations and steps for each stage based on the purpose of each stage in the design science approach (Arnott and Pervan 2012). Hence, our study contributes to the application of texting mining methods and econometric methods in online healthcare platforms, and provides some insights for researchers to segment the steps of data analysis with a design science approach.

7.2.2 Practical Implications

With the increasing effect of digital technology on the way patients communicate with each other, from a practical perspective, the findings of our study primarily serve three groups of individuals:

diabetic patients, pharmaceutical companies, and the platform provider. Our findings can provide guidance for diabetic patients, advance the practice of online healthcare platform operations, and offer valuable suggestions for pharmaceutical companies to improve diabetic drugs. Furthermore, pharmaceutical companies and platform providers can expand their business value propositions.

First, WebMD managers must understand how to operate WebMD further. They need to understand that the emotions, review length, perceived effectiveness, and *PS* expressed by patients also have an intuitive influence on potential patients in judging helpfulness. To the best of our knowledge, WebMD does not currently contain a function with emotional language. Accordingly, WebMD platform managers can utilize this study's findings to improve websites to be more user-friendly and easy to use by creating a review assistance system that provides reviewers with emotional language (such as sad, happy, etc.). This would help them express their emotions more accurately, which, in turn, would be more helpful to potential patients. Moreover, we suggest that the WebMD platform promotes more reviews with extreme *PS* and long length for potential patients rather than neutral and short ones, which can help potential patients find useful information faster among numerous online reviews and improve the operational efficiency for the platform providers.

Second, our study indicated that reviews covering the topics of blood sugar, family medical history, dose time, and body changes helped potential patients. Consequently, we suggest that WebMD platform managers design new functions and items according to the topics of reviews and guide patients to write reviews from these aspects, such as writing/browsing reviews by topic and scoring the performance of drugs by subject. Thus, potential patients can directly browse the reviews of interest rather than aimlessly browse the reviews.

Finally, for pharmaceutical companies offering drug treatment for diabetes, the effectiveness of drugs, dosing time of taking drugs, injection of drugs, blood sugar, and medical history are elements that need to be emphasized and considered in the development of next-generation drugs. This is because they all significantly influence on the judgment of potential patients.

7.3 Limitations and Future Work

Limitations: Although this study presents several comprehensive insights into *RH* and opens doors for several research areas in online drug treatment reviews using text mining, econometrics, and design science, there are still some concerns and limitations in our work. First, we focus on unpacking the influence of emotions, review length, *PS*, perceived effectiveness, and perceived ease of use on *RH* based on the online review from WebMD. Hence, our study is limited to only the WebMD platform and diabetics. Second, the attributes of topics and emotions were extracted from textual data using texting mining techniques. Although this study's measurements of emotions follow the NRC word-emotion association lexicon referring to Gour et al. (2022), it is still a difficult task to measure different levels of emotions objectively. Finally, some factors, such as patients' financial ability, patients' self-management, etc. may have a considerable influence on patients' decision-making (Lu et al., 2021), but the lack of efficient data on these factors has limited us to further explore the influence of them on *RH*.

Future work: Several directions are worth pursuing in the future based on the aforementioned limitations. First, future research can further investigate and revalidate our findings on online reviews from patients with other chronic diseases and other healthcare platforms, such as the Drugs website. We would like to validate the generalizability of our results in other types of chronic diseases, such as high blood pressure, depression, coronary heart disease, etc. Second, the emotions expressed by reviewers in online reviews have long been an important object in studies related to online reviews (Siering et al. 2018; Gour et al., 2022), thus it is worthwhile to construct an advanced deep learning model to more objectively measure emotions in online reviews. Finally, it is interesting to incorporate more factors (patients' income, patients' dietary habits, etc.) to explore the perceived helpfulness of the potential patients in the future. Furthermore, self-management, including self-management includes drug management, family doctor communication management, lifestyle management, etc. (Dadgar and Joshi 2018), is an effective way to manage diabetes (Dadgar and Joshi 2018). Thus, it is worthwhile to incorporate multiple self-management programs to study *RH* for future research.

Acknowledgments

This paper was supported in part by the National Natural Science Foundation of China under grant numbers 72171161, 71971041, and 71871148; by the Outstanding Young Scientific and Technological Talents Foundation of Sichuan Province under grant number 2020JDJQ0035; and by the Sichuan University to Building a World-class University under Grant No. SKSYL2021-08.

References

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

Arnott, D., & Pervan, G. (2012). Design science in decision support systems research: An assessment using the hevner, march, park, and ram guidelines. *Journal of the Association for Information Systems*, *13*(11), 923-949.

Aschner, P., Karuranga, S., James, S., Simmons, D., Basit, A., Shaw, J. E., et al. (2021). The international diabetes federation's guide for diabetes epidemiological studies. *Diabetes Research and Clinical Practice*, *172*, 108630.

Baechle, C., Huang, C. D., Agarwal, A., Behara, R. S., & Goo, J. (2020). Latent topic ensemble learning for hospital readmission cost optimization. *European Journal of Operational Research*, *281*(3), 517–531.

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.

Cao, Q., Duan, W., & Gan, Q. (2011). Exploring determinants of voting for the "helpfulness" of online user reviews: A text mining approach. *Decision Support Systems*, *50*(2), 511–521.

Cavanaugh, L. A., MacInnis, D. J., & Weiss, A. M. (2016). Perceptual dimensions differentiate emotions. *Cognition and Emotion*, *30*(8), 1430–1445.

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

Journal of Hospitality Management, *85*, 102356.

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.

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.

Chua, A. Y. K., & Banerjee, S. (2015). Understanding review helpfulness as a function of reviewer reputation, review rating, and review depth. *Journal of the Association for Information Science & Technology*, *66*(2), 354–362.

Craciun, G., Zhou, W., & Shan, Z. (2020). Discrete emotions effects on electronic word-of-mouth helpfulness: The moderating role of reviewer gender and contextual emotional tone. *Decision Support Systems*, *130*, 113226.

Dadgar, M., & Joshi, K. D. (2018). The role of information and communication technology in self management of chronic diseases: An empirical investigation through value sensitive design. *Journal of the Association for Information Systems*, 86–112.

Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *Journal of Abnormal & Social Psychology*, *51*(3), 629–636.

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. (2016). What makes an online consumer review trustworthy? *Annals of Tourism Research*, *58*, 46–64.

Filieri, R., Raguseo, E., & Vitari, C. (2018). When are extreme ratings more helpful? Empirical evidence on the moderating effects of review characteristics and product type. *Computers in Human Behavior*, *88*, 134–142.

Fink, L., Rosenfeld, L., & Ravid, G. (2018). Longer online reviews are not necessarily better. *International Journal of Information Management*, *39*, 30–37.

Forman, C., Ghose, A., & Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information Systems Research*, *19*(3), 291–313.

Gour, A., Aggarwal, S., & Kumar, S. (2022). Lending ears to unheard voices: An empirical analysis of user generated content on social media. Production and Operations Management.

Hautala, G. S., Comadoll, S. M., Raffetto, M. L., Ducas, G. W., Jacobs, C. A., Aneja, A., & Matuszewski, P. E. (2021). Most orthopaedic trauma patients are using the internet, but do you know where they're going? *Injury*, *52*(11), 3299–3303.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly*, *28*(1), 75–105.

Hong, Y., Huang, N., Burtch, G., & Li, C. (2016). Culture, conformity, and emotional suppression in online reviews. *Journal of the Association for Information Systems*, *17*(11), 737–758.

Jacoby, J. (1977). Information load and decision quality: Some contested issues. *Journal of Marketing Research*, *14*(4), 569–573.

Johnson, E. J., & Meyer, R. J. (1984). Compensatory choice models of noncompensatory processes: the effect of varying context. *Journal of Consumer Research*, *11*(1), 528–541.

Kordzadeh, N., & Warren, J. (2017). Communicating personal health information in virtual health communities: an integration of privacy calculus model and affective commitment. *Journal of the Association for Information Systems*, *18*(1), 45–81.

Kuan, 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.

Lamano, J. B., Riestenberg, R., Mendoza, A. H., Lee, D., Li, Y., & Bloch, O. (2021). Social media utilization by academic neurosurgery departments correlates with higher online patient satisfaction scores. *Neurosurgery*, *67*(1), 131–132.

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.

Lind, J. T., & Mehlum, H. (2010). With or without U? The appropriate test for a U-shaped relationship. *Oxford Bulletin of Economics and Statistics*, *72*(1), 109–121.

Liu, Z., & Park, S. (2015). What makes a useful online review? Implication for travel product websites. *Tourism Management*, *47*(47), 140–151.

Lu, W., Hou, H., Ma, R., Chen, H., Zhang, R., Cui, F., et al. (2021). Influencing factors of patient satisfaction in teleconsultation: A cross-sectional study. *Technological Forecasting and Social Change*, *168*, 120775.

Lutz, B., Pröllochs, N., & Neumann, D. (2022). Are longer reviews always more helpful? Disentangling the interplay between review length and line of argumentation. *Journal of Business Research*, *144*, 888–901

Malhotra, N. K. (1982). Information load and consumer decision making. *Journal of Consumer Research*, *9*(4), 419–430.

Mallipeddi, R. R., Janakiraman, R., Kumar, S., & Gupta, S. (2021). The effects of social media content created by human brands on engagement: Evidence from Indian general election 2014. *Information Systems Research*, *32*(1), 212–237.

Meek, S., Wilk, V., & Lambert, C. (2021). A big data exploration of the informational and normative influences on the helpfulness of online restaurant reviews. *Journal of Business Research*, *125*, 354–367.

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

Olbrich, Rainer, Holsing, & Christian. (2012). Modeling consumer purchasing behavior in social shopping communities with clickstream data. *International Journal of Electronic Commerce*, *16*(2), 15–40.

Park, S. H., & Hong, S. H. (2018). Identification of primary medication concerns regarding thyroid hormone replacement therapy from online patient medication reviews: Text mining of social network data. *Journal of Medical Internet Research*, *20*(10), e11085.

Pavlou, P. A., & Dimoka, A. (2006). The nature and role of feedback text comments in online marketplaces: Implications for trust building, price premiums, and seller differentiation. *Information Systems Research*, *17*(4), p.392–414.

Pumawirawan, N., Eisend, M., Pelsmacker, P. D., & Dens, N. (2015). A meta-analytic investigation of the role of valence in online reviews. *Journal of Interactive Marketing*, *31*, 17–27.

Ray, A., Bala, P. K., & Rana, N. P. (2021). Exploring the drivers of customers' brand attitudes of online travel agency services: A text-mining based approach. *Journal of Business Research*, *128*, 391–404.

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.

Risselada, H., de Vries, L., & Verstappen, M. (2018). The impact of social influence on the perceived helpfulness of online consumer reviews. *European Journal of Marketing*, *52*(3/4), 619–636.

Saeedi, P., Petersohn, I., Salpea, P., Malanda, B., Karuranga, S., Unwin, N., et al. (2019). Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition. *Diabetes Research and Clinical Practice*, *157*, 107843.

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.

Schwenk, C. R. (1986). Information, cognitive biases, and commitment to a course of action. *Academy of Management Review*, *11*(2), 298.

Siering, M., Deokar, A. V., & Janze, C. (2018). Disentangling consumer recommendations: Explaining and predicting airline recommendations based on online reviews. *Decision Support Systems*, *107*, 52–63.

Srivastava, V., & Kalro, A. D. (2019). Enhancing the helpfulness of online consumer reviews: The role of latent (content) factors. *Journal of Interactive Marketing*, *48*, 33–50.

Tversky, A., & Kahneman, D. (1978). Judgment under uncertainty: heuristics and biases. *Science*, *185*(4157), 17–34.

Wang, X., Tang, L. (Rebecca), & 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., & Xu, W. (2018). Leveraging deep learning with LDA-based text analytics to detect automobile insurance fraud. *Decision Support Systems*, *105*, 87–95.

Wen, H., Park, E., Tao, C. W., Chae, B., & Kwon, J. (2019). Exploring user-generated content related to dining experiences of consumers with food allergies. *International Journal of Hospitality Management*, *85*, 102357.

Wu, J. J., & Chang, S. T. (2020). Exploring customer sentiment regarding online retail services: A topic-based approach. *Journal of Retailing and Consumer Services*, *55*, 102145.

Yang, Z., Gao, Y., & Fu, X. (2021). A decision-making algorithm combining the aspect-based sentiment analysis and intuitionistic fuzzy-VIKOR for online hotel reservation. *Annals of Operations Research*. https://doi.org/10.1007/s10479-021-04339-y

Yin, D., Mitra, S., & Han, Z. (2014). When do consumers value positive vs. negative reviews? An empirical investigation of confirmation bias in online word of mouth. *MIS Quarterly*, *38*(2), 539–560.

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

Zeithaml, V. (1988). Consumer perceptions of price, quality, and value: A means-end model and synthesis of evidence. *Journal of Marketing*, *52*, 2–22.

Zhang, J. Q., Craciun, G., & Shin, D. (2010). When does electronic word-of-mouth matter? A study of consumer product reviews. *Journal of Business Research*, *63*(12), 1336–1341.

Zhao, Y., Xu, X., & Wang, M. (2019). Predicting overall customer satisfaction: Big data evidence from hotel online textual reviews. *International Journal of Hospitality Management*, *76*, 111–121.

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