Quality assessment of temporomandibular disorders-related information on Chinese social media: A cross-sectional study

Yifei Deng [ID] , Jianing Zhou [ID] , Ming Yang [ID] , Yaxin Weng and Xin Xiong' [ID]

Abstract

Background: Temporomandibular disorders (TMDs) affect people's quality of life greatly, and precise understanding of TMDs contributes to a proper treatment choice. Social media is an access to health information, hence it is needed to evaluate the quality of relative information on social media.

Objective: This research aims to assess the quality of information about TMDs understood by the public on two mainstream social media platforms, WeChat and Zhihu. They will be evaluated from four aspects: readability, credibility, concreteness, and accuracy.

Methods: Researchers searched for relative articles on WeChat and Zhihu and selected the samples. The readability was evaluated separately, and the DISCERN instrument was employed to evaluate the credibility and concreteness. Accuracy was measured by comparing samples with authoritative journals and textbooks. The Health On the Net code of conduct for medical and health websites (HONcode) and Global Quality Scale (GQS) were used as supplemental tools. Two researchers conducted this process independently, the intraclass correlation coefficient was used to examine the consistency.

Results: One hundred and eleven articles were included, with 47 articles from WeChat, 64 from Zhihu. For readability, the articles received a mean score of 27.79 (standard deviation (SD) 2.99) out of 35. The DISCERN instrument reported a mean score of 38.52 (SD 7.13) out of 80. As for accuracy, most articles (92 of 111) got 3.5 or more out of 5, demonstrating that the two platforms did well in this area. HONcode reported a mean score of 6.29 out of 16 (SD 1.42) while GQS showed a mean score of 2.91 out of 5 (SD 0.77), indicating the reliability needs improvements, and these articles can only provide limited help to the public.

Conclusions: The quality of TMDs-related information from WeChat and Zhihu is generally low. Although they do well on accuracy and readability, the credibility and concreteness still need further improvements. And different improvements and suggestions are recommended for uploaders and platforms.

Introduction

Temporomandibular disorders (TMDs) are defined as a group of clinical conditions affecting the muscles of mastication, the temporomandibular joint, and the related structures.[1] It is estimated that 29% of Chinese population is suffering from TMDs.[2] The typical symptoms include joint pain, joint popping, joint dislocation, and joint movement disfunction, all of which significantly impair the quality of life for patients and exert a profound impact on

State Key Laboratory of Oral Diseases & National Center for Stomatology & National Clinical Research Center for Oral Diseases, West China Hospital of Stomatology, Sichuan University, Chengdu, Sichuan, China

Corresponding author:
Xin Xiong, State Key Laboratory of Oral Diseases & National Center for Stomatology & National Clinical Research Center for Oral Diseases, West China Hospital of Stomatology, Sichuan University, Chengdu 610041, Sichuan, China.
Email: drxiongxin@scu.edu.cn

society as a whole.[3,4] Moreover, TMDs are among the chronic pain conditions that constitute a significant global health issue. Chronic pain turns out to be a major source of suffering, as it interferes with daily functioning.[5] Additionally, it has been demonstrated through epidemiological research that individuals suffering from TMDs are prone to experiencing a range of comorbidities. These encompass headaches, widespread pain, fibromyalgia, as well as neck and back pain. Apart from mental pain, individuals suffering from TMDs may encounter more psychosocial issues like stress, anxiety, depression, and pain catastrophizing.[6] As a form of hypochondriasis, TMDs manifest as an excessive preoccupation with one's well-being, ranging from minor concerns to deep distress.[7] This kind of concern is distinguished by a sense of dread regarding illness, not an unwavering belief in its presence, which is called health anxiety.[8] Health anxiety is a usual health issue that results in significant distress, puts pressure on the relationship between doctors and the public, and raises the amount spent on healthcare.[9,10] Presently, it is acknowledged that TMDs' symptoms may spontaneously abate over a variable period, ranging from several days to several months, and some treatments can yield favorable outcomes in the early stage of this disease. Early intervention has the potential to significantly alleviate TMDs-related pain and substantially improve the individuals' quality of life.[11] The absence of appropriate treatment may lead to a deterioration of the condition, and more intense treatments are required.[4] Given the multifaceted and intricate etiology of TMDs, the correct understanding of TMDs is crucial for enhancing the capacity to detect the initial stages of the condition. Once the public acquire relevant knowledge, they will be able to recognize the situation and seek help from the professionals when they encounter the early stages of TMDs, which leads to a better prognosis. Moreover, such knowledge can influence people's subjective attitudes and perceptions of their health status. Individuals can be educated to adopt proper behaviors to protect their joints and avoid harmful actions. From this perspective, fostering awareness of relevant knowledge stands as one of the most effective preventative measures against TMDs. However, the dissemination of inaccurate information could potentially mislead the public, leading to erroneous choices in treatment options.

Getting accurate information and starting conservative treatments early in the disease can make a big difference. Many articles contained in this research highlight numerous cases where individuals failed to seize the optimal window for TMDs treatment. This phenomenon emphasizes the necessity of awareness of TMDs as well. Despite the public, even professionals lack the proper understanding of TMDs. Most Chinese dental students have not taken any TMDs-related courses, and 75% postgraduate students and 85% dentists admitted that they have not undergone the systematic training in TMDs.[12] The current scenario unequivocally establishes the critical significance of high-quality online data.

In addition, China has the biggest number of Internet users in the world, with social media emerging as one of the most vital resources of information. WeChat is one of the most popular social media platforms, attracting a diverse user base including individuals, organizations and media. Tencent, parent company of WeChat, declared that by the end of the first quarter of 2022, WeChat has 1.2 billion monthly active users in many different countries, making it one of the biggest social media in the world. As a social media platform, WeChat encourages all its users to post different information about nearly every aspect of life on the platform. In addition, the majority of WeChat users are the young, with a high likelihood of being diagnosed with TMDs. According to existing data, 20.9% of its users are under 24, while 25.65% fall between the ages of 24 and 30. More specifically, in the case of medical field, 98.35% of the participants admitted that they have seen health-related information on WeChat, making WeChat one of the most popular resources of health-related information.[13] Zhihu, a question-and-answer platform whose users upload abundant information as well, is popular among the young. According to the available data, Zhihu has an average monthly active user base of 1.013 hundred million. Zhihu's user base is not limited to China, with the advancement of its internationalization strategy, Zhihu is also attracting overseas users, especially the Chinese diaspora and foreigners interested in Chinese culture. Among these users, 30.86% fall into the age group of 24 years and below, while 29.38% are aged between 25 and 30 years old. These statistics suggest that a significant portion of Zhihu users could potentially be at a higher risk of being diagnosed with TMDs. On Zhihu, users can access posts pertaining to diverse aspects of TMDs, like science popularization, treatment experience sharing and so on.

In recent years, TMDs are gaining more and more attention of the researchers, and the annual number of publications and citations keep increasing greatly.[14,15] As for the public, it is more and more access to the TMDs-related knowledge. But in contrast to the great amount of information on social media, the quality of information about TMDs is mostly unclear. In fact, quality assessments of information about cancer on WeChat showed an unsatisfying result, the majority of the information on WeChat is not regulated and can only ranked as "fair,"[16,17] indicating that such platforms may not be ideal enough to gain professional knowledge, and this problem can also be true for TMDs-related information.

For this reason, a significant conflict has emerged concerning the public's requirements for accurate and reliable information on the platforms and the enigmatic nature of the information's quality. To address this issue, it is

important to conduct such an assessment on the quality of TMDs-related information, otherwise, the public are at the risk of being misled. This evaluation serves as a guidance, directing individuals toward the acquisition of accurate and reliable information. Extending its impact beyond the general public, this study can also influence various professional groups. Doctors can obtain supplementary knowledge from social media, which can enhance their overall understanding of TMDs. For the platforms, the study provides a framework to assess the caliber of content, enabling them to implement targeted policies or enhancements to ensure the dissemination of high-quality information.

As mentioned above, the rationale for this study is to assess the quality of information about TMDs. With a significant portion of the Chinese population affected and social media being a primary source of health information, evaluating the quality of TMDs-related content on platforms like WeChat and Zhihu is essential. These platforms have a vast user base, but the quality of health information provided is mostly unknown, posing a risk of misinformation. This study aims to assess and improve the quality of TMDs information, guiding users to reliable sources and aiding in the development of educational strategies. It also contributes to the broader field by establishing an evaluation framework for health information on social media.

Methods

This research is a cross-sectional study, and aims to conduct an assessment of the TMDs-related information. The evaluation process in conducted from December 10th, 2023 to March 15th, 2024, at the West China School of Stomatology, Sichuan University. All the researchers involved in this study are undergraduate and graduate students from the West China School of Stomatology, Sichuan University. These individuals have undergone extensive training in subjects related to this research and possess a wealth of professional expertise necessary to successfully undertake this task.

The DISCERN instrument[18,19] was employed to appraise the credibility and concreteness. Authoritative journals and textbooks were utilized as references to assess the accuracy, precise disease diagnoses and treatment standards were also included in this step.[1,4,20–23] Health On the Net code of conduct for medical and health websites (HONcode)[24] and Global Quality Scale (GQS)[25] were used as supplementary tools.

Ethical considerations and the consent statement

This research has been reviewed and approved by the Institutional Review Board of West China Hospital of Stomatology for ethical exemption.

This research utilizes existing medical records or data, all of which are sourced from the public internet and voluntarily uploaded and shared by the authors. The study does not involve any privacy of the authors, nor does it pose any risk or harm to them. Additionally, the study does not involve any commercial interests. The Institutional Review Board of West China Hospital of Stomatology has waived the requirement for the consent statement.

The project was discussed in December, 2022, and the Institutional Review Board reference number is WCHSIRB-D-2022-240.

Search strategy and data extraction

In the beginning, the researchers conducted a search on WeChat and Zhihu to retrieve articles related to TMDs. Researchers entered the Chinese terms "Temporomandibular Joint," "Temporomandibular Disorders," "Temporomandibular Disorders syndrome," "Temporomandibular Disorders Treatment," and "Temporomandibular Disorders Symptoms" into WeChat and Zhihu for retrieval of relevant articles. To ensure a representative sample, researchers ranked the search results according to the posted time. If following the acquiescent order, the better articles (better accuracy, better completeness, more likes, more views, etc.) were given priority and listed ahead of others. These samples may not be indicative of the overall quality of information available on the platforms. Then researchers determined whether an article should be included by certain inclusion criteria: (a) articles containing information about any aspect of TMDs; (b) articles posted between January 1st, 2022 and January 6th, 2023; (c) article contents covering more than two pages in PDF documental form. Since this research focuses on assessing the overall quality of information related to TMDs, any theme pertaining to TMDs is acceptable. This includes introduction of TMDs, treatments, experience, or cases sharing and so on. It should be noted that other academic factors, such as the specific types of TMDs, are not considered in the inclusion process. The exclusion criteria were as follows: (a) in any form other than article like picture, video, or voice record; (b) most part of the articles concentrating on irrelevant information like complaints or advertisements; (c) duplicated articles; and (d) professional publications for the professional.

After applying these criteria, 111 articles were included to conduct this study, and researchers recorded the characters of each article like uploader, and main contents. The entire process is depicted in Figure 1.

Evaluation of dimensions and methods

Overview

To measure readability, credibility, concreteness, and accuracy in this study, the researchers divided the whole process into the following four steps. First, researchers set up a scale to evaluate the readability. Then, researchers
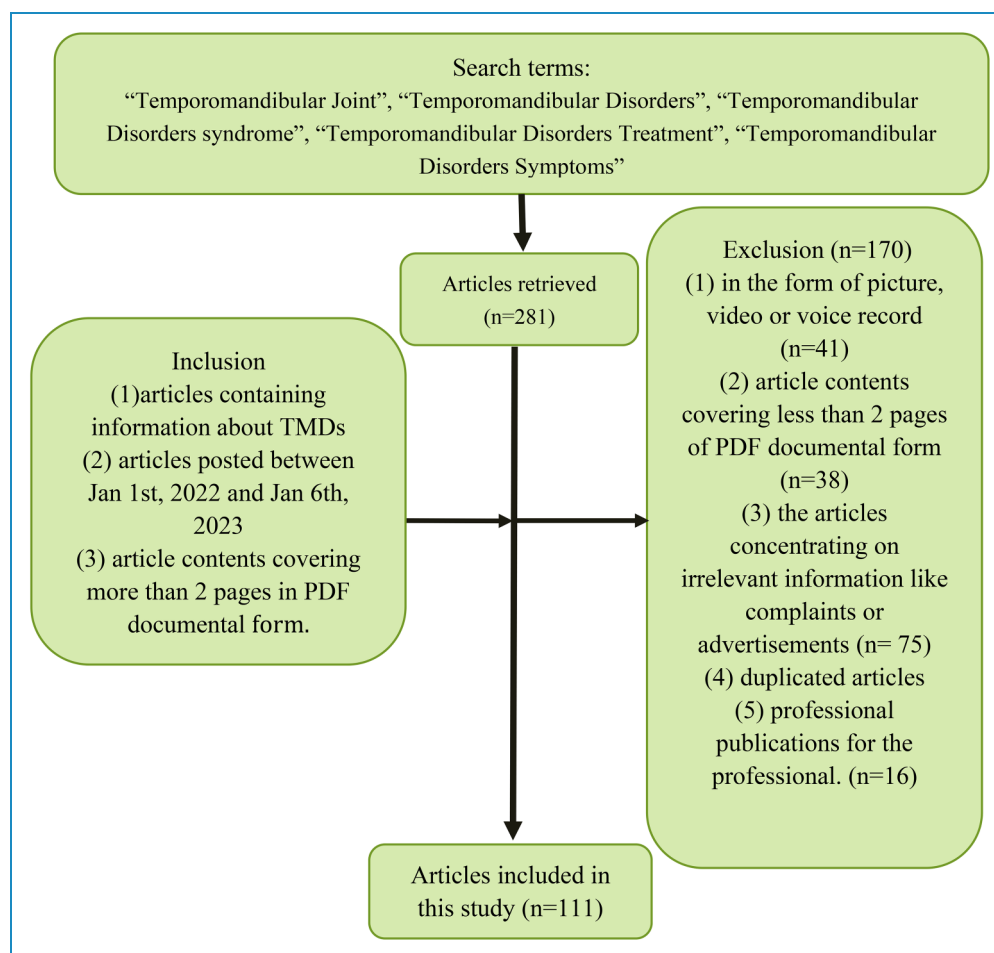
Figure 1. Search and samples selection flow in this study.

utilized the DISCERN instrument, a well-established method used globally, to evaluate the credibility and concreteness of the samples.[19] Third, researchers measured the accuracy by comparing samples with authoritative journals and textbooks, as well as precise disease diagnoses and treatment standards.[1,4,20–23] Schiffman et al.,[1] Liu and Steinkeler,[4] Gauer RL and Semidey,[20] Okeson,[21] Nelson,[22] and Yost et al.[23] are the actual materials utilized for the accuracy assessment. Finally, the HONcode and GQS were employed to evaluate the reliability and the utility of the information as supplement.[19]

Two researchers conducted the whole process and scored all the articles independently, then the ICC was used to measure the consistency between the researchers. ICC is a value ranging from 0 and 1 used to measure the consistency, with higher value indicating better consistency. And it is widely admitted that ICC > 0.80 indicates very strong consistency, ICC from 0.41 to 0.80 means medium consistency, ICC < 0.40 reports poor consistency.[26,27] When the difference in scores between two researchers exceeds 1 point, a third researcher is introduced into the rating process. The entire process is depicted in

Figure 2. In addition, researchers used the test method kurtosis-skewness for Gaussian distribution to judge the distribution of article scores, and the results were normally distributed (Kurtosis and Skewness are within ±1.96). Analysis of Variance (ANOVA) was used to assess the significance of the differences in scores between articles in different categories. Results were considered statistically significant at a level of $p < .05$. Finally, for the correlation between the data on the number of reviews and the article scores, the researchers employed linear regression as a means of calculation. Outcomes that showcased a correlation between these two variables were deemed statistically significant, reaching a threshold of $p < .05$. All analyses were performed using Excel 2021 (2303 Build 16.0.16227.20202) and SPSS 24.0.[28–30]

Evaluation of readability

The aim of most articles in this study is to cater to the public and various groups with limited knowledge in the field. Consequently, these articles should not be overly specialized or academically inclined.
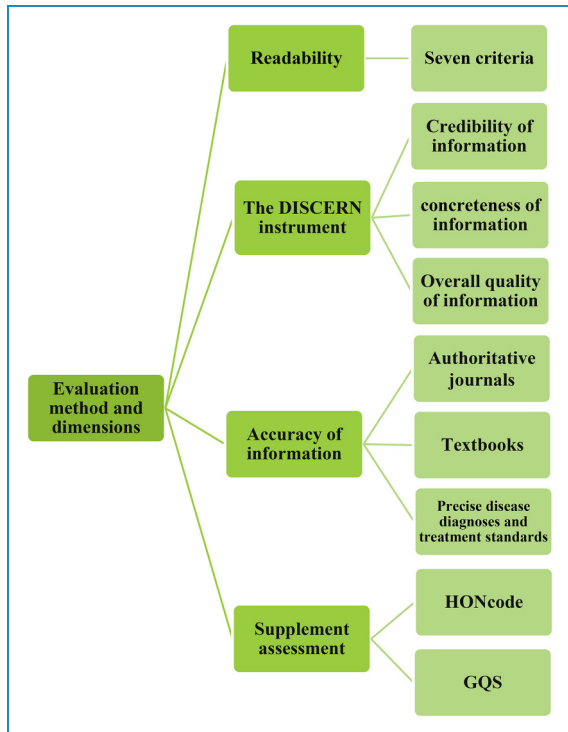
Figure 2. Flow chart of the whole assessment process.

However, the typical tools employed to evaluate readability and comprehensibility are not accessible for this particular study. The most widely recognized and validated instruments are the Gunning-Fog Index (GFI), Flesh Reading Ease Scores (FRES), and Flesch-Kincaid Grade Level. They are calculated through a formula that integrates the syllable count per word and sentence length, hence, the number of syllables is a crucial factor.[31] However, in Chinese, each character inherently represents a single syllable. The syllable count is solely determined by the number of characters in a word, and most Chinese words consist of either two or four characters. No matter how difficult or how easy to read they are, the number of syllables is the same. Consequently, the GFI and FRES of different articles can be strikingly similar in Chinese, even if their actual readability varies significantly.

To assess readability and comprehensibility, researchers established seven criteria. Each article was assigned a score ranging from 1 to 5 for each criterion, with the total points representing its overall readability and comprehensibility.

The seven criteria and their respective standards are outlined below:

1. Proper noun ratio: Except for necessary proper nouns, a higher inclusion of proper nouns leads to a lower score.
2. Rare characters/complex sentences ratio: Fewer rare characters and complex sentences contribute to a higher score.

3. Relevance of professional knowledge to the topic: The degree of relevance between the professional knowledge provided and the topic determines whether the public require additional expertise. The closer the relevance, the higher the score.
4. Number of illustrations: A greater number of illustrations results in a higher score.
5. Interactivity: Increased interactivity within the article leads to a higher score.
6. Layout rationality: A visually appealing and well-structured layout enhances the reading experience, thus contributing to a higher score.
7. Grammar mistakes and typos: Errors in grammar and typos can undermine the quality of an article. Consequently, a higher number of mistakes results in a lower score.

Two researchers carried out this process separately, and ICC was used to measure the consistency of the results. The ICC values for WeChat and Zhihu were 0.905 and 0.935 respectively, indicating strong consistency between the researchers.

DISCERN: Assessment of credibility and concreteness of TMDs related information

DISCERN serves as a tool utilized to evaluate a broad range of consumer health information on treatment in a variety of formats. It is, in essence, a questionnaire containing 16 questions. Questions 1–8 center around the reliability of the publications, and Questions 9–15 pay attention to the details of the information concerning treatment choice. Question 16 is a subjective question about the overall quality of the article. An article ought to be given 1 to 5 points for each question. One point means the quality criterion has not been fulfilled at all, while 5 points means the quality criterion has been completely fulfilled.[32] Depending on the score of each article, they were divided into five categories: excellent (score > 63), good (62.5 > score > 51), fair (50.5 > score > 39), poor (38.5 > score > 27), and very poor (score < 26.5), with the maximum score being 80.[33]

Two researchers carried out this process separately, and ICC was used to measure the consistency of the results. In this study, the ICC for the credibility of content on WeChat and Zhihu were calculated to be 0.999 and 0.998 respectively, while for concreteness, the ICC values for WeChat and Zhihu were 0.988 and 0.996 respectively. Thus, it is believed that there's very strong consistency between two researchers.

Assessment of accuracy

Four categories were set up and each assigned a different point value: perfectly accurate (5 points); partially accurate

(4 points); only a few mistakes (3 points); many mistakes (1–2 points). Authoritative journals and textbooks were used as reference, as well as precise disease diagnoses and treatment standards.

Two researchers performed this process separately and ICC was used to examine the consistency.

HONcode and GQS

HONcode, established by the Health On the Net Foundation, was a set of principles aimed at regulating the health-related websites.[24,34] Eight principles were included in this tool and each representing 2 points. Zero means the criterion has not been fulfilled at all, 1 means the criterion has been partly fulfilled, and 2 means the criterion has been completely fulfilled. The utility of the information was assessed using the GQS, a five-point scale that assigned articles a score ranging from 1 to 5.[25] The scores were then used to categorize the articles into different levels, ranging from excellent to poor. Figures 3 and 4 show the scales and standards.

Categories of samples and uploaders

At the very beginning of the process, researchers recorded the main contents and uploaders of the samples and divided them into different groups. For contents, researchers set up four groups: (a) introduction to TMDs or the cause of TMDs; (b) effects of TMDs on orthodontic treatment; (c) specific treatment, prevention method, and relative reminding; (d) experience or cases sharing. Besides, researchers divided the uploaders into four groups: (a) medical organizations; (b) doctors; (c) individuals; and (d) science popularization institutions. Differences between groups were analyzed.

Statistical analysis

Excel 2021 (2303 Build 16.0.16227.20202) and SPSS 24.0 were used to analyze and represent the data. Researchers calculated the mean score and standard deviation (SD) of each question and used frequencies and percentages to illustrate the group division. The ICC was calculated between two researchers.

Result

There were 111 articles from WeChat and Zhihu which met the inclusion criteria. In terms of the information types, 66.67% (74/111) of articles including introduction to TMDs or the cause of TMDs, 9.91% (11/111) referred to the effects of TMDs on orthodontic treatment, 41.44% (46/111) involved specific treatment, prevention method, and relative reminding, and 36.94% (41/111) of them are experience or cases sharing. With regard to the uploading source, the majority of the articles have been posted by medical institutions (41/111, 36.94%), followed by individuals (37/111, 33.33%), doctors (26/111, 23.42%), and science popularization institutions (7/111, 6.31%) (Table 1).

In terms of readability, all articles (with scores ranging from 7 to 35 points) were equidistant into four levels, most of the articles scored above Good (more than 95% of the articles). Among the 3 evaluate dimensions of DISCERN, the scores of TMD-related articles on social media were mostly graded Fair and Poor (98.2% of the articles on these 2 grades or lower), except for Accuracy on a Good grade (82.88% of the articles achieved good grade or higher on this dimension) (Table 2).

In addition, Table 3 shows that the p-value of normality test (K-S test) of frequencies of all articles scores in the DISCERN are more than 0.05, indicating that these scores are in normal distribution, while Readability, Accuracy, HON, and GQS are not.

The comparison of quality scores of WeChat and Zhihu and the p-value of the test on the average scores of articles from these two platforms in different dimensions is shown in Table 4. In the 3 evaluation dimensions of DISCERN, articles on WeChat got 23.33, 13.79, and 3.22 points respectively, articles on Zhihu got 21.88, 13.25, 2.86 points respectively. The resulting scores shows that articles on Wechat scored higher in DISCERN, Accuracy and HON scales ($p < .05$). Zhihu's article scored higher for many dimensions of readability ($p < .05$).

Evaluation scores according to the uploader are shown in Table 5. For articles on WeChat, in the scores of DISCERN and HON scales, which were statistically different ($p < .05$) after a multifactor ANOVA, articles uploaded by doctors scored the highest, while those uploaded by individuals scored the lowest. And it is worth noting that articles uploaded by science organizations scored the lowest on both the Proper Noun Ratio and the Rare Characters/ Complex Sentences Ratio in the refinement scores assessing readability ($p < .05$).

For articles on Zhihu, articles from doctors received the highest scores for readability, while science popularization institutions received the highest scores for HONcode ($p < .05$).

Articles with a single aspect of the content on WeChat scored higher than articles with multiple aspects of content in DISCERN and GQS scales ($p < .05$).

As for Zhihu, articles with a single aspect of content also scored higher than articles containing multiple aspects of content in accuracy ($p < .05$), and there were no statistically significant differences in any of the other dimensions (Table 6).

Discussion

During this study, researchers collected articles on WeChat and Zhihu, and assessed their quality from four aspects:

The HONcode	
Principle	
1	Any advice is given by the professionals
2	Information aims at supporting the relationship between patients and doctors
3	Respect privacy and personal data
4	Information is supported by reference, and have HTML links to the data
5	Claims will be supported by principle (4)
6	Information is presented in the clearest way and contact information is showed
7	The support and founding are clearly identified
8	Advertisement is distinguished form other content

Figure 3. HONcode in details.
HONcode: Health On the Net code of conduct for medical and health websites.

readability, credibility, concreteness, and accuracy. Different scales were used for this process including the DISCERN instrument, HONcode, GQS and textbooks. Two researchers conducted the whole process independently and the ICC was used to test the consistency between the two researchers. The results showed that the credibility and concreteness need significant improvements, while the accuracy and readability performed well.

Association and comparison with existing research

Compared to previous studies, this study found similar results. For example, a study evaluating the quality of TMDs-related videos on YouTube reported low quality with a mean score of 31.1 (SD 4.16) using the DISCERN instrument.[35] Another study assessing the quality of websites providing health information on minimal invasive repair of pectus excavatum found a mean score of 42.5 (SD 12.2) for the DISCERN instrument with serious shortcomings. In addition, the study also concluded that the overall quality is low to moderate.[36] Similarly, a study on hand osteoarthritis found online information is low in quality generally and was difficult for the general public to read.[37] These studies, in conjunction with the present study, collectively suggest that the general health-related information available on the internet is mostly of poor quality. This phenomenon can be attributed to the unfettered accessibility of the internet, enabling indiscriminate publication of articles regardless of the author's competence or specialization. In this study, many of the articles consisted are uploaded by individuals, which are often subjective and unprofessional, leading to potential misinformation.

As for the difference, this study found good accuracy in information from WeChat and Zhihu, the previous study reported poor accuracy for TMDs-related videos on YouTube.[35] The reasons for the difference are as follows: (a) many articles were given high scores in this field: 23 articles cannot be judged as "accurate" or "inaccurate" since they were personal experiences and feelings shared by the public. For this reason, they were given nearly full marks in this field. (b) Creating these articles were easy: articles in this study primarily served the purpose of educational dissemination and scientific popularization. The knowledge presented is basic and simply explained, making it easy to avoid mistakes. (c) Articles are more convenient to edit: compared to videos, articles can be exam and corrected more easily, contributing to high scores in case of accuracy.

In terms of the HONcode and GQS, the results of this study were consistent with previous studies. Videos on YouTube got 5.64 (SD 1.00) for HONcode and 2.43 (SD 0.64) for GQS.[35] While the mean score in this study was 6.03 (SD 1.16) and 2.92 (SD 0.54) separately, indicating a similar lack of adherence to formal academic regulations on social media platforms. The fifth, seventh, and eighth questions in HONcode, which focus on support for claim-solving, identity of support for the website, and statement of advertisement, were worst performed with a mean score of 0 (SD 0). The lack of academic regulations was a common phenome on social media, mainly because the awareness of copyright, funding, and so on was particularly lacking for the whole society. The main contents were users' first priority, and users hardly pay attention to these regulations. This result suggested that the adherence of formal academic regulations can contribute to the overall quality and reliability of health-related information available online.

Potential improvements for these two platforms

When initially searching for samples, it was obvious that the two platforms can be improved in different aspects.

The GQS	
Score	Description
1	Poor quality, not useful for the patients
2	Generally poor and limited use for patients
3	Moderate quality and partly useful
4	Good quality and useful for patients
5	Excellent quality and very useful

Figure 4. GQS in details.
GQS: Global Quality Scale.

Table 1. Characteristics of articles related to TMD on internet ($n=$ 111).

Variable	Values
Category, *n* (%)	
Information category	
Introduction to TMDs or the cause of TMDs	74 (66.67%)
Effects of TMDs on orthodontic treatment	11 (9.91%)
Specific treatment, prevention method, and relative reminding	46 (41.44%)
Experience or cases sharing	41 (36.94%)
Uploading source	
Medical institutions	41 (36.94%)
Doctors	26 (23.42%)
Individuals	37 (33.33%)
Science popularization institutions	7 (6.31%)

TMD: temporomandibular disorder.

Some articles included a substantial amount of advertising content promoting treatments for TMDs, medical checkups, hospitals, and other relevant services. They make up a big part of the content, but the platform does not clearly indicate when an article contained an advertisement. Even worse, advertisement showed little regard for the public seeking help, treating them solely as a means to profit. This phenomenon is not rare, six articles (5.41%) in this study contain advertisements. If the platforms can mark whether an article has ads, it will be beneficial.

Furthermore, while anyone can upload health-related information on these platforms, there is no guarantee of its accuracy and credibility. This is mainly due to the lack of the review process, and the difficulty in balancing the conflict between the cost and more professional reviewers for the platforms. If articles undergo a review process before being published on the websites, mistakes can be reduced in the samples. In this study, 11 articles (13.75%) were scored no more than 3, which means a few or even more mistakes were found, like "this is an uncommon disease", "you can eat whatever you like in case of TMDs". This result showed that such a review process is needed. Unlike other topics, health-related information directly impacts people's well-being, and incorrect information can lead to disastrous consequences. To improve the overall quality of the information, the platforms can consider implementing the following suggestions.

Firstly, the platforms should pay more attention to the supervision and regulation of health-related contents. Unsatisfying factors like mistakes and advertisements can be reduced if the platform reviews them before publishing or the writer obeyed the regulations. To solve the conflict between cost and the need for more professional reviewers, platforms can conduct research on leveraging artificial intelligence technologies, such as natural language processing and machine learning algorithms, to automatically review content prior to publication. By training models to detect false information, incorrect diagnoses, and treatment suggestions, these models can prompt authors to correct misleading content. This approach not only enhances the accuracy and credibility of information but also allows platforms to explore the integration of human and artificial intelligence review strengths. By doing so, they can establish a more efficient information quality supervision system. Secondly, the platforms should guarantee that professional crew or institutions were involved when creating health-related information for science popularization purposes. They can examine the content during the process and be responsible for it. Third, the platforms should give warnings to its user, reminding them of the potential risks and consequences of

Table 2. Scores and grade distribution of the 111 articles.

Category	Grade	Scores	Values	Percentage
			n	%
Readability				
	Excellent	(28, 35]	48	43.24
	Good	(21, 28]	61	54.95
	Fair	(14, 21]	2	1.80
	Poor	[7, 14]	0	0.00
DISCERN				
	Excellent	(62, 80]	0	0.00
	Good	(51, 62]	2	1.80
	Fair	(39, 51]	52	46.85
	Poor	(27, 39]	51	45.95
	Very poor	[16, 27]	6	5.41
Accuracy				
	Excellent	(4, 5]	22	19.82
	Good	(3, 4]	70	63.06
	Fair	(2, 3]	17	15.32
	Poor	[1, 2]	2	1.80
HON				
	Excellent	(12, 16]	0	0
	Good	(8, 12]	5	4.50
	Fair	(4, 8]	96	86.49
	Poor	[0, 4]	10	9.01
GQS				
	Excellent	(4, 5]	4	3.60
	Good	(3, 4]	28	25.23
	Fair	(2, 3]	54	48.65
	Poor	[1, 2]	25	22.52

GQS: Global Quality Scale; HON: Health On the Net.

Table 3. The normality test (K-S test) of frequencies of all articles scores.

	Statistics	p-value
Readability		
	0.084	.050*
DISCERN		
	0.059	.200
Accuracy		
	0.219	<.001***
HON		
	0.105	.004**
GQS		
	0.165	<.001***

GQS: Global Quality Scale; HON: Health On the Net.

Table 4. The average score of articles on Zhihu and WeChat in all aspects.

Category	Zhihu	WeChat	p-value
Readability	30.96	27.16	.09
Proper noun ratio	4.63	4.01	.005**
Rare characters/complex sentences ratio	4.84	4.68	.010**
Relevance of professional knowledge to the topic	3.87	4.27	.585
Number of illustrations	3.15	4.15	.001***
Interactivity	2.23	2.74	.415
Layout rationality	4.25	4.5	.004**
Grammar mistakes and typos	4.25	4.5	.004**
DISCERN	37.99	40.33	.369
Credibility	21.88	23.33	.394
Concreteness	13.25	13.79	.650
Overall evaluation	2.86	3.22	.099
Accuracy	3.62	4.16	<.001***
HON	6.00	6.84	.012*
GQS	2.81	3.14	.12

GQS: Global Quality Scale; HON: Health On the Net.

relying on health-related information from unprofessional sources. Users should be cautious that information obtained from the internet cannot replace professional diagnosis, and they should seek help from doctors. Currently, it is up to the uploader to include such a statement, but researchers believe the platforms should make it mandatory for every article. Lastly, the promotion of establishing high-quality content should be actively encouraged by the platforms. The lack of high-quality content can be attributed to a lack of motivation, as writers receive minimal tangible benefits. Therefore, the platforms should encourage them to be more productive, with the rewards such as monetary wage or enhanced social status playing a crucial role.

Potential improvements in related articles

In addition to the platform's supervisory role, authors have several methods to improve the quality of their articles. Firstly, they ought to adhere to the aforementioned regulations, such as the statement. Secondly, interactivity with the audience can enhance their interest and create a sense of warmth and concern. Authors can achieve this by actively engaging with the public within their articles, for instance, by incorporating questions like "Have you ever experienced…," "Do you know…," or "Let's learn this point

together". Moreover, authors should pay careful attention to the layout of their work. Articles lacking a well-designed layout can appear as an overwhelming stack of information, thereby diminishing readability and comprehension. Conversely, a visually appealing and logically organized layout can facilitate a seamless reading experience while showcasing clear levels of structure and coherence. Furthermore, the inclusion of more illustrations can significantly elevate the quality of articles. For instance, complex topics like the anatomy of the temporomandibular joint require extensive explanations, but a carefully crafted illustration can effectively convey the information in a clear and easily digestible manner.

Table 5. The scores of articles from different uploaders on the two platforms on the four scales.

Scores	Medical institution	Doctors	Individuals	Science popularization institution	*p*-value
Wechat					
Readability	29.05	30.36	28.99	25.38	.068
Proper noun ratio	4.33	3.90	4.31	2.88	.026*
Rare characters/complex sentences ratio	4.94	4.57	4.81	4.25	.023*
Relevance of professional knowledge to the topic	4.33	4.10	4.46	4.38	.097
Number of illustrations	3.61	4.02	4.65	4.38	.292
Interactivity	3.28	2.67	2.73	2.00	.128
Layout rationality	4.00	4.60	4.69	4.50	.332
Grammar mistakes and typos	4.00	4.60	4.69	4.50	.332
DISCERN	39.21	43.68	33.58	40.12	.032*
Credibility	22.45	24.07	20.25	23.50	.277
Concreteness	13.68	15.93	10.67	13.50	.082**
Overall evaluation	3.08	3.68	2.67	3.13	.097**
Accuracy	4.17	4.64	3.75	4.375	.07
HON	6.97	7.29	4.67	6.50	<.01**
GQS	2.98	3.57	2.58	3.25	.156
Zhihu					
Readability	26.64	28.92	26.49	25	<.01**
Proper noun ratio	4.86	4.55	4.65	4.00	.036*
Rare Characters/complex sentences ratio	4.91	4.76	4.89	4.50	.308
Relevance of Professional knowledge to the topic	4.36	4.21	3.48	3.83	.001***
Number of illustrations	2.18	3.82	3.13	2.67	.033*
Interactivity	1.91	2.87	1.97	2.00	.044*
Layout rationality	4.14	4.37	4.24	4.00	.397
Grammar mistakes and typo**s**	4.14	4.37	4.24	4.00	.397

(continued)

Table 5. Continued.

Scores	Medical institution	Doctors	Individuals	Science popularization institution	p-value
DISCERN	36.91	36.34	39.42	39.83	.51
Credibility	21.95	21.87	21.90	22.17	.999
Concreteness	12.14	11.66	14.62	14.50	.131
Overall evaluation	2.82	2.82	2.9	3.17	.928
Accuracy	3.64	3.79	3.55	3.67	.335
HON	6.82	5.79	5.72	7.33	.049*
GQS	2.91	2.71	2.87	2.67	.854

GQS: Global Quality Scale; HON: Health On the Net.

Table 6. A comparison of different scores on the amount of aspects covered by articles on the two platforms.

Scores	Articles containing a single aspect of content	Articles containing multiple aspects of content	p-value
WeChat	n = 12	n = 35	
Readability	28.17	28.81	.591
DISCERN	33.29	41.27	<.01**
Accuracy	3.92	4.30	.069
HON	6.08	6.89	.065
GQS	2.58	3.2	.023*
Zhihu	n = 42	n = 22	
Readability	26.77	27.95	.060
DISCERN	37.51	39.05	.466
Accuracy	3.44	4.00	<.01**
HON	5.92	6.03	.775
GQS	2.73	3.00	.162

GQS: Global Quality Scale; HON: Health On the Net.

Development of a systematic evaluation method

During this study, a comprehensive evaluation system was not available. This study, in conjunction with other referenced studies, can merely combine diverse evaluation tools or frameworks to derive at the ultimate outcome. However, each tool or table can only contain a limited number of evaluation dimensions, resulting in a broken and puzzling process. Hence, it is essential to devise a systematic evaluation approach that can ascertain the quality of a wide spectrum of information from various perspectives, with the final score serving as a direct representation of the overall quality. Furthermore, the collaborative endeavors of experts are indispensable in developing assessment instruments that are suitable for easily recognizable content. Research teams with multidisciplinary backgrounds, including dentistry, psychology, sociology, computer science, etc., can be assembled to explore various aspects of the dissemination of TMDs-related information from different perspectives. For examples, dentists, especially specialists dealing with TMDs, can contribute in setting evaluation criteria based on the current best evidence in their fields. The psychologists can study the psychological responses and cognitive biases of the public toward TMDs information, sociologists focus on the impact of social environment and cultural factors on information dissemination, computer scientists are committed to developing more advanced information processing and analysis tools. Together, they can provide comprehensive support and solutions to the evaluation process, and can also improve the quality of TMDs information on social media platforms, thereby facilitating the general populace's

access to accurate information without encountering any difficulties.

Reasons for the poor quality

This study find that the quality of TMDs-related information is generally poor, the possible reasons are as follows:

1. The lack of a systematic TMDs curriculum in dental education means that most dental students and practitioners have not received formal training in TMDs diagnosis and treatment.[12] This situation leads to insufficient knowledge and confidence in managing TMDs problems.
2. For the professionals, the translation of relative knowledge into practice may be influenced by individual factors such as clinical experience and practicing specialty, which can further contribute to the variability in the quality of information.
3. The social media platforms allow the publication of articles by individuals with varying levels of expertise, which can result in the spread of misinformation.

The combination of these factors contributes to the overall poor quality of TMDs-related information available to the public and dental professionals.

Potential directions for future research

Future research could explore the following two directions. Firstly, cross-platform and cross-cultural comparative studies could be conducted. Regarding platforms, the scope of research could be broadened to encompass a wider array of representative social media platforms for comparative analysis. This would facilitate an in-depth examination of the differences and underlying reasons across various dimensions of information quality on different platforms. In terms of culture, researchers could investigate the variations in users' needs and perceptions concerning medical and health information across diverse cultural contexts. Secondly, studies assessing the effectiveness of intervention measures could be undertaken. By monitoring the intervention actions implemented by platforms, researchers could compare the information quality pre- and postintervention, evaluate the tangible impact of each measure, pinpoint the most efficacious intervention strategies, and further refine these approaches to optimize the enhancement of health information quality on social media platforms. Moreover, examining the effects of educational interventions on both authors and the public represents another valuable research idea. Researchers could determine whether such interventions can augment authors' awareness and capability to disseminate high-quality information, as well as bolster the public's proficiency in discerning information quality and accessing reliable sources of information.

Limitation

There are several limitations to this approach as follows:

1. The evaluation process only focuses on a limited number of aspects, such as the uploader category and main content. Other important factors, such as the post date, number of views, and uploading purpose, are not taken into consideration. Some data is not available to this study, making the evaluation process limited and not complete. Limited evaluation dimensions may not provide a perfect comprehensive assessment of the article quality.
2. Secondly, the method employed for readability evaluation is original. However, it is uncertain whether this method is sufficiently reliable and comprehensive enough. Therefore, it is imperative that this method undergoes further refinement and modification. More research and testing should be conducted to enhance its validity and applicability.
3. For the platforms, only WeChat and Zhihu are included in this study, other representative platforms, apps, TV media and paper-based media can be also included for further study.
4. Besides, some medical institutions post articles as a form of self-promotion rather than providing accurate TMD-related information. As such, these articles can be seen as advertisements that should be excluded. However, these articles were still included in this study because they did mention some knowledge of TMDs like treatment methods and treatment cases; these parts were still considered educational.

Conclusion

These articles from Chinese social media, WeChat and Zhihu, received low scores on DISCERN, HONcode, and GQS, indicating a lack of comprehensiveness and depth. However, the information provided was clear, accurate and easy to understand for the public. These findings suggest that the current quality of TMDs-related articles available on Chinese internet platforms is generally poor. Furthermore, there is significant variation in the quality of articles from different authors and publishers. Therefore, the public should be cautious when reading such articles.

ORCID iDs: Yifei Deng https://orcid.org/0009-0004-5906-7565
Jianing Zhou https://orcid.org/0009-0004-5760-5921
Ming Yang https://orcid.org/0009-0001-7819-3151
Xin Xiong https://orcid.org/0000-0003-2175-0970

References

1. Schiffman E, Ohrbach R, Truelove E, et al. Diagnostic criteria for temporomandibular disorders (DC/TMD) for clinical and research applications: recommendations of the international RDC/TMD consortium network* and orofacial pain special interest group†. *J Oral Facial Pain Headache* 2014; 28: 6–27.
2. Xie C, Lin M, Yang H, et al. Prevalence of temporomandibular disorders and its clinical signs in Chinese students, 1979–2017: a systematic review and meta-analysis. *Oral Dis* 2019; 25: 1697–1706.
3. List T and Jensen RH. Temporomandibular disorders: old ideas and new concepts. *Cephalalgia* 2017; 37: 692–704.
4. Liu F and Steinkeler A. Epidemiology, diagnosis, and treatment of temporomandibular disorders. *Dent Clin North Am* 2013; 57: 465–479.
5. Shrivastava M, Battaglino R and Ye L. A comprehensive review on biomarkers associated with painful temporomandibular disorders. *Int J Oral Sci* 2021; 13: 23.
6. Aggarwal VR, McBeth J, Zakrzewska JM, et al. The epidemiology of chronic syndromes that are frequently unexplained: do they have common associated factors? *Int J Epidemiol* 2006; 35: 468–476.
7. Longley SL, Broman-Fulks JJ, Calamari JE, et al. A taxometric study of hypochondriasis symptoms. *Behav Ther* 2010; 41: 505–514.
8. Tyrer P. Recent advances in the understanding and treatment of health anxiety. *Curr Psychiatry Rep* 2018; 20: 49.
9. Sun W, Fang S, Wang Y, et al. Prevalence and associated factors of health anxiety in patients with temporomandibular disorders. *Oral Dis* 2024; 31: 181–192.
10. Mohan G, Joseph AM, Philip CK, et al. Psychological and physiological relaxation using yoga in patients undergoing dental implant surgery: a pilot randomized clinical trial with biomarker assessment. *Int J Yoga Therap* 2022; 32: Article 14. DOI: 10.17761/2022-d-21-00056
11. Wu J, Huang Z, Chen Y, et al. Temporomandibular disorders among medical students in China: prevalence, biological and psychological risk factors. *BMC oral Health* 2021; 21: 549.
12. Xiong X, Xiao C, Zhou X, et al. Knowledge and attitudes regarding temporomandibular disorders among postgraduate dental students and practicing dentists in western China: a questionnaire-based observational investigation. *Pain Res Manag* 2023; 2023: 7886248.
13. Zhang X, Wen D, Liang J, et al. How the public uses social media WeChat to obtain health information in China: a survey study. *BMC Med Inform Decis Mak* 2017; 17: 66.
14. Wang Q, Jia J, Zhou C, et al. A bibliometric analysis of research on temporomandibular joint disc displacement from 1992 to 2022. *Healthcare (Basel, Switzerland)* 2023; 11. DOI: 10.3390/healthcare11142108
15. Wadhokar OC, Kulkarni CA and Joshi M. Changing trends in publication regarding temporomandibular joint dysfunction from 1975 to 2021: a bibliometric analysis. *Cureus* 2023; 15: e46179.
16. Pan P, Yu C, Li T, et al. Evaluating the quality of cancer-related WeChat public accounts: cross-sectional study. *JMIR cancer* 2024; 10: e52156.
17. Yang W, Li B, Liu M, et al. Quality evaluation of health information about breast cancer treatment found on WeChat public accounts. *Arch Public Health* 2023; 81: 170.
18. Charnock DJQcfchiotcRUoO and Library TB. The DISCERN handbook. 1998: 7-51.
19. Davaris MT, Dowsey MM, Bunzli S, et al. Arthroplasty information on the internet: quality or quantity? *Bone Joint Open* 2020; 1: 64–73.
20. Gauer RL and Semidey MJ. Diagnosis and treatment of temporomandibular disorders. *Am Fam Physician* 2015; 91: 378–386.
21. Okeson JJSLE. Management of temporomandibular disorders and occlusion: Jeffrey P. Okeson. 2019.
22. Nelson SJ. *Wheeler's dental anatomy, physiology and occlusion-e-book*. Elsevier Health Sciences, 2019.
23. Yost O, Liverman CT, English R, et al. Temporomandibular disorders: priorities for research and care. 2020.
24. Boyer C, Selby M and Appel RD. The health on the net code of conduct for medical and health web sites. *Stud Health Technol Inform* 1998; 52: 1163–1166.
25. Bernard A, Langille M, Hughes S, et al. A systematic review of patient inflammatory bowel disease information resources on the world wide web. *Am J Gastroenterol* 2007; 102: 2070–2077.
26. Azer SA and Azer S. Bibliometric analysis of the top-cited gastroenterology and hepatology articles. *BMJ open* 2016; 6: e009889.
27. Chong YM, Fraval A, Chandrananth J, et al. Assessment of the quality of web-based information on bunions. *Foot Ankle Int* 2013; 34: 1134–1139.
28. Jensen JH and Helpern JA. MRI quantification of non-Gaussian water diffusion by kurtosis analysis. *NMR Biomed* 2010; 23: 698–710.
29. Yu Z, Guindani M, Grieco SF, et al. Beyond t test and ANOVA: applications of mixed-effects models for more rigorous statistical analysis in neuroscience research. *Neuron* 2022; 110: 21–35.
30. Churchill GA. Using ANOVA to analyze microarray data. *BioTechniques* 2004; 37: 173–175, 177.
31. Spiers H, Amin N, Lakhani R, et al. Assessing readability and reliability of online patient information regarding vestibular schwannoma. *Otology & Neurotology: Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology* 2017; 38: e470–e475.
32. <DISCERN - The DISCERN Instrument.pdf>.

33. Yang Y, Hou M, Gong X, et al. Quality assessment of hypertension treatment-related information on WeChat: cross-sectional study. *J Med Internet Res* 2022; 24: e38567.

34. Starman JS, Gettys FK, Capo JA, et al. Quality and content of internet-based information for ten common orthopaedic sports medicine diagnoses. *The Journal of Bone and Joint Surgery American Volume* 2010; 92: 1612–1618.

35. Kim MJ, Kim JR, Jo JH, et al. Temporomandibular disorders-related videos on YouTube are unreliable sources of medical information: a cross-sectional analysis of quality and content. *Digit Health* 2023; 9: 20552076231154377. DOI: 10.1177/20552076231154377

36. Zuidema WP, Graumans MJ, Oosterhuis JWA, et al. The quality of web sites' health information on minimal invasive repair of pectus excavatum using the DISCERN instrument. *Eur J Pediatr Surg* 2021; 31: 157–163.

37. Hong SW, Kang JH, Park JH, et al. Quality and readability of online information on hand osteoarthritis. *Health Informatics J* 2023; 29, 14604582231169297. DOI: 10.1177/14604582231169297