

Emotions and Appeals: An Analysis of Online Comments on HPV Vaccine under Social Media Publicity

— A Case Study of the Weibo Account of "People's Daily"

Yinqi Zhang^{1,a,*}

¹*Department of Journalism and Communication, South China University of Technology,
Guangzhou, 511431, China*

a. yeahzyq@163.com

**corresponding author*

Abstract: In 2018, the human papillomavirus vaccine (HPV) was approved for sale in China, and the topic of HPV vaccination has attracted the attention of the general public. Vaccines are biological agents made of bacteria, viruses, tumor cells, etc., which can make the body produce specific immunity, and are the most effective measures to control infectious diseases and prevent infection with viruses. Vaccination is of great significance for protecting people's health, maintaining social stability and national security. Media content and comments on social media provide an important basis for understanding public opinion and people's demands. In this study, data from Weibo platform were used, "HPV" was searched as a keyword under the official Weibo account of People's Daily, and relevant information was collected and analyzed with python. The results of this public opinion mining play a crucial role in discovering the problems existing in HPV vaccination in China, and also provide important ideas for the government to solve this problem.

Keywords: HPV, social media, data mining, sentiment analysis, use and gratification theory

1. Introduction

At present, the discussion about vaccination is endless on social media. Under the publicity and popularization of HPV vaccine, more and more people have accepted HPV vaccine. In the past 10 years, a large number of vaccine sentiment studies have also appeared [1]. In foreign studies, some scholars have paid attention to the negative emotional expression of vaccination [2], and some scholars have studied the different reactions of general audiences to news under different frameworks [3].

However, the situation of HPV vaccination in foreign countries is different from that in China. The premise of foreign research is that there is sufficient vaccine supply, but China's nine-valent HPV vaccine faces the problem of short supply. Therefore, this study tends to discuss the localization of HPV vaccination sentiment.

In the academic research in China, the attention on HPV vaccine has made considerable progress in the past three years, and the research focus is more focused on the research on vaccination intention and vaccine popularization [4]. However, the research is not in-depth enough, there is a lack of

representative core journal papers, the research objects are mostly gathered in the general public such as women and college students [5], and the research on public institutions such as the government is insufficient.

Therefore, this study focuses on the official Weibo account of "People's Daily", classifies people's vaccine emotions and conducts word frequency analysis to obtain the information that people are most concerned about, laying a foundation for relevant departments to make reasonable use of information and actively respond to policy news and public emotions, so as to improve the quality of social media information service, which has substantial research significance and practical significance.

2. Theoretical framework and research questions

2.1. Public opinion

Public opinion is one of the main opinions. Major public opinion events cover almost every aspect of social life. Song Yuchao and Chen.J categorize the important ideas available online. According to the research of Song Yuchao and Chen.J., online public opinion can be divided into politics, economy, culture, society, accidents and disasters. Online public opinion topics can be divided into seven categories: government management, such as anti-corruption, information transparency, etc.; Social livelihood, such as health care, employment, social security, etc.; Defense diplomacy, such as foreign relations, ethnic and religious; Education and teaching, such as educational equity, academic integrity, etc. Corporate finance, such as employment conflicts and price fluctuations; Accidents and disasters, such as work safety, natural disasters, traffic accidents, etc.; Cultural construction, such as sports events, celebrations, etc. [6]. The discussion of the medical aspects associated with the HPV vaccine is an important part of public opinion.

2.2. Use and Gratification theory

Based on certain social and psychological needs, individuals and groups have expectations for mass media or other sources, which leads to different modes of media contact or participation in other activities [7]. In this study, social media users expressed their thoughts in the comments section. the key point is, active audiences have their own demands, they are not just simple recipients, they are more disseminators to a certain extent, they seek social media to express their views. Users who comment on social media have utility, intentionality and imperviousness to influence, making their own demands in the hope of pushing for an official solution.

2.3. Major public opinion analysis framework

Time trend analysis and text mining are effective methods to analyze public opinion.

The first thing is to analyze the time trend of public opinion. Through trend analysis, we can see the changes in the level of public attention to a certain topic. Through text analysis, we can see more deeply what people are thinking and talking about at different stages [8].

Based on this theory, the following research questions are proposed:

RQ1: What are the Baidu search and attention trends for HPV vaccine from April 2018 to June 2023?

RQ2: Is the peak of HPV vaccine searches in the past five years related to the People's Daily publication?

RQ3: Are there any differences among the focus on HPV bivalent, quadrivalent and nine-valent vaccines?

Through text mining, we can discover how the public feels about a particular social issue. Digging

into the text of Internet users and media reports can give an in-depth understanding of the concepts and thoughts of public opinion. Text is the product of public opinion expression and the way to understand public opinion.

For the public, netizens use social media to express a social problem [8]. Since the comments on mainstream media account reflect the way people think about an issue, it is of great significance to analyze media texts on a social topic.

Based on this view, the following research questions are proposed:

RQ4: What are people's emotions about HPV vaccination?

RQ5: What are the biggest concerns about the HPV vaccine?

3. Research method

This study uses "HPV" as the keyword and the time span is from April 2018 to June 2023:

Analysis of Internet users' attention trend: Search keywords using Baidu index, limit the time span, analyze the Internet users' attention trend of HPV vaccine.

Text analysis of netizens' opinions: Select the Weibo account of People's Daily, use python to capture posts and comments data, and use python to conduct emotion analysis to obtain emotional tendencies; After cleaning the data, python was used for word frequency analysis and word cloud generation.

3.1. Data collection

Baidu Index is a data sharing platform based on Baidu's massive Internet user behavior (search, click, share) data. It is one of the most important statistical analysis platforms in the current Internet and even the whole data era. The search index is based on the search volume of netizens in Baidu, and the keyword is taken as the statistical object to analyze and calculate the search frequency of each keyword in Baidu web search. In the online public opinion research conducted by Chinese scholars, Baidu Index is often used to study the degree of netizens' attention to a certain phenomenon or a certain keyword. For example, scholars Zheng Jietao, Yuan Yifan and Li Li used Baidu Index to study the search trend of monkeypox epidemic in China [9].

Weibo, China's largest social media platform, allows users to post short articles, pictures, videos and other content, interact and share with other users. Weibo has a huge user base, including ordinary users, celebrities, media organizations and so on and users can follow other users to get the content they post, and can comment, like, forward and other operations on Weibo. Therefore, Weibo is an ideal platform for public opinion mining on social media.

In this study, the Weibo platform was used to study the "HPV vaccine" related content published by the official account of People's Daily, and python was used to capture posts and comments.

3.2. data preprocessing

Before conducting the analysis procedures, we preprocessed the text. The steps we performed on the database were as follows:

- tokenization of the text that involves the division into words of the text itself;
- the removal of stopwords defined as words that do not carry significant information for the analysis;
- stemming and rooting, which consist of reducing words to the root;
- the removal of extra numbers and spaces;
- the removal of punctuation, etc. [10].

3.3. data analysis

In this study, Baidu index was used to generate a time line of search volume to more intuitively show the concern trend of Internet users on HPV vaccine; python was used for data capture and data cleaning on the Weibo web page, emotion dictionary developed by Dalian University of Technology was used for emotion analysis, some errors were manually corrected, and word frequency and word cloud were generated by python, which indicated the most concerned problems of Internet users in the discussion of HPV vaccine.

Sentiment analysis is a branch of computational linguistics that focuses on determining the opinions expressed in a written document. It aims to identify how emotions are conveyed in the text and whether they indicate a positive or negative viewpoint. To accomplish this, sentiment analysis relies on a predetermined lexicon that assigns scores to words. When analyzing a text, each relevant word is assigned a score based on its proximity to positive or negative words in the lexicon. Therefore, it is crucial to have a lexicon that assigns a polarity (positive or negative) and a score to each word [10]. In this study, refer to the research of Ruan D Nong and other scholars in China, sentiment analysis was conducted using the Chinese emotion dictionary developed by Dalian University of Technology [11].

4. Result

4.1. Baidu Index Timeline

People's attention to HPV vaccine has shown an overall growth trend in the past five years, and the change of search volume has a certain periodicity (Figure 1). From 2018 to 2023, the number of Internet users searching for HPV on Baidu has increased significantly, indicating that people are more and more concerned about their own health and HPV vaccine vaccination. However, with the popularity of HPV vaccine and the increase of vaccination rate, the number of HPV Baidu searches has decreased significantly since 2023.

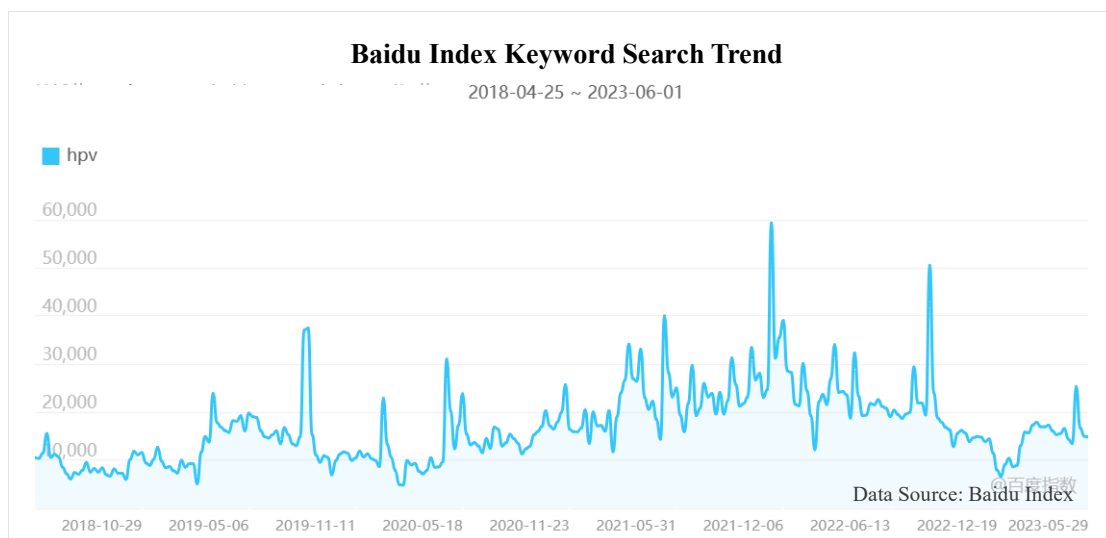


Figure 1: Baidu Index Timeline

4.2. The correlation between the peak of Baidu Index and the content published by People's Daily

The peak of people's attention to HPV has little to do with the time when "People's Daily" publishes

HPV topics, but the amount of discussion and attention will be greater in the years when "People's Daily" publishes content intensively (Figure 2).

In the figure below, the red dot is the approximate time when the HPV vaccine content was released by the Weibo account of "People's Daily", and the closer time is uniformly represented by a dot. It can be seen that people's search volume for HPV vaccine does not change with the publication content of People's Daily to a large extent, but in the year when People's Times frequently publishes HPV-related news (2021), people's attention to HPV does increase on the whole.

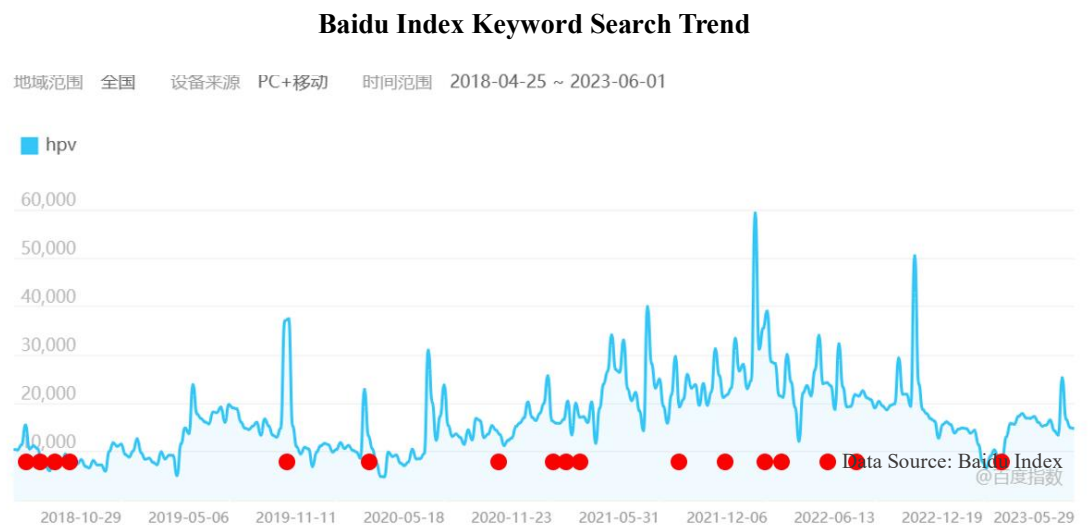


Figure 2: Search Volume and Posting Nodes

4.3. Differences in the level of concern among different types of vaccines

The nine-valent HPV vaccine has received the most attention, much more than the bivalent and quadrivalent vaccines (Figure 3).

HPV vaccines are divided into bivalent, quadrivalent and nine-valent, and there are differences between them in the type and scope of HPV virus prevention, of which the nine-valent HPV vaccine can prevent the most virus types. Due to the late launch time of the nine-valent vaccine, good preventive effect, and small supply, it has received the largest attention, far exceeding the bivalent vaccine and quadrivalent vaccine, and the search volume and attention degree of the bivalent vaccine are roughly the same.

There was no search volume for the HPV nine-valent vaccine until 2020, because the HPV nine-valent vaccine was not administered in bulk in China until the end of 2019.

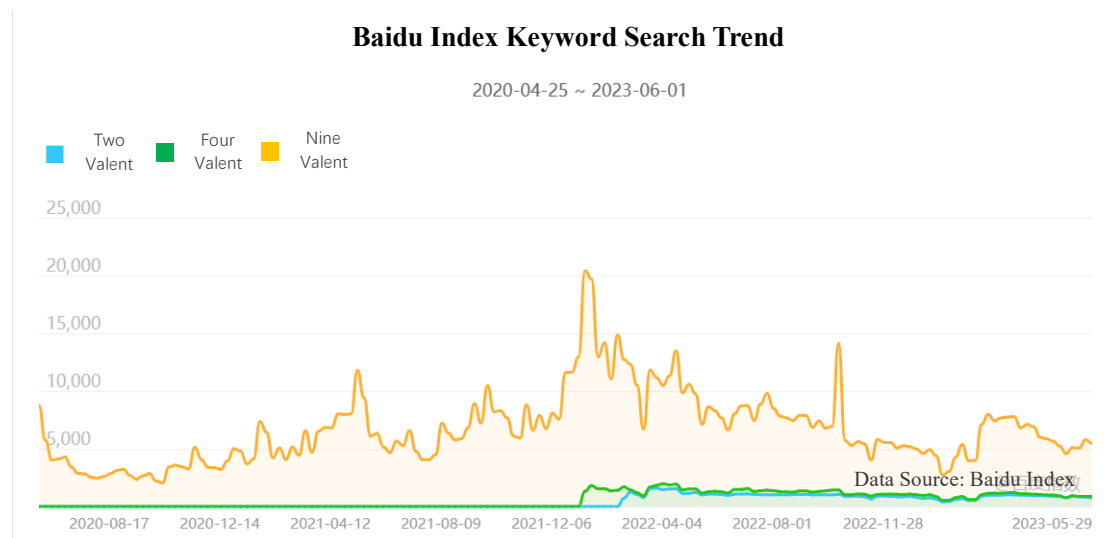


Figure 3: Search Comparison of Different Categories

4.4. Sentiment Pie Chart

The vast majority of Chinese netizens are pessimistic about HPV vaccination in China, and "sad" is the main sentiment (Figure 4).

After deleting meaningless content such as facial expressions and emojis, 23,605 pieces of valid data were obtained. These data were imported into the sentiment dictionary of Dalian University of Technology for emotion analysis. 54% of Internet users expressed sadness, and the typical keywords were "unable to make an appointment" and "unable to vaccinate"; 33% of people showed no obvious emotion, and the representative comment was "on which platform can I book an appointment"; 7% of users expressed anger over this issue, with representative comments such as "If I do not use technology to cheat, I will not be able to get an appointment when I am older than the age to vaccinate"; Only 4% of users were happy, mostly because of appointments or successful vaccinations.

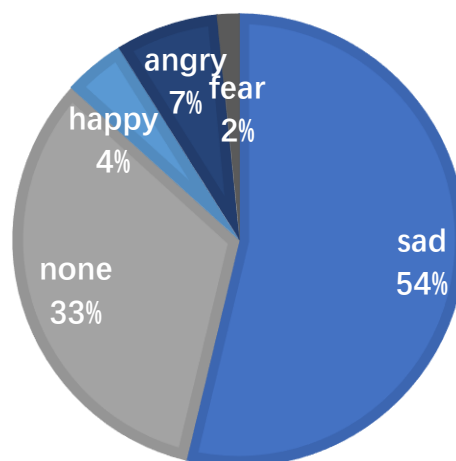


Figure 4: Sentiment Pie Chart

4.5. Weibo based Word Cloud

In the issue of HPV vaccination, Chinese people are most concerned about the supply of vaccines, and pay more attention to the nine-valent vaccines than the two-valent and four-valent ones (Figure

5).

The high-frequency word "appointment" appeared 5437 times, and "no" appeared 2931 times. From the point of view of the content, people are more concerned about the problem of "appointment is not available" and "no vaccine", which is also the main appeal of people's voice on the HPV issue through the Internet social platform. Therefore, if the government wants to obtain more positive evaluation, the first task is to increase the supply of nine-valent HPV vaccine and prioritize the vaccination needs of the appropriate age group.



Figure 5: Weibo based Word Cloud

5. Conclusion

With the development of medical science and the advancement of a series of public health events, people are paying more and more attention to their health problems. In the topic of HPV vaccine, although people's attention does not fluctuate with each publication of People's Daily, it is affected by the frequency of publication of People's Daily. At present, Internet users pay more attention to the nine-valent vaccine than other HPV vaccines, and the inadequate supply of nine-valent vaccine in mainland China and the high frequency of official publicity have aggravated people's sadness about the vaccine. This kind of vaccine sadness is not the resistance to vaccines, but the hope of being able to vaccinate, but encountered the realistic dilemma of being unable to make an appointment. Such sentiments, reflected in a large number of People's Daily comments on Weibo, have also encouraged more people who could not get an appointment for the nine-valent vaccine to speak out, hoping to draw the government's attention to meet the demand for vaccines.

6. Discussion

The advent of the era of big data and the progress of technology provide people the opportunity to use the Internet to express opinions, and also make it possible to mine online public opinion, helping relevant institutions to make decisions against public opinion. Based on the topic of HPV, this study took the official Weibo account of People's Daily as a case, mined the content and comments about HPV vaccine published by the account, and conducted sentiment analysis and word frequency analysis of public opinion to obtain the issues that people are most concerned about. This study fills the gap in the research of HPV vaccine in China in the field of social media, innovatively adopts sentiment analysis and word frequency analysis to study people's most important needs on this issue, and provides realistic basis and data support for the subsequent decision-making of the government

and relevant institutions and is of practical significance and academic significance.

However, subject to the existing technical and cultural barriers, there are still the following problems to be solved and improved:

1. In the data mining stage, due to the anti-mining mechanism of Weibo, some comments, especially secondary comments, were not completely collected, and about 2000 comments were not collected, which may lead to errors between the conclusion of this study and the actual situation. At present, no perfect method has been found to solve this problem, waiting for follow-up researchers to supplement.

2. In the stage of data preprocessing, due to the particularity of Chinese expression and the particularity of the topic, some illogical combinatorial words are retained in the stage of removing stop-words, which are frequently used in comments, and their deletion will cause errors in the research. However, due to the manual processing of more than 23,000 data, it is inevitable that there will be omissions and insufficient consideration, which is also the regret left by this study.

3. In the data analysis stage, due to the existence of some obscure expressions in Chinese expressions, such as positive words combined with some specific expressions will become sarcasm, which also brings certain difficulties to Chinese sentiment analysis. Although manual correction has been made and the new code has been handed over to machine learning, due to the large amount of data, every comment cannot be completely corrected. This also brings some research errors.

HPV vaccine is an important medical topic related to people's livelihood. With the development of economy, people are paying more and more attention to their own health. By analyzing the attitude of Internet users towards HPV vaccine and collecting statistics on the high-frequency words spoken online, we can better identify the difficulties and problems faced by China in the process of promoting the popularization of HPV vaccine, so that the government can make better decisions and create a better life for the people.

References

- [1] Xiao Luo, Gregory Zimet, and Setu Shah. *HUMAN VACCINES & IMMUNOTHERAPEUTICS. A natural language processing framework to analyse the opinions on HPV vaccination reflected in twitter over 10 years (2008 - 2017)*. 2019, VOL. 15, NOS. 7-8, 1496-1504
- [2] Zhou X, Coiera E, Tsafnat G, Arachi D, Ong MS, Dunn AG. *Using social connection information to improve opinion mining: Identifying negative sentiment about HPV vaccines on Twitter*. *Stud Health Technol Inform*. 2015;216:761-5. PMID: 26262154.
- [3] MacKay M, Cimino A, Yousefinaghani S, McWhirter JE, Dara R, Papadopoulos A. *Canadian COVID-19 Crisis Communication on Twitter: Mixed Methods Research Examining Tweets from Government, Politicians, and Public Health for Crisis Communication Guiding Principles and Tweet Engagement*. *Int J Environ Res Public Health*. 2022 Jun 6;19(11):6954. doi: 10.3390/ijerph19116954. PMID: 35682537; PMCID: PMC9180105.
- [4] Li T, Chen C & Chen Y. (2021). *Research status of HPV vaccine under the threshold of health communication: A review of relevant research at home and abroad*. *Science and Technology Communication* (09),24-27. doi:10.16607/j.cnki.1674-6708.2021.09.012.
- [5] Hsu, P., & Rui, (2021). *Risk Aversion or Compliance? -- An empirical study on the influence of social media on the intention of female college students to receive HPV vaccine*. *Chinese Internet Communication Research* (04),145-165.
- [6] Song Yuchao and Chen. J. *Intelligence Magazine, Research on the Formation Mechanism of Online Public Opinion Game for Subject Classification—Based on the Research of Three Subjects*. 2015. 34 (8): p. 122-126.
- [7] Richard West and Lynn H. Turner: *Introduction to Communication Theory*. P331-335.
- [8] Lian Zhixuan, et al., *Research on the Multidimensional Dynamic Classification and Prediction Model of Internet Public Opinion for Big Data*. 2018(5): p. 20.
- [9] Zheng J T, Yuan Y F & Li L. (2023). *Research on global monkeypox epidemic network public opinion in 2022 based on Baidu Index*. *Modern Medicine and Hygiene* (10),1626-1629.
- [10] Barbierato, E., Bernetti, I., & Capecechi, I. (2021). *Analyzing TripAdvisor reviews of wine tours: an approach based on text mining and sentiment analysis*. *International Journal of Wine Business Research*.
- [11] Ruan D Nong, Ma H J & Wang Y F. (2023). *Agenda setting and emotion analysis of cancer trending on Weibo*. *Hospital Management Forum* (04),35-40+74.