

Impact Analysis of COVID-19 Pandemic on Hospital Reviews on "Dianping" Website in Shanghai and China: Empirical Study

Weixue Huo, Mengwei He, Zhaoxiang Zeng, Xianhao Bao, Ye Lu, Wen Tian, Jiaxuan Feng, Rui Feng

Submitted to: Journal of Medical Internet Research
on: September 22, 2023

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript..... 5

Supplementary Files..... 26

 Figures 27

 Figure 1..... 28

 Figure 2..... 29

 Figure 3..... 30

 Figure 4..... 31

 Figure 5..... 32

 Figure 6..... 33

 Multimedia Appendixes 34

 Multimedia Appendix 1..... 35

 Multimedia Appendix 2..... 35

Impact Analysis of COVID-19 Pandemic on Hospital Reviews on "Dianping" Website in Shanghai and China: Empirical Study

Weixue Huo^{1*} Dr med; Mengwei He^{1*} Dr med; Zhaoxiang Zeng² Dr med; Xianhao Bao¹ Dr med; Ye Lu¹ Dr med; Wen Tian¹ Dr med; Jiaxuan Feng³ Dr med; Rui Feng¹ Dr med

¹Department of Vascular Surgery?Shanghai General Hospital, Shanghai Jiaotong University Shanghai CN

²Department of Vascular Surgery, Changhai Hospital, Navy Medical University, Shanghai, P. R. China. Shanghai CN

³Department of Vascular Surgery?Shanghai General Hospital, Shanghai Jiaotong University Department of Vascular Surgery, Changhai Hospital, Navy Medical University, Shanghai, P. R. China. Shanghai CN

*these authors contributed equally

Corresponding Author:

Rui Feng Dr med

Department of Vascular Surgery?Shanghai General Hospital, Shanghai Jiaotong University

650 New Song Jiang Road, Shanghai 200433, P.R. China

Shanghai

CN

Abstract

Background: In the era of the Internet, individuals have increasingly accustomed themselves to gathering necessary information and expressing their opinions on public online platforms. The healthcare sector is no exception, as these evaluations to a certain extent influence people's healthcare decisions. In China, the widespread outbreak of COVID-19 has had an impact on the strained doctor-patient relationships and patients' healthcare-seeking behaviors. This, in turn, raises the question of how patient evaluations of hospitals have undergone transformations as a result.

Objective: To investigate the variances in patient evaluations received by hospitals during different periods and across different types of healthcare facilities amidst the COVID-19 pandemic. Utilizing ChatGPT technology, negative comments pertaining to hospitals were analyzed, and potential actionable measures were identified to enhance patient satisfaction. This study aims to provide insights for hospital administrators when dealing with unforeseen public health crises.

Methods: A Java-based program was developed to gather patient evaluations from Dianping, a popular consumer review website. Evaluations were collected from the top 50 comprehensive hospitals nationwide, as well as nominated specialized hospitals in pediatric care, tumor, and obstetrics and gynecology. Statistical analysis of the rating scores was conducted using SPSS. Furthermore, ChatGPT was utilized to categorize the content of the evaluations.

Results: A total of 30,317 valid evaluation records were collected, including 7,696 negative evaluations. Analysis of these data revealed a significant impact of the pandemic on hospital evaluation scores. Overall, there was a significant increase in average evaluation scores during the outbreak ($P < .001$). Furthermore, there were notable differences in the composition of negative comments among different types of hospitals ($P < .001$). Children's hospitals received sensitive feedback regarding waiting times and treatment effectiveness, while patients at maternity hospitals showed a greater concern for the attitude of healthcare providers. Patients at tumor hospitals expressed a desire for timely examinations and treatments, especially during the pandemic period.

Conclusions: The COVID-19 pandemic has a significant impact on patient evaluation scores. There are variations in the scores and content of evaluations among different types of specialized hospitals. Utilizing ChatGPT to analyze patient evaluation content represents an innovative approach for statistically assessing factors contributing to patient dissatisfaction. The findings of this study could provide valuable insights for hospital administrators to foster more harmonious physician-patient relationships and enhance hospital performance during public health emergencies.

(JMIR Preprints 22/09/2023:52992)

DOI: <https://doi.org/10.2196/preprints.52992>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to all users.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <http://www.jmir.org/>

Original Manuscript

Impact Analysis of COVID-19 Pandemic on Hospital Reviews on "Dianping" Website in Shanghai and China: Empirical Study

Abstract

Background: In the era of the Internet, individuals have increasingly accustomed themselves to gathering necessary information and expressing their opinions on public online platforms. The healthcare sector is no exception, as these comments to a certain extent influence people's healthcare decisions. During the onset of the COVID-19 pandemic, how the medical experience of Chinese patients and their evaluations of hospitals have changed remains to be studied. Therefore, we plan to collect patient medical visit data from the internet to reflect the current status of medical relationships under specific circumstances.

Objective: The aim of this study is to explore the differences in patient comments across various stages (during, before, and after) of the COVID-19 pandemic, as well as among different types of hospitals (Children's Hospital, Maternity Hospital, Tumor Hospital). Additionally, by leveraging ChatGPT, the study categorizes the elements of negative hospital evaluations. Analysis is conducted on the acquired data, and potential solutions that could improve patient satisfaction are proposed. This study is intended to assist hospital managers in providing a better experience for patients who are seeking care amidst an emergent public health crisis.

Methods: Selecting the top 50 comprehensive hospitals nationwide and the top specialized hospitals (children's hospitals, tumor hospitals, and maternity hospitals), we collect patient reviews from these hospitals on the Dianping website. Utilizing ChatGPT, we classify the content of negative reviews. Additionally, we conduct statistical analysis using SPSS to examine the scoring and composition of negative evaluations.

Results: A total of 30,317 pieces of effective comment information were collected from January 1, 2018 to August 15, 2023, including 7,696 pieces of negative comment information. Manual inspection results indicate that ChatGPT has an accuracy rate of 92.05%. The F1 score is 0.914. The analysis of this data revealed a significant correlation between the comments and ratings received by hospitals during the pandemic. Overall, there was a significant increase in average comment scores during the outbreak ($P < .001$). Furthermore, there were notable differences in the composition of negative comments among different types of hospitals ($P < .001$). Children's hospitals received sensitive feedback regarding waiting times and treatment effectiveness, while patients at maternity hospitals showed a greater concern for the attitude of healthcare providers. Patients at tumor hospitals expressed a desire for timely examinations and treatments, especially during the pandemic period.

Conclusion: The COVID-19 pandemic has some association with patient comment scores. There are variations in the scores and content of comments among different types of specialized hospitals. Utilizing ChatGPT to analyze patient comment content represents an innovative approach for statistically assessing factors contributing to patient dissatisfaction. The findings of this study could provide valuable insights for hospital administrators to foster more harmonious physician-patient relationships and enhance hospital performance during public health emergencies.

Key Words

patient satisfaction; physician-patient relationship; ChatGPT; patient complaint.

Introduction

The tense relationship between Chinese patients and healthcare service providers.

In recent years, medical issues have been a highly focused and hot topic in China.[1, 2] In 2016, 2017, and 2019, China experienced severe conflicts between patients and healthcare providers, including extreme cases of violence where doctors were harmed or even killed by patients or their family members.[3, 4] The occurrence of this series of events highlights the challenging state of patients-healthcare service providers relationships in China. Moreover, the significance of these relationships cannot be underestimated. On the one hand, doctor-patient relationships have a profound impact. They influence patients' satisfaction with healthcare services and their understanding of medical information. Additionally, these relationships affect patients' behavior in response to illness and their adherence to prescribed medication, ultimately impacting their overall quality of life and health outcomes.[5]

Therefore, it is important for us to improve the current state of patients-healthcare service providers relationships. Against this backdrop, the COVID-19 pandemic has further exacerbated the complexity of this issue. [6] The outbreak of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), causing COVID-19, has been an unprecedented event affecting all sectors of society, particularly the healthcare sector. The spread of the pandemic has resulted in a strained allocation of medical resources, potentially leading to shortages in certain healthcare services. Hospitals are required to prioritize patients based on varying levels of urgency and may need to implement isolation measures. These measures have the potential to impact the mental well-being of both healthcare workers and patients to varying degrees.[7, 8] These factors could potentially lead to an even more strained relationship between Chinese patients and healthcare service providers.

Web based comments and ChatGPT

In the era of the Internet, individuals are becoming increasingly accustomed to expressing their viewpoints on public online platforms and utilizing public comments as a means to inform their own rational decision-making processes.[9, 10] The field of healthcare is no exception, as previous research on quality management has indicated that due to factors such as complex complaint processes and lack of transparency in the resolution, only a third of patients opt to make a formal complaint when they encounter unsatisfactory service. [11-13] However, with the increasing prevalence of internet technology, public online platforms have provided patients with a more efficient platform to express their personal opinions. When faced with unsatisfactory service, the option of online reviews provides an alternative approach for patients to address their concerns. Through internet platforms, they can discuss their negative medical experiences and share them with others. [14] Research has demonstrated that the public dissemination of negative reviews has the potential to undermine or tarnish the image of physicians and amplify patient dissatisfaction with doctors or hospitals.[15] Undoubtedly, these comments also have the potential to exert an underlying influence on subsequent patients seeking medical care. Hence, the identification and documentation of negative patient comments, followed by targeted improvements, are of paramount importance in enhancing the quality of healthcare services and fostering positive physician-patient relationships. [16] Over the past year, considerable attention has been drawn to a novel AI model called ChatGPT due to its remarkable proficiency in accomplishing varied natural language tasks.[17] [18-21] ChatGPT is a general large language model (LLM) recently developed by OpenAI. Many studies have shown that[17, 22-25] ChatGPT has demonstrated impressive analytical and observational abilities, approaching or surpassing the threshold required to simulate the United States Medical

Licensing Examination (USMLE) multiple-choice questions. Leveraging the analytical capabilities of ChatGPT, this study aims to conduct content analysis and organization of publicly available negative reviews to explore the specific reasons and composition of patients' negative feedback.

Current research findings indicate that during the COVID-19 pandemic, the rating of doctor-patient relationships among patients (4.18 ± 0.51) has increased compared to the pre-pandemic period (3.86 ± 0.67), suggesting an improvement in the doctor-patient relationship. However, this study did not differentiate between specific types of hospitals. General hospitals and specialized hospitals may exhibit differences in doctor-patient relationships due to variances in the types of patients they treat. Additionally, when collecting rating data prior to the pandemic, this study relied on participants' retrospective memory for scoring, which could introduce certain biases and deviate from the actual circumstances. Moreover, a patient's experience during a hospital visit is composed of multiple factors; in addition to doctor-patient relationships, the quality of service provided by nurses and other hospital staff should also be taken into consideration. Even the cleanliness of the hospital environment and the quality of service offered by automated service equipment can affect the patient's experience during a hospital visit.[26, 27]

This study goes beyond the conventional doctor-patient relationship. Instead, it starts by exploring and analyzing the patients' public comments of the hospitals. By examining the assessments published on public platforms, we aim to investigate the impact of COVID-19 factors and hospital categorization differences on the healthcare experiences of patients in China, taking into account the perspectives of the patients. The research questions posed in this survey encompass the following aspects:

RQ1: During the pandemic, how have patient ratings for hospitals changed?

RQ2: Since the outbreak of the pandemic, what changes have occurred in the composition of negative reviews from patients about hospitals?

RQ3: What kind of negative feedback have different types of hospitals—such as children's, maternity, and tumor hospitals—received, and what are the characteristics of the negative reviews for these various institutions?

Methods

Sample and data collection

This study relies on the open business review platform "Dianping" website.[28] Established in Shanghai in April 2003, this platform is the earliest independent third-party review website globally, boasting an active user base of 40.987 million people. The sample for this study consists of a total of 88 hospitals, including those ranked in the top 50 nationally by Fudan University and those nominated in the Fudan University Specialist Hospital rankings for children's hospitals, maternity hospitals, and tumor hospitals. The collection of comment information was limited to the period from January 1, 2018, to August 15, 2023. The collected content includes the hospital's region, type, name, date, score (ranging from 1 to 5), content, and user ID.(Figure1 A) In order to ensure patient privacy, the user ID is only used to judge whether the comment information meets the requirements, and will not appear in this article.



Figure 1. A: Negative Comment Samples from 'Dianping' Website.; B: ChatGPT Handling Demonstration.

Comment inclusion criteria

To ensure the quality of the included comment information, it is necessary to conduct preliminary screening, labeling, and exclusion of meaningless comment content. Comment information with a total character count less than 10 and comment information with fewer than 10 different Chinese characters will be directly deleted. Comment information corresponding to user IDs that appear more than five times will undergo manual review to confirm their authenticity and validity before being included.

Ethical considerations

The data collected for this study consists of de-identified information that has been publicly released online, adhering to ethical standards while fully respecting and safeguarding the rights of all

participants. Furthermore, this study has undergone a rigorous review and received approval from the Ethics Committee of Shanghai General Hospital.(20240417094101248)

Processing and analysis of comment scores

Based on the comment dates, the comment information is categorized as follows: a. Before the pandemic (January 1, 2018 - April 30, 2018), b. During the pandemic (January 1, 2020 - April 30, 2020), c. After the pandemic (January 1, 2023 - April 30, 2023). The Kruskal-Wallis test is employed to compare the comment score data among the three categories: a. Before the pandemic, b. During the pandemic, c. After the pandemic. Subsequently, pairwise Mann-Whitney U tests are conducted to compare the three groups, and the statistical analysis results are recorded. The significance level is set at $P < .05$ (two-tailed).

Categorizing the comment information based on hospital location: a. Shanghai region, b. Other regions. A pairwise Mann-Whitney U test is employed to compare the comment score data between hospitals in the Shanghai region and hospitals in other regions, and the statistical analysis results are recorded. The significance level is set at $P < .05$ (two-tailed).

Categorizing the comment information based on hospital type: a. General hospitals, b. Children's hospitals, c. Maternity hospitals, d. Tumor hospitals. A Kruskal-Wallis test is employed to compare the comment score data among children's hospitals, maternity hospitals, and tumor hospitals. Subsequently, pairwise Mann-Whitney U tests are conducted to compare the three groups, and the statistical analysis results are recorded. The significance level is set at $P < .05$ (two-tailed).

Processing and analysis of comment content

Three clinical experts [Rui F, JX F, ZX Z] were initially assigned to randomly review 500 comments with scores less than or equal to 3. After a thorough discussion among the experts, a compilation of 10 reasons was derived. This compilation is named as Hospital-Patient Factor (HPF).

Utilizing ChatGPT, a comprehensive analysis and extraction of all comments with scores less than or equal to 3 was conducted. For each negative comment, up to two reasons behind the low scores were extracted from the patients' feedback. The instructional prompts provided to ChatGPT were as follows:

"Hello, I have a collection of patient dissatisfaction comment tags regarding the hospital. I will provide you with a patient's comment and you will assign appropriate tags indicating the reasons for their dissatisfaction. Please note that you can only select from a predefined set of ten tags. [HPF-A: Long waiting times ,HPF-B: Dissatisfaction with the hospital's self-service equipment ,HPF-C: Dissatisfaction with the treatment effect ,HPF-D: Poor attitude of nurses ,HPF-E: Cumbersome medical treatment process ,HPF-F: Concerns about overtreatment ,HPF-G: Perception of understaffed medical personnel ,HPF-H: Poor hospital environment ,HPF-I: Poor attitude of doctors ,HPF-J: Poor attitude of other hospital staff] Please select the two most relevant tags. And please respond with only the tags, without any additional content. Here is the comment: "+ (comment content)". (Figure 1 B)

After processing all the data, comments that cannot be tagged with specific HPF labels by ChatGPT will be deleted. A systematic sampling survey of the ChatGPT output is conducted. Out of every 100 comments, 10 are extracted and given to three clinical experts for manual review in order to assess the authenticity and validity of the ChatGPT results.

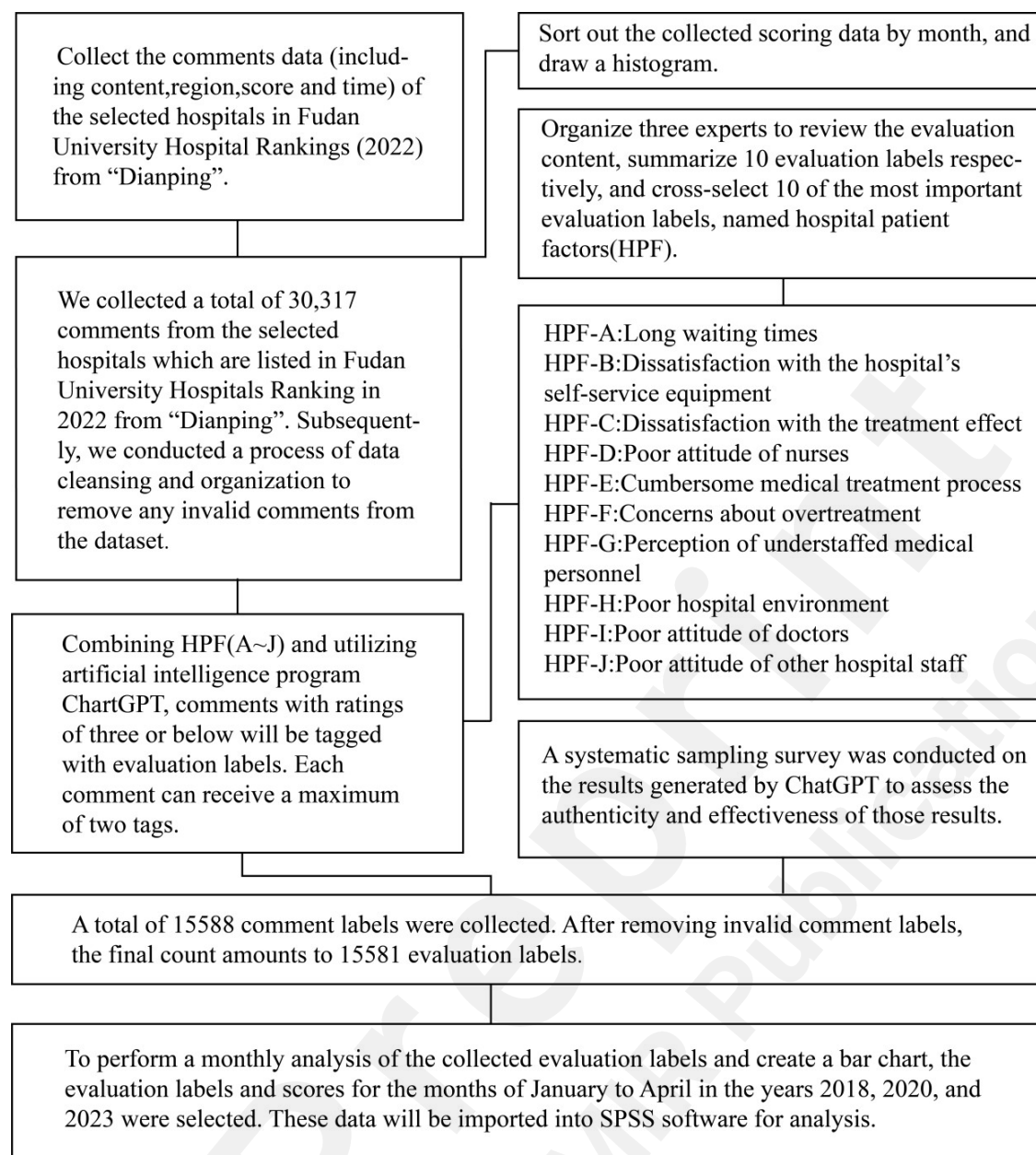


Figure 2. Flowchart depicting the process of data collection and organization in this study.

Results

The general flow of comment information from collection to processing is shown in (Figure 2). In this study, a total of 32,085 pieces of comment data were gathered, and after careful screening, 1,768 comments that did not meet the inclusion criteria were excluded. This resulted in a final collection of 30,317 valid comment records. (Multimedia Appendix 1) Among them, 7,696 were categorized as negative comments (with scores less than or equal to 3). After undergoing processing via ChatGPT, the negative comments yielded a total of 15,588 HPF tags, and after further scrutiny, 7 invalid tags were removed.

A systematic random sampling survey was conducted on the remaining output results to confirm their authenticity and effectiveness. A total of 1519 labels were manually reviewed, and the results show that the accuracy of ChatGPT's identifications reached 92.05%, and a confusion matrix for ChatGPT's automatic classification assessment was plotted. (Figure 3) Based on this, the performance of GPT in text classification was evaluated using manually encoded data. Among them,

the macro F1 score was 0.914.(Table1).[29, 30] Ultimately, 15,581 valid HPF tags were obtained. An overview of the collected data is presented in(Table2).

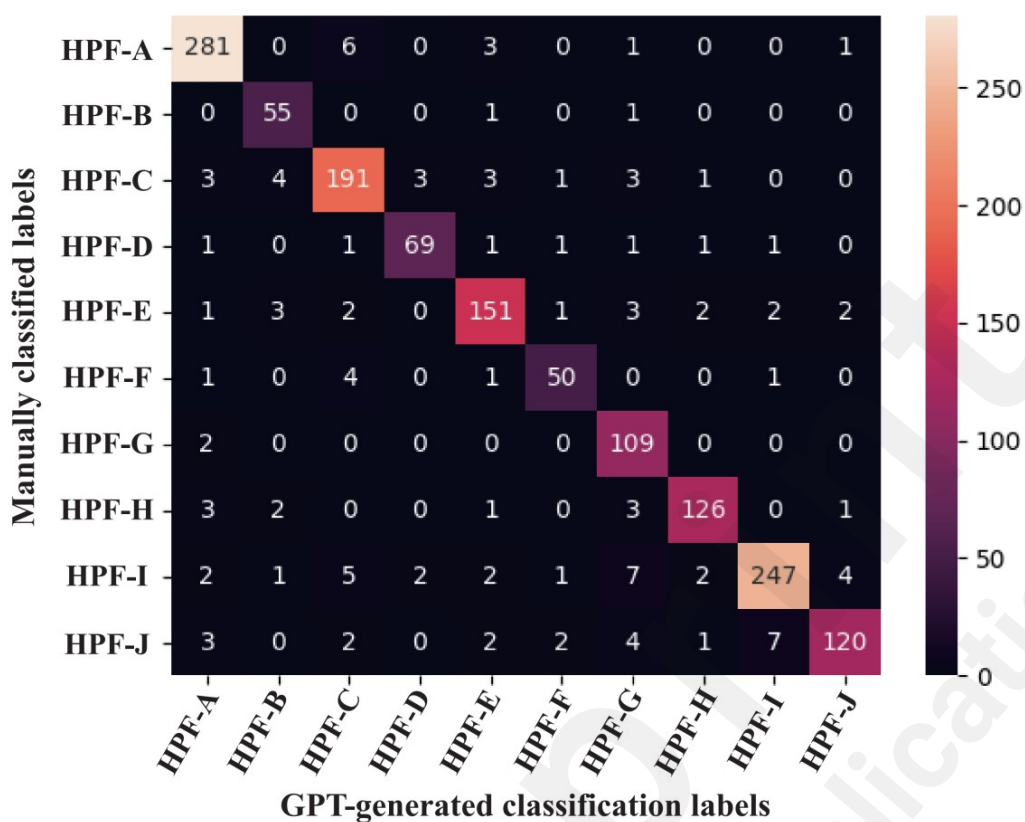


Figure3.Confusion Matrix: Manual Inspection of Chat-GPT's Automatic Classification Results.

Table1 Results of GPT text classification performance evaluation

Categorization of HPF		Precision	Recall	F1 score
Single HPF				
	HPF-A	0.946	0.962	0.954
	HPF-B	0.846	0.965	0.902
	HPF-C	0.905	0.914	0.909
	HPF-D	0.932	0.907	0.921
	HPF-E	0.915	0.904	0.909
	HPF-F	0.893	0.877	0.885
	HPF-G	0.826	0.982	0.897
	HPF-H	0.947	0.926	0.937
	HPF-I	0.957	0.905	0.931
	HPF-J	0.938	0.851	0.892
Aggregate				
	Macro	0.911	0.919	0.914
	Micro	0.921	0.921	0.921

Table2 Overview of Hospital Grouping, Number of Hospitals and Comments Count.

Categorization of Hospitals		Number of Hospitals, n(%)	Number of Comments, n(%)
Area			
	All Hospitals	88(100)	30317(100)

	Hospitals in Shanghai	19(22)	14511(48)
	Hospitals in Other Areas	69(78)	15806(52)
Period			
	Before the pandemic (2018.01~04)	88(100)	1904(6)
	During the outbreak of Covid19(2020.01~04)	88(100)	1949(6)
	After the pandemic (2023.01~04)	88(100)	1016(3)
Type of Hospital			
	Children's Hospital	17(19)	4910(16)
	Maternity Hospital	13(15)	6685(22)
	Tumour Hospital	13(15)	978(3)

Trend of comment scores in all hospitals

The results indicate that the comment scores of all hospitals experienced a gradual increase from January 2018 to January 2019 and remained at a high level from February 2019 to November 2019. After the outbreak of the pandemic in December 2019, the comment scores of all hospitals rose again and reached a peak in February 2020. Subsequently, from March 2020 to March 2022, the comment scores of all hospitals gradually declined over a period of two years, reaching the lowest value since 2018 in March 2022. However, from April 2022 to January 2023, the comment scores experienced a sudden increase followed by a continuous decline. In October 2022, the comment scores sharply rose again, only to plummet after reaching its peak in January 2023. During the period from April 2022 to January 2023, the comment scores of all hospitals exhibited two drastic fluctuations, forming an "M" shaped waveform. Subsequently, from January 2023 to August 2023, the comment scores remained at a lower level.(Figure4A)

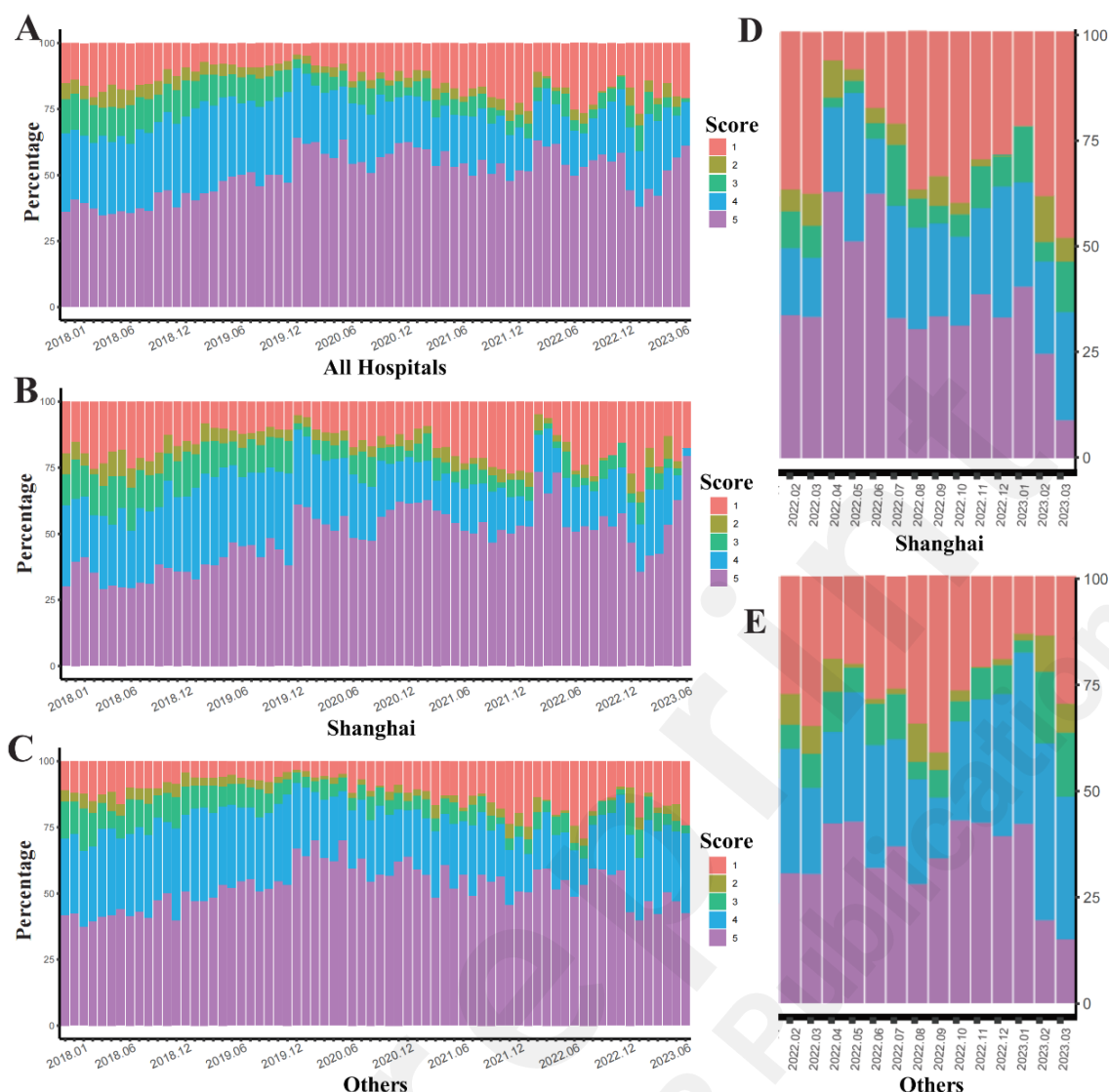


Figure 4. A-C: Trend chart of score changes for all hospitals, hospitals in Shanghai, and hospitals in other areas from January 2018 to August 2023.;D-E:The "M"-shaped fluctuation graph of hospitals in Shanghai, and hospitals in other areas.

The comment scores trend of hospitals in Shanghai and other regions

The comment score trends of hospitals in Shanghai and other regions follow a similar pattern to the average score changes of all hospitals, encompassing a sequence of "gradual increase-maintenance-sudden increase-decline-"M" shaped waveform-low score maintenance." (Figure 4 B,C) During the "M" shaped waveform period from April 2022 to January 2023, the "M" shaped waveform in Shanghai was steeper and experienced more drastic fluctuations compared to other regions. (Figure 4 D,E) This study conducted a separate analysis of the comment scores and content of comments for hospitals in Shanghai and other regions during the "M" shaped waveform period (April 2022 to January 2023). It was found that there were significant differences in the comment scores between hospitals in Shanghai and other regions during this period ($P=.03$). However, there was no significant difference in the composition of the comment content ($P=.59$). (Table 3) (Multimedia Appendix 2 A-C)

Table 3. Score forms for hospitals in the Shanghai region and other areas.

Characteristic	Values
Shanghai,mean(SD)	4.027(1.49)
Rank,n(%)	481(100)
1	74(15)
2	19(4)
3	24(5)
4	67(14)
5	297(62)
Others,mean(SD)	3.874(1.54)
Rank,n(%)	752(100)
1	140(19)
2	21(3)
3	46(6)
4	132(18)
5	413(54)
P.value	.03

The comparison between the period before and after the outbreak of the pandemic and during its peak.

Given the significant impact of the sudden outbreak of the pandemic on hospitals, we selected three periods for comparison: Period A, before the outbreak (January 1, 2018 - April 30, 2018); Period B, during the outbreak (January 1, 2020 - April 30, 2020); and Period C, after the outbreak (January 1, 2023 - April 30, 2023). We compared the comment scores and analyzed the composition of negative feedback during these periods. Clearly, there were significant differences in the scores between Periods A, B, and C ($P<.001$). Pairwise Mann-Whitney U tests further confirmed significant differences among all groups ($P=.007$). The highest average score was observed in Period B, at 4.274. Period A had an average score of 3.647, while Period C had an average score of 3.737. (Figure5 A,B,C,K,L)

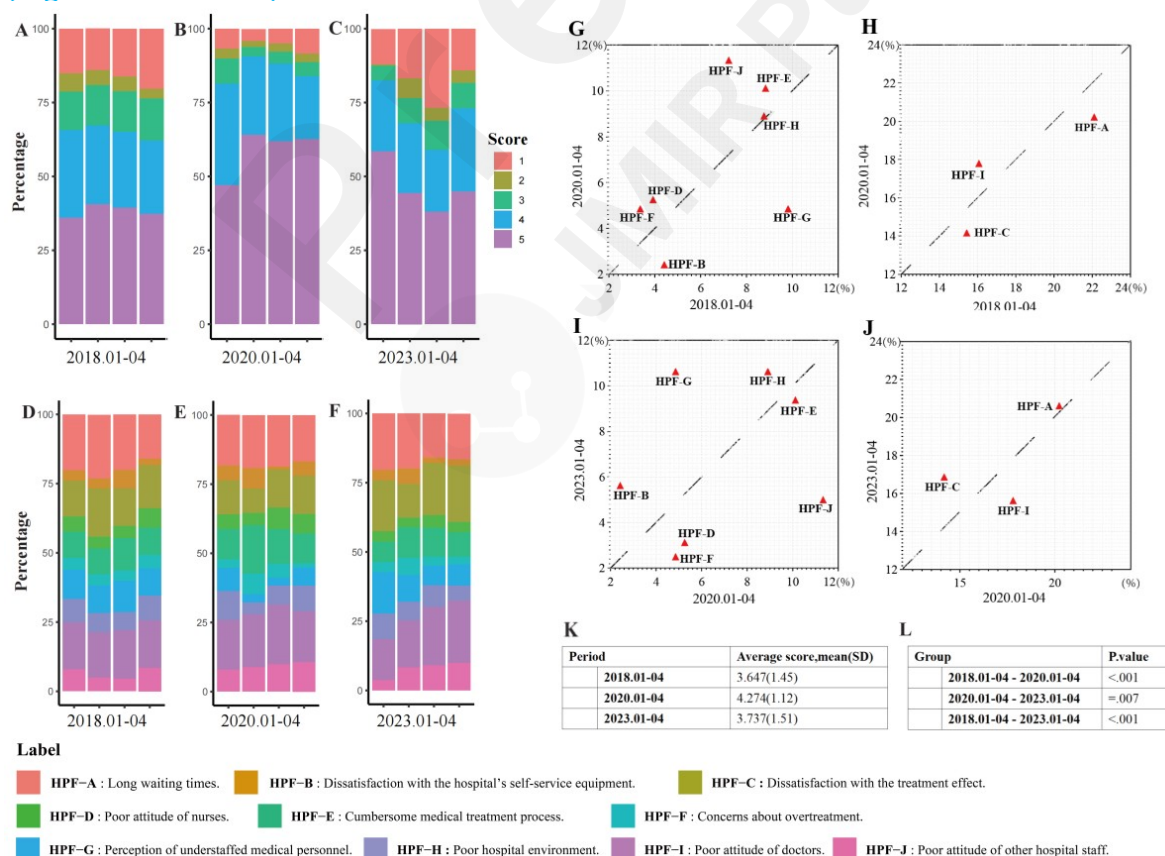


Figure 5.

A-C: Hospitals' score diagrams for the months of January to April in 2018, 2020, and 2023.;D-F: Composition chart of hospital evaluation content for January-April in 2018, 2020, and 2023.;G-H: Comparison chart of hospital evaluation content composition for January to April in 2018 and 2020.;I-J: Comparison chart of hospital evaluation content composition for January to April in 2020 and 2023.;K: Average hospital scores for January to April in 2018, 2020, and 2023.;L: P.value.

We conducted an analysis of the composition of negative feedback during Periods A, B, and C, revealing significant differences among the three groups ($P = .03$). Pairwise comparisons using the chi-square test for two proportions showed no significant difference between Periods A and C ($P = .20$). However, there was a relatively significant difference between Periods A and B ($P = .04$), as well as between Periods C and B ($P = .07$). (Figure 5 D,E,F) (Multimedia Appendix 2 D) Period B, which corresponds to the outbreak period, exhibited a relatively higher proportion of negative comments categorized as HPF-D/E/F/I/J. (Figure 5 G,H,I,J) During periods A and C, which correspond to non-outbreak periods, there was a relatively higher proportion of negative comments classified as HPF-B/G/C.

Comparison of Children's Hospital, Maternity Hospital and Tumor Hospital

We compared the comment scores of these children's hospitals, maternity hospitals, and tumor hospitals, and found significant differences in the comment scores among these three distinct types of specialized medical institutions ($P < .001$). Furthermore, pairwise Wilcoxon rank-sum tests were conducted for each pair of datasets, demonstrating significant differences between any two groups ($P < .001$). Maternity hospitals had the highest average score, at 4.121. Children's hospitals were next, with an average score of 3.812. Tumor hospitals had the lowest average score, at 3.572. Additionally, unlike hospitals of other types, tumor hospitals received lower comment scores during the outbreak compared to before the outbreak. On the other hand, children's hospitals and maternity hospitals received the highest comment scores during the outbreak. (Table 4) (Multimedia Appendix 2 E-G)

Table 4. Score forms for children's hospitals, maternity hospitals and tumor hospitals

Characteristic	Average score, mean(SD)
Before the pandemic (2018.01~04)	N/A
Children's Hospital	3.847(1.48)
Maternity Hospital	3.958(1.31)
Tumour Hospital	3.797(1.42)
During the outbreak of Covid19 (2020.01~04)	N/A
Children's Hospital	4.003(1.35)
Maternity Hospital	4.461(1.19)
Tumour Hospital	3.404(1.61)
After the pandemic (2023.01~04)	N/A
Children's Hospital	3.757(1.49)
Maternity Hospital	4.241(1.34)
Tumour Hospital	3.394(1.51)

We conducted an analysis of the composition of negative comments for children's hospitals, maternity hospitals, and tumor hospitals, and found significant differences in the composition of comments among the three groups ($P < .001$). Further pairwise comparisons using chi-square tests revealed significant differences in the composition of negative comments between children's hospitals and maternity hospitals, as well as between children's hospitals and tumor hospitals ($P < .001$). (Multimedia Appendix 2 H-K)

Compared to other hospitals, maternity hospitals tend to have a relatively higher proportion of negative comments related to HPF-D/J/I, while children's hospitals tend to have a relatively higher proportion of negative comments related to HPF-A/C/F/G. (Figure 6)

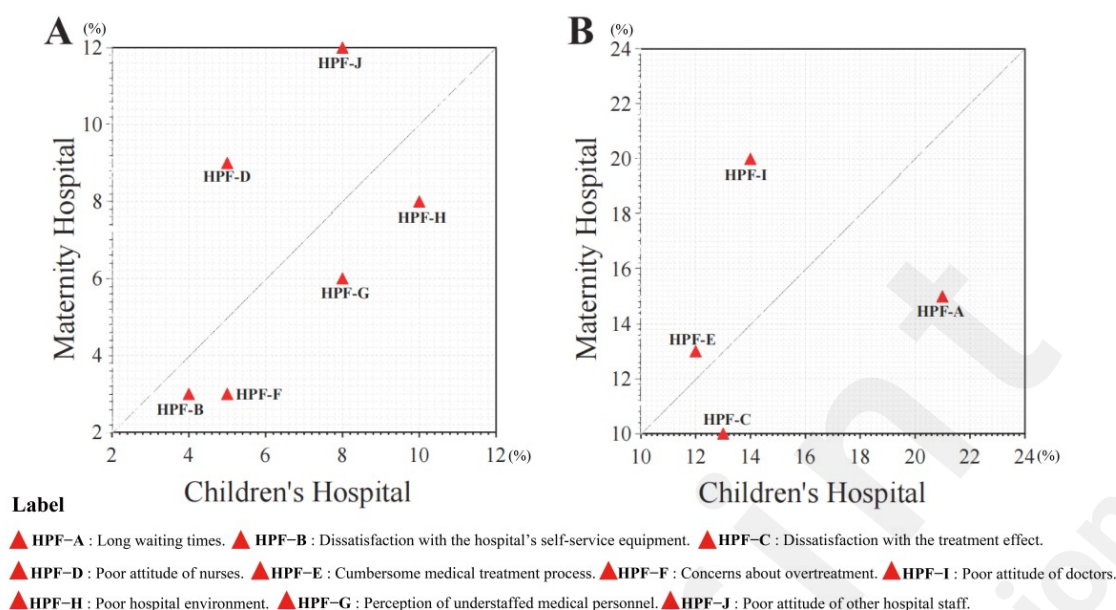


Figure 6. A-B: Comparison chart of the composition of evaluation content between children's hospitals and maternity hospitals.

Discussion

Principal Findings

In this study, we compared the trends in comment scores and composition of comments exhibited by different hospitals in different regions and of different types during the periods before, during, and after the pandemic. We found that at the early stages of the Covid-19 outbreak, there was a certain increase in the patient satisfaction scores for hospitals. (Average score: 4.274, $P < .001$) However, as the pandemic continued, there was a gradual decline in the hospitals' satisfaction scores from patients. (Average score: 3.737, $P < .001$) Moreover, during the pandemic, patients were more likely to express dissatisfaction with the attitudes of doctors, nurses, and other hospital staff.

In addition, in our survey of specialized hospitals, we discovered that, unlike other hospitals, patients at tumor hospitals tended to give lower scores during the pandemic. (Average score: 3.404, $P < .001$) This may be related to the anxiety that cancer patients feel about their treatments. The negative feedback provided by patients at maternity hospitals and children's hospitals also had their own distinctive characteristics.

The comment score trends for all hospitals, hospitals in Shanghai, and hospitals in other regions

Our research demonstrates that during the COVID-19 pandemic in China, which spanned a three-year period, patient comments of hospitals exhibited distinct characteristics across different stages. Existing findings indicate that during the peak of the outbreak (January to June 2020), the physician-patient relationship showed signs of improvement, reaching its optimal state in February 2020. Positive media coverage of healthcare workers, the adoption of telemedicine, and supportive healthcare policies had a significant and positive impact on the physician-patient relationship during

the pandemic.[6, 31-33] Physicians have perceived an increased level of trust from patients in their clinical practice compared to the past. For instance, patients voluntarily turn their heads away when sneezing and display greater respect towards doctors during conversations.

However, the general public does not share the belief that the physician-patient relationship has improved during the pandemic. According to a survey, 61% of the population expressed opposition to the idea of any improvement in the physician-patient relationship during this time.[34] The variations in attitudes and behaviors can be explained by several factors. Firstly, the Chinese government has undertaken the entire cost of treating COVID-19 patients, relieving them of financial concerns and subsequently fostering improved attitudes towards seeking medical care and willingness to undergo treatment. This has contributed to a reduction in the sense of mistrust between doctors and patients.[35, 36] The healthcare providers may perceive an improvement in the physician-patient relationship due to these factors. However, the general public, who are not affected by COVID-19, does not benefit from waived medical expenses. On the contrary, the scarcity of medical resources during the pandemic makes it even more challenging for the general population to seek medical care.

Positive media coverage[37, 38] and the positive impacts of healthcare policies during the pandemic may not necessarily be sustained in the long run. Over a two-year period from March 2020 to March 2022, patient ratings of hospitals gradually declined, reaching their lowest point in March 2022. Multiple factors contributed to this decline. According to Roubille et al., as the duration of the pandemic prolonged, the confidence of the general public may have waned or suffered, ultimately leading to a deterioration in physician-patient relationships due to the adverse effects of COVID-19. [39]

M-shaped pattern

As shown in the (Figure 4 A) During the period of April 2022 to January 2023, the comment scores of all hospitals experienced two significant fluctuations, forming a distinctive "M" shaped pattern. Analyzing the timing, research indicates that the formation of the two peaks of the "M" shaped pattern correlates with the outbreak in Shanghai in the first half of 2022 and the subsequent comprehensive relaxation of COVID-19 control measures in China towards the end of 2022. We have observed that these two incidents have caused varying degrees of strain on medical resources, [40] while simultaneously leading to higher comment scores from patients towards hospitals. This finding aligns with previous research in this area.[6, 31-33]

As shown in the (Figure 4 D,E) The "M" shaped waveform formed during the outbreak in Shanghai from April to July 2022 was steeper, with the highest comment scores reached in April 2022. In contrast to the Shanghai region, hospitals in other areas exhibited a smoother "M" shaped waveform, with scores lower than the peak scores in Shanghai. During this period, both Shanghai and other regions implemented the same healthcare policies, and the media coverage received by the public was consistent. In contrast, the situation in Shanghai is characterized by a more severe outbreak and a greater strain on medical resources.

Based on these observations, we speculate that under similar circumstances, the scarcity of medical resources and the severity of the epidemic may lower patients' expectations for medical outcomes. Consequently, they would value the precious medical resources even more, leading to an increase in comment scores and a decrease in negative reviews.

Comparison between the outbreak and before and after the COVID-19 pandemic

We selected data from January to April 2020, as well as data from January to April 2018 and 2023, representing the periods of the pandemic outbreak, pre-pandemic, and post-pandemic respectively. The results reveal that hospitals received the highest comment scores during the outbreak, with an

average score of 4.274. In comparison, the average scores for the pre-pandemic and post-pandemic periods were 3.647 and 3.737 respectively. We have already discussed these findings in earlier sections.(Figure5 A-C,K,L)Furthermore, through further analysis of the composition of negative comments during these three periods, we have found certain differences between the pre-pandemic and pandemic outbreak periods, as well as the post-pandemic and pandemic outbreak periods. During the pandemic outbreak, there was a relatively higher proportion of HPF-D/E/F/I/J in negative comments, while HPF-B/G/C showed a relatively lower proportion compared to the pre-pandemic and post-pandemic periods. (Figure5 D-J)(Multimedia Appendix 2 D)

HPF-D, HPF-I, and HPF-J respectively represent patient dissatisfaction with nurses, physicians, and other hospital staff. The highest proportion of negative comments is "dissatisfaction with physician attitude."For a long time, "dissatisfaction with physician attitude" has been a significant factor in the occurrence of doctor-patient disputes. These disputes not only directly lead to conflicts between doctors and patients but also erode the level of trust between them, resulting in a stalemate in the diagnosis and treatment process and affecting the quality and standard of healthcare.[41, 42] The reasons for this situation are multifaceted, encompassing significant information asymmetry between doctors and patients, the authoritative status of doctors, poor doctor-patient communication, and a lack of empathy towards patients.[43] During the pandemic, the limited availability of medical resources has led to reduced communication time between doctors and each patient. Consequently, effective communication skills have become crucial in such circumstances. Numerous studies have demonstrated that fostering a positive doctor-patient communication environment can enhance patient trust, contribute to superior treatment experiences, and ultimately improve patient satisfaction and treatment outcomes.[42-44] Hospitals should also strengthen training in clinical and communication skills. [44] When dealing with critically ill patients, physicians must promptly communicate the prognosis to the patients. It is important to alleviate any unrealistic expectations the patients may have regarding the outcome and share the decision-making process with them, in order to prevent medical disputes and deterioration of the doctor-patient relationship.[45, 46]

Similarly, nurses and other hospital staff members are crucial components of the healthcare system. During the pandemic, there has been a notable increase in patients' dissatisfaction with nurses and other staff members. Hospital administrators should prioritize their care and training to enhance service quality. Patient-centered nursing care can provide a better treatment experience for patients. [47] Moreover, there has been an increase in the HPF-E during the pandemic, which emphasizes the need for hospitals to optimize complex healthcare processes to minimize negative patient experiences. This result reminds us that simplifying complicated medical procedures during a pandemic may help reduce the generation of negative emotions in patients.

Difference between Children's Hospital, Maternity Hospital and Tumor Hospital

We selected two time points, December 2019 and November 2022, as reference periods to categorize the comment data into three stages: pre-pandemic, during the pandemic, and post-pandemic. It is noteworthy that while most hospitals experienced an improvement in their comment scores during the pandemic, the scores for cancer hospitals instead showed a decline. From a rating of 3.797 before the pandemic, it decreased to 3.404 during the pandemic (Table4)(Multimedia Appendix 2 G). This anomalous phenomenon may be attributed to the unique characteristics of cancer patients. Compared to ordinary patients, cancer patients often bear a heavier psychological burden, manifested through severe depressive and anxiety symptoms.[48]

Furthermore, during the pandemic, the implementation of containment policies and the scarcity of healthcare resources may have significantly disrupted the diagnosis and treatment of cancer patients.

Research by Takeshi et al. demonstrates that the COVID-19 pandemic, with its strain on healthcare resources, can result in delayed diagnoses, leaving cancer patients already in advanced stages by the time they have the opportunity for diagnosis. The pandemic may also lead to treatment delays for cancer patients, such as postponed surgeries due to circumstances beyond their control.[49-51] Given the circumstances described above, the impact of encountering the COVID-19 pandemic can be disastrous for cancer patients who are in urgent need of diagnosis and treatment. Research by Ye et al. has elucidated that during the COVID-19 outbreak, cancer patients may experience heightened psychological distress due to treatment delays. This suggests that the consequences for cancer patients, when faced with the dual challenge of seeking timely diagnosis and navigating the pandemic, can be profound and deeply distressing.[52, 53] This could be a contributing factor to the lower comment scores of cancer hospitals during the pandemic.

Therefore, for cancer hospitals, ensuring timely diagnosis and treatment for cancer patients in the face of potential future outbreaks is crucial for enhancing healthcare service quality.

When analyzing the composition of negative comments, children's hospitals and maternity hospitals exhibit prominent characteristics. The categories of HPF-A/C/F/G in negative comments are relatively high in children's hospitals. (Figure6 A,B) These categories respectively represent "excessive waiting time," "dissatisfaction with treatment effectiveness," "perceived excessive medical procedures and examinations," and "insufficient healthcare staff." On one hand, the negative comments of "excessive waiting time" and "insufficient healthcare staff" authentically reflect the current shortage of pediatric doctors in China.[54, 55] Based on the basic data from the Chinese Pediatric Medical Resources White Paper, China has approximately 260 million children aged 0-14, but currently, there are only around 100,000 pediatric doctors nationwide. On average, each pediatric doctor is responsible for the care of 2,000 children. This indicates a shortfall of at least 200,000 pediatric doctors in China.[56] The shortage of pediatric doctors is a significant factor contributing to negative comments of children's hospitals. Furthermore, the negative comments of "dissatisfaction with treatment effectiveness" and "perceived excessive medical procedures and examinations" reflect the anxiety and lack of trust that parents of sick children commonly experience toward doctors.[55]

Given the aforementioned context, addressing the shortage of pediatric doctors and alleviating the negative emotions of parents of sick children have emerged as two crucial focal points for enhancing the quality of healthcare services in children's hospitals. Focusing on the aforementioned dimensions, hospital administrators should prioritize training programs for pediatric doctors. Additionally, it is essential to enhance training for pediatric doctors and improve their salary levels without compromising their passion and dedication to caring for children.[56, 57]

When examining negative comments of obstetric and gynecological hospitals, it becomes apparent that the proportions of HPF-D/J/I are relatively high. (Figure6 A,B) These refer respectively to "poor nurse attitude," "poor doctor attitude," and "poor attitude of other hospital staff." It is intriguing to observe that negative assessments in the field of obstetrics and gynecology mainly revolve around "poor attitude." This phenomenon can perhaps be explained from two perspectives. On one hand, during the outbreak of a pandemic, the number of patients has increased significantly, while the number of doctors has not kept pace. As a result, the amount of time and energy they can allocate to each patient is either reduced or can only be sustained by extending their working hours. A study revealed that 73% of obstetrics and gynecology practitioners work over 50 hours per week. In such long and intense work environments, the occupational burnout rate in the field of obstetrics can even reach 56.6%.[58] This will make obstetricians feel irritable and irritable, and reduce their patience with the objects they serve or contact.[59] On the other hand, due to the abnormal hormones in pregnant women, fear of childbirth and serious anxiety, these will make pregnant women more sensitive and uneasy.[60-62]

In conclusion, in order to provide a positive medical experience for pregnant women and reduce the likelihood of negative comments, hospital administrators can address the following two issues. Firstly, they can decrease the working hours of each healthcare professional, increase the number of specialized healthcare personnel, and improve the salary and benefits for healthcare workers. This will help alleviate the occupational burnout among obstetricians and gynecologists. Secondly, timely psychological interventions should be provided for pregnant women with fear of childbirth, including evidence-based treatments such as psychoeducation, cognitive restructuring, and relaxation exercises. These interventions are believed to contribute to reducing the rate of cesarean section deliveries and enhancing overall satisfaction with the childbirth experience.[63, 64]

The limitations of this study

This study has certain limitations. Firstly, the selection of hospitals primarily focused on those with high national rankings, which are mainly located in economically developed cities. Therefore, the conclusions may not be applicable to hospitals in smaller cities. Secondly, unlike conventional surveys, this study only captured publicly available patient comment information, thus lacking demographic data on the individuals who provided these comments. Thirdly, it was not possible to completely eliminate the possibility of some negative comments being maliciously posted by competitors as false information.

Conclusion

The COVID-19 pandemic has some association with patient comment scores. In this study, comment scores of general hospitals, children's hospitals, and maternity hospitals increased during the pandemic, excluding tumor hospitals. The scores and content of patient comments varied among different specialized hospitals. Analyzing patient comment content using ChatGPT is an innovative method for identifying factors contributing to patient dissatisfaction and provides a user-friendly approach. This study has certain reference value for hospital administrators in establishing harmonious relationships between healthcare providers and patients, and in improving hospital performance during public health emergencies.

Acknowledgements

This work was supported by the National Natural Science Foundation of China project (82270505); Shanghai Municipal Hospital Development Center (SHDC12022107);

Data availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Conflict of interest

None declared.

Abbreviations

HPF : Hospital-Patient Factor;

Multimedia Appendix 1

Original Data: 30,317 pieces of hospital review information collected from the Dianping website.

Multimedia Appendix 2

Supplementary picture.

Reference

1. Xiao Y, Liu L, Chen TT, Zhu SY. Rebuilding trust in the physician-patient relationship: addressing causes and implementing reforms in China. *Fam Pract*. 2023.PMID: 37942660
2. Duan G, Liao X, Yu W, Li G. Classification and Prediction of Violence Against Chinese Medical Staff on the Sina Microblog Based on a Self-Organizing Map: Quantitative Study. *J Med Internet Res*. 2020;22(5):e13294.PMID: 32348253
3. Peng W, Ding G, Tang Q, Xu L. Continuing violence against medical personnel in China: A flagrant violation of Chinese law. *Biosci Trends*. 2016;10(3):240-3.PMID: 27245351
4. Zhang L, Wang A, Xie X, Zhou Y, Li J, Yang L, et al. Workplace violence against nurses: A cross-sectional study. *Int J Nurs Stud*. 2017;72:8-14.PMID: 28412581
5. Kaplan SH, Greenfield S, Ware JE, Jr. Assessing the effects of physician-patient interactions on the outcome s of chronic disease. *Med Care*.27(3 Suppl):S110-27.PMID: 2646486
6. Li J, Pang PC, Xiao Y, Wong D. Changes in Doctor-Patient Relationships in China during COVID-19: A Text Mining Analysis. *Int J Environ Res Public Health*. 2022;19(20).PMID: 36294022
7. Zhang J, Lu H, Zeng H, Zhang S, Du Q, Jiang T, et al. The differential psychological distress of populations affected by the COVID-19 pandemic. *Brain Behav Immun*. 2020;87:49-50.PMID: 32304883
8. Parsons Leigh J, Fiest K, Brundin-Mather R, Plotnikoff K, Soo A, Sypes EE, et al. A national cross-sectional survey of public perceptions of the COVID-19 pandemic: Self-reported beliefs, knowledge, and behaviors. *PLoS One*. 2020;15(10):e0241259.PMID: 33095836
9. Howard M, Fleming ML, Parker E. Patients do not always complain when they are dissatisfied: implications for service quality and patient safety. *J Patient Saf*. 2013;9(4):224-31.PMID: 24257066
10. Phau I, Baird M. Complainers versus non-complainers retaliatory responses towards service dissatisfactions. *Marketing Intelligence & Planning*. 2008;26(6):587-604.doi:10.1108/02634500810902848
11. Gurung G, Derrett S, Gauld R, Hill PC. Why service users do not complain or have 'voice': a mixed-methods study from Nepal's rural primary health care system. *BMC Health Serv Res*. 2017;17(1):81.PMID: 28122552
12. Jiang Y, Ying X, Zhang Q, Tang SR, Kane S, Mukhopadhyay M, et al. Managing patient complaints in China: a qualitative study in Shanghai. *BMJ Open*. 2014;4(8):e005131.PMID: 25146715
13. Skalen C, Nordgren L, Annerback EM. Patient complaints about health care in a Swedish County: characteristics and satisfaction after handling. *Nurs Open*. 2016;3(4):203-11.PMID: 27708831
14. Zhang W, Deng Z, Hong Z, Evans R, Ma J, Zhang H. Unhappy Patients Are Not Alike: Content Analysis of the Negative Comments from China's Good Doctor Website. *J Med Internet Res*. 2018;20(1):e35.PMID: 29371176
15. Gordon HS. Consider Embracing the Reviews from Physician Rating Websites. *J Gen Intern Med*. 2017;32(6):599-600.PMID: 28299603
16. Hsieh SY. A system for using patient complaints as a trigger to improve quality. *Qual Manag Health Care*. 2011;20(4):343-55.PMID: 21971031
17. Kung TH, Cheatham M, Medenilla A, Sillos C, De Leon L, Elepano C, et al. Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models.

PLOS Digit Health. 2023;2(2):e0000198.PMID: 36812645

18. Kuroiwa T, Sarcon A, Ibara T, Yamada E, Yamamoto A, Tsukamoto K, et al. The Potential of ChatGPT as a Self-Diagnostic Tool in Common Orthopedic Diseases: Exploratory Study. *J Med Internet Res*. 2023;25:e47621.PMID: 37713254

19. Li J. Security Implications of AI Chatbots in Health Care. *J Med Internet Res*. 2023;25:e47551.PMID: 38015597

20. Zhu L, Mou W, Chen R. Can the ChatGPT and other large language models with internet-connected database solve the questions and concerns of patient with prostate cancer and help democratize medical knowledge? *J Transl Med*. 2023;21(1):269.PMID: 37076876

21. Zhu L, Mou W, Yang T, Chen R. ChatGPT can pass the AHA exams: Open-ended questions outperform multiple-choice format. *Resuscitation*. 2023;188:109783.PMID: 37349064

22. Strong E, DiGiammarino A, Weng Y, Basaviah P, Hosamani P, Kumar A, et al. Performance of ChatGPT on free-response, clinical reasoning exams. *medRxiv*. 2023.PMID: 37459090

23. Haman M, Skolnik M. Testing ChatGPT's Capabilities for Social Media Content Analysis. *Aesthetic Plast Surg*. 2023.PMID: 37605022

24. Liu J, Wang C, Liu S. Utility of ChatGPT in Clinical Practice. *J Med Internet Res*. 2023;25:e48568.PMID: 37379067

25. Sun H, Zhang K, Lan W, Gu Q, Jiang G, Yang X, et al. An AI Dietitian for Type 2 Diabetes Mellitus Management Based on Large Language and Image Recognition Models: Preclinical Concept Validation Study. *J Med Internet Res*. 2023;25:e51300.PMID: 37943581

26. Conroy RM, Teehan M, Siriwardena R, Smyth O, McGee HM, Fernandes P. Attitudes to doctors and medicine: The effect of setting and doctor-patient relationship. *Br J Health Psychol*. 7(Pt 1):117-25.PMID: 14596722

27. Yan X, Li J. Better doctor-patient relationships start with the small things. *BMJ*. 2023;383:2935.PMID: 38101920

28. Anonymous users' evaluation and rating page for Fudan University Affiliated Zhongshan Hospital (West Campus). Webpage in Chinese. Shanghai Hantao Information Consulting Co., Ltd. 2023 Jun 30. URL: <https://www.dianping.com/review/1317174326> [accessed 2024-01-23]

29. Schopow N, Osterhoff G, Baur D. Applications of the Natural Language Processing Tool ChatGPT in Clinical Practice: Comparative Study and Augmented Systematic Review. *JMIR Med Inform*. 2023;11:e48933.PMID: 38015610

30. Domingos P. A few useful things to know about machine learning. *Communications of the ACM*. 2012;55(10):78-87.doi:10.1145/2347736.2347755

31. Zhou Y, Ma Y, Yang WFZ, Wu Q, Wang Q, Wang D, et al. Doctor-patient relationship improved during COVID-19 pandemic, but weakness remains. *BMC Fam Pract*. 2021;22(1):255.PMID: 34937550

32. Zhou Y, Yang WFZ, Ma Y, Wu Q, Yang D, Liu T, et al. Doctor-Patient Relationship in the Eyes of Medical Professionals in China During the COVID-19 Pandemic: A Cross-Sectional Study. *Front Psychiatry*. 2021;12:768089.PMID: 34777069

33. Zhou Y, Chen S, Liao Y, Wu Q, Ma Y, Wang D, et al. General Perception of Doctor-Patient Relationship From Patients During the COVID-19 Pandemic in China: A Cross-Sectional Study. *Front Public Health*. 2021;9:646486.PMID: 34295863

34. Gao B, Dong J. Does the Impact of COVID-19 Improve the Doctor-Patient Relationship in China? *Am J Med Sci*. 360(3):305-6.PMID: 32563570

35. Semenov AV, Pshenichnaya NY. Born in Wuhan: lessons from COVID-19 epidemic in China. *Russian Journal of Infection and Immunity*. 2020;10(2):210-20.doi: 10.15789/2220-7619-BIW-1453

36. Liu S, Dai L, Xu J. Tell health stories comprehensively and accurately: A case study of health edition of People's Daily. *Int J Nurs Sci*. 2020;7(Suppl 1):S46-S51.PMID: 32995379

37. Jiang L, Broome ME, Ning C. The performance and professionalism of nurses in the fight against the new outbreak of COVID-19 epidemic is laudable. *Int J Nurs Stud*. 2020;107:103578.PMID: 32446015

38. Gan Y, Chen Y, Wang C, Latkin C, Hall BJ. The fight against COVID-19 and the restoration of trust in Chinese medical professionals. *Asian J Psychiatr*. 2020;51:102072.PMID: 32334408

39. Roubille C, Ribstein J, Hurpin G, Fesler P, Fiat E, Roubille F. Confidence vanished or impaired until distrust in the doctor-patient relationship because of COVID-19: Confidence vanished or impaired until distrust: "COVID" in relationship. *Rev Med Interne*. 2021;42(1):58-60.PMID: 33168352

40. Lyu B, Xu M, Lu L, Zhang X. Burnout syndrome, doctor-patient relationship and family

support of pediatric medical staff during a COVID-19 Local outbreak in Shanghai China: A cross-sectional survey study. *Front Pediatr*. 2023;11:1093444.PMID: 36861079

41. Jiang MM, Wu ZY, Tu AX. Research on the Cooperative Governance Path of Multiple Stakeholders in Doctor-Patient Disputes under the Environment of Information Asymmetry. *Int J Environ Res Public Health*. 2023;20(2).PMID: 36674353

42. Wang Y, Wu H, Lei X, Shen J, Feng Z. The Influence of Doctors' Online Reputation on the Sharing of Outpatient Experiences: Empirical Study. *J Med Internet Res*. 2020;22(12):e16691.PMID: 33306028

43. Drossman DA, Chang L, Deutsch JK, Ford AC, Halpert A, Kroenke K, et al. A Review of the Evidence and Recommendations on Communication Skills and the Patient-Provider Relationship: A Rome Foundation Working Team Report. *Gastroenterology*. 2021;161(5):1670-88 e7.PMID: 34331912

44. Tavakoly Sany SB, Behzad F, Ferns G, Peyman N. Communication skills training for physicians improves health literacy and medical outcomes among patients with hypertension: a randomized controlled trial. *BMC Health Serv Res*. 2020;20(1):60.PMID: 31973765

45. Liu Y, Wang P, Bai Y. The influence factors of medical disputes in Shanghai and implications - from the perspective of doctor, patient and disease. *BMC Health Serv Res*. 2022;22(1):1128.PMID: 36071431

46. Labrie N, Schulz PJ. Does argumentation matter? A systematic literature review on the role of argumentation in doctor-patient communication. *Health Commun*. 2014;29(10):996-1008.PMID: 24359318

47. Liang Z, Xu M, Liu G, Zhou Y, Howard P. Patient-centred care and patient autonomy: doctors' views in Chinese hospitals. *BMC Med Ethics*. 2022;23(1):38.PMID: 35395761

48. Morrison EJ, Novotny PJ, Sloan JA, Yang P, Patten CA, Ruddy KJ, et al. Emotional Problems, Quality of Life, and Symptom Burden in Patients With Lung Cancer. *Clin Lung Cancer*. 2017;18(5):497-503.PMID: 28412094

49. Terashima T, Tsutsumi A, Iwami E, Kuroda A, Nakajima T, Eguchi K. Delayed visit and treatment of lung cancer during the coronavirus disease 2019 pandemic in Japan: a retrospective study. *J Int Med Res*. 2022;50(5):3000605221097375.PMID: 35579175

50. Liu C, Piao H, Zhang T, Yang D, Li X, Tang X. Delayed Diagnosis and Treatment of Cancer Patients During the COVID-19 Pandemic in Henan, China: An Interrupted Time Series Analysis. *Front Public Health*. 2022;10:881718.PMID: 35685763

51. Ribeiro R, Wainstein AJA, de Castro Ribeiro HS, Pinheiro RN, Oliveira AF. Perioperative Cancer Care in the Context of Limited Resources during the COVID-19 Pandemic: Brazilian Society of Surgical Oncology Recommendations. *Ann Surg Oncol*. 2021;28(3):1289-97.PMID: 32980998

52. Ye Y, Wang J, Cai S, Fu X, Ji Y. Psychological distress of cancer patients caused by treatment delay during the COVID-19 pandemic in China: A cross-sectional study. *Psychooncology*. 2022;31(9):1607-15.PMID: 35506550

53. Bolger JC, Donlon NE, Butt W, Neary C, Al Azzawi M, Brett O, et al. Successful maintenance of process and outcomes for oesophageal cancer surgery in Ireland during the first wave of the COVID-19 pandemic. *Ir J Med Sci*. 2022;191(2):831-7.PMID: 33728528

54. Hu KJ, Sun ZZ, Rui YJ, Mi JY, Ren MX. Shortage of paediatricians in China. *Lancet*. 2014;383(9921):954.PMID: 24629297

55. Zhang Y, Huang L, Zhou X, Zhang X, Ke Z, Wang Z, et al. Characteristics and Workload of Pediatricians in China. *Pediatrics*.144(1):e20183532.PMID: 31253739

56. Hu Y, Xu J, Dong W, Yuan Z, Sun X. Strategies to correct the shortage of paediatricians in China. *Lancet*. 2018;392(10145):385.PMID: 30102171

57. Deng W, Feng Z, Yao X, Yang T, Jiang J, Wang B, et al. Occupational identity, job satisfaction and their effects on turnover intention among Chinese Paediatricians: a cross-sectional study. *BMC Health Serv Res*. 2021;21(1):6.PMID: 33397391

58. Ye J, Wang H, Wu H, Ye L, Li Q, Ma XY, et al. Burnout among obstetricians and paediatricians: a cross-sectional study from China. *BMJ Open*. 2019;9(1):e024205.PMID: 30782730

59. Rosenstein AH. Physician stress and burnout: what can we do? *Physician Exec*.38(6):22-6, 8, 30.PMID: 23888648

60. Fenwick J, Toohill J, Creedy DK, Smith J, Gamble J. Sources, responses and moderators of childbirth fear in Australian women: a qualitative investigation. *Midwifery*. 2015;31(1):239-46.PMID: 25440298

61. Jokic-Begic N, Zigic L, Nakic Rados S. Anxiety and anxiety sensitivity as predictors of fear of childbirth: different patterns for nulliparous and parous women. *J Psychosom Obstet Gynaecol*. 2014;35(1):22-8.PMID: 24328559
62. Koc AE, Colak S, Colak GV, Pusuroglu M, Hocaoglu C. Investigating fear of childbirth in pregnant women and its relationship between anxiety sensitivity and somatosensory amplification. *J Obstet Gynaecol*. 2021;41(2):217-23.PMID: 32314633
63. Rouhe H, Salmela-Aro K, Toivanen R, Tokola M, Halmesmaki E, Saisto T. Obstetric outcome after intervention for severe fear of childbirth in nulliparous women - randomised trial. *BJOG*. 2013;120(1):75-84.PMID: 23121002
64. Watt MC, Stewart SH, Lefaivre MJ, Uman LS. A brief cognitive-behavioral approach to reducing anxiety sensitivity decreases pain-related anxiety. *Cogn Behav Ther*. 2006;35(4):248-56.PMID: 17189242



Supplementary Files

Figures

A: Negative Comment Samples from 'Dianping' Website.; B: ChatGPT Handling Demonstration.

A

B

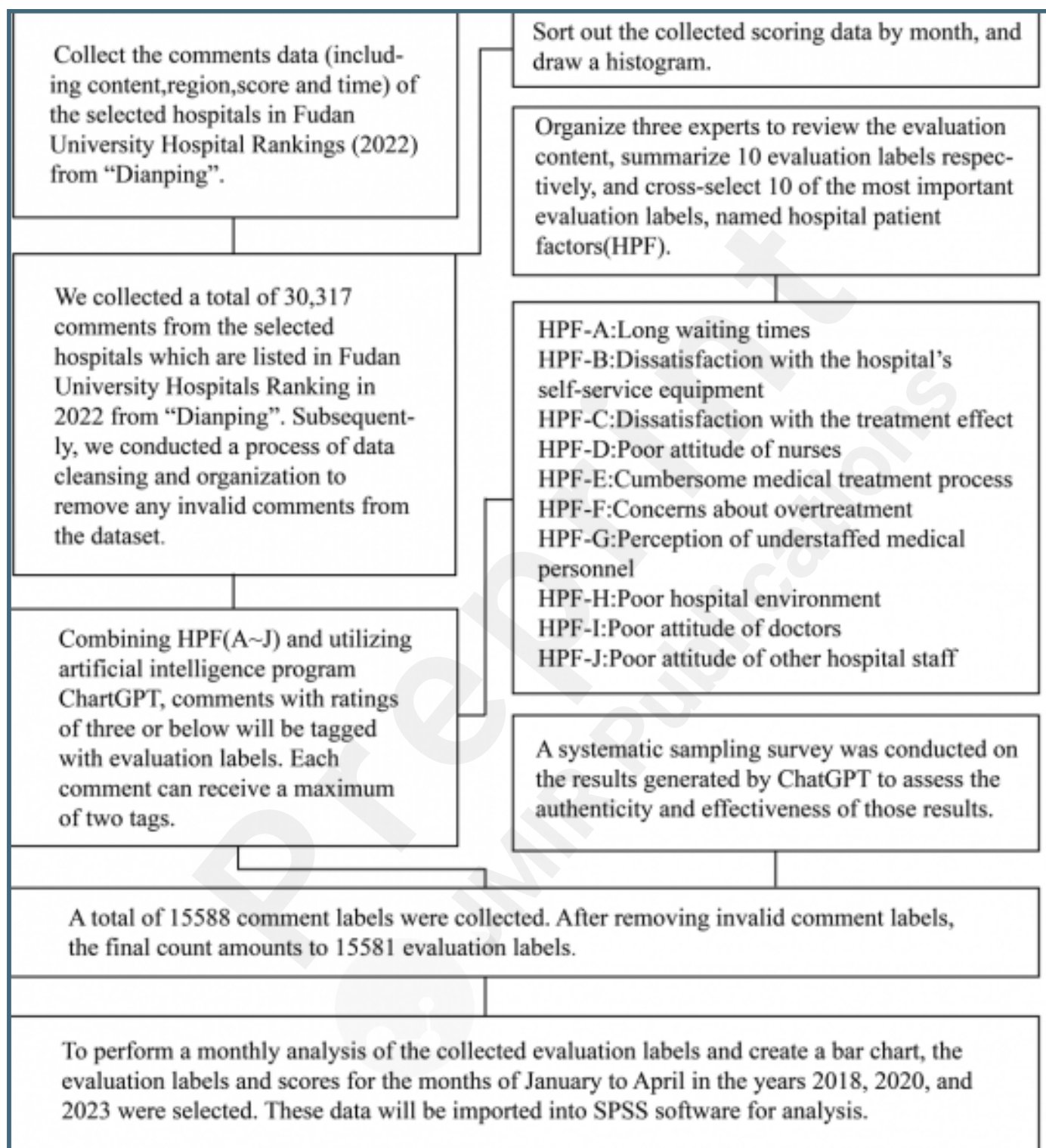
The Interrogative Instructions for GPT

你好, 我这里有一些患者对医院不满意的评价标签, 可不可以我发一段患者的评价信息, 你给这段评价打上患者不满意的原因的标签。注意, 只能在我规定的十个标签范围内 (1.流程繁琐 2.医生态度不佳 3.护士态度不佳 4.医院其他工作人员态度不佳 5.认为医院环境脏乱差或者吵闹拥挤 6.对医院自助服务设备感到不满意 7.等待过长 8.对治疗效果不满意 9.认为存在过度医疗和过度检查 10.认为医护人员人手不足, 或者难以找到医生进行有效沟通) 中选出最为相关的两项。并且, 只回复标签就可以了, 不要说其他的内容。同时, 请在标签前面加上数字编号, 例如“1.”, “2.”。可以吗? 以下是评价内容: 2楼普外科, 从医生、医导台到发票机那坐的老阿姨, 都脾气非常重 看病的医生等下班了看得非常不耐烦, 病历也没打, 却在系统里打勾说打了 后来再找去打, 发现已经下班了 去向导医台这种情况怎么办, 导医台的更是可怕, 直接态度就非常差, 好像我问她是个很冒犯她的事情, 不懂, 导医台不就是答疑的吗? 心情不好就请假不要上班了吧, 大家来看病, 本来心情也不咋好了, 我对你客客气气, 你凶我, 我欠你啥了, 凭什么接受你的情绪

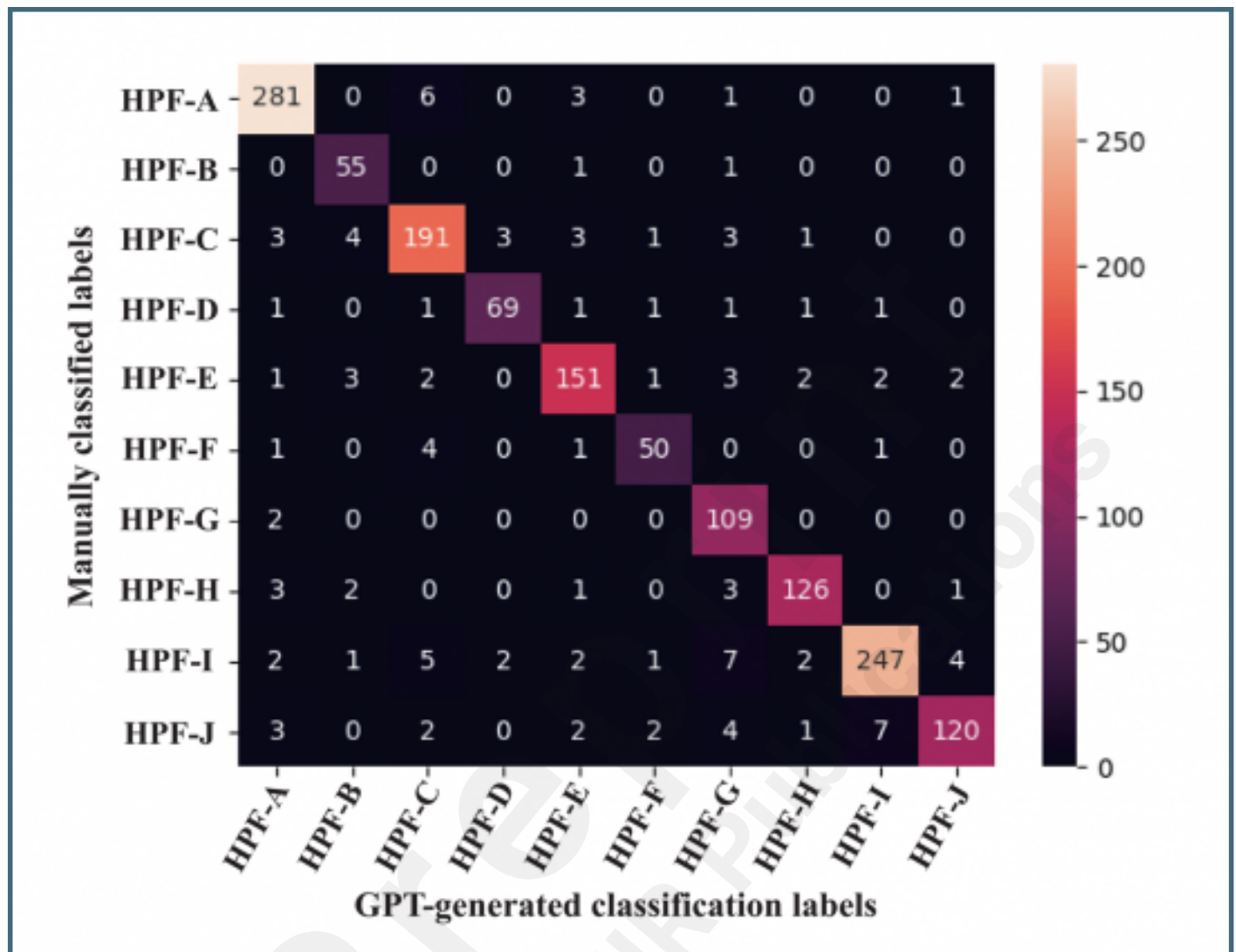
Negative Comments Labels Generated by GPT

2. 医生态度不佳
3. 医院其他工作人员态度不佳

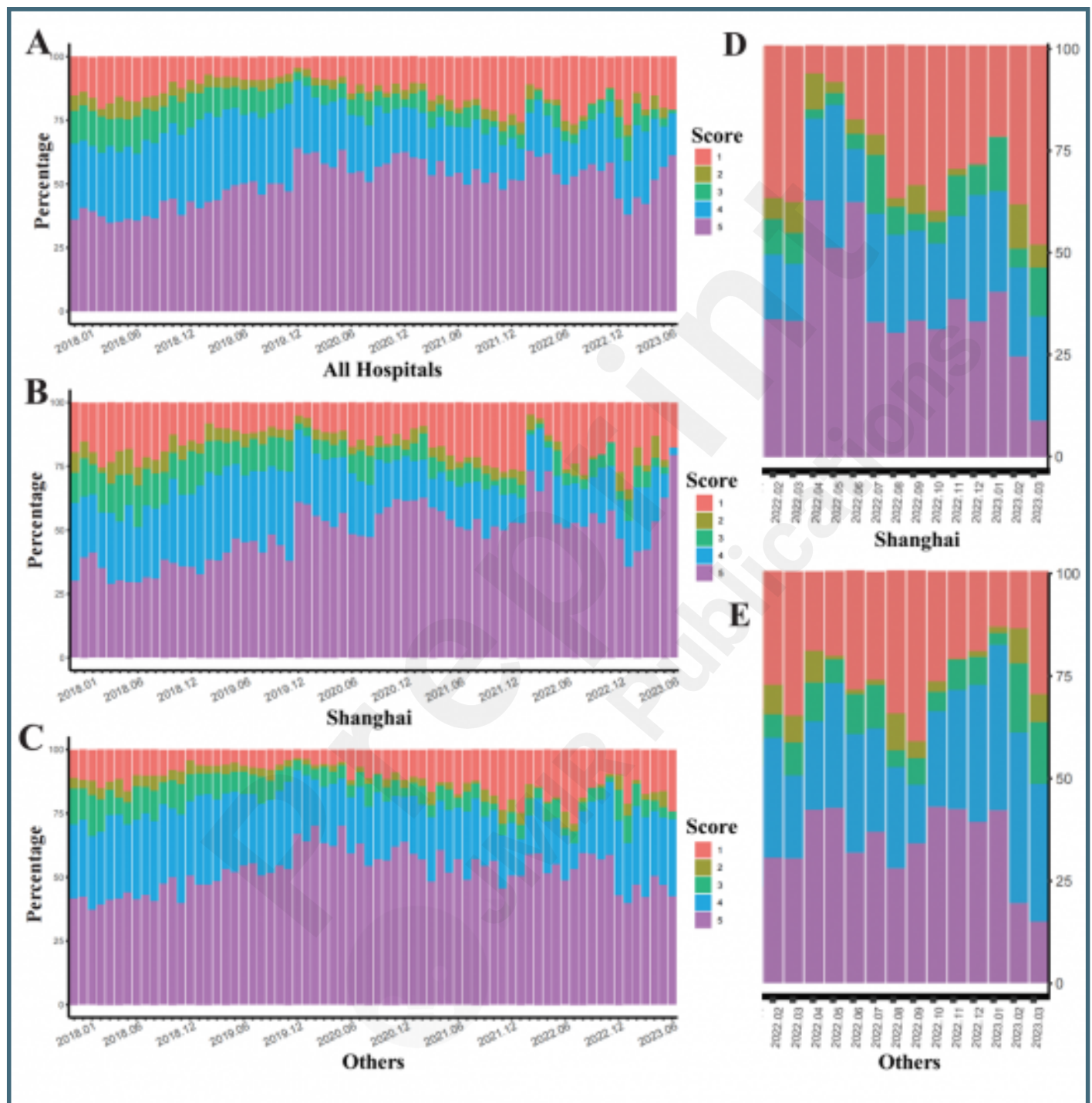
Flowchart depicting the process of data collection and organization in this study.



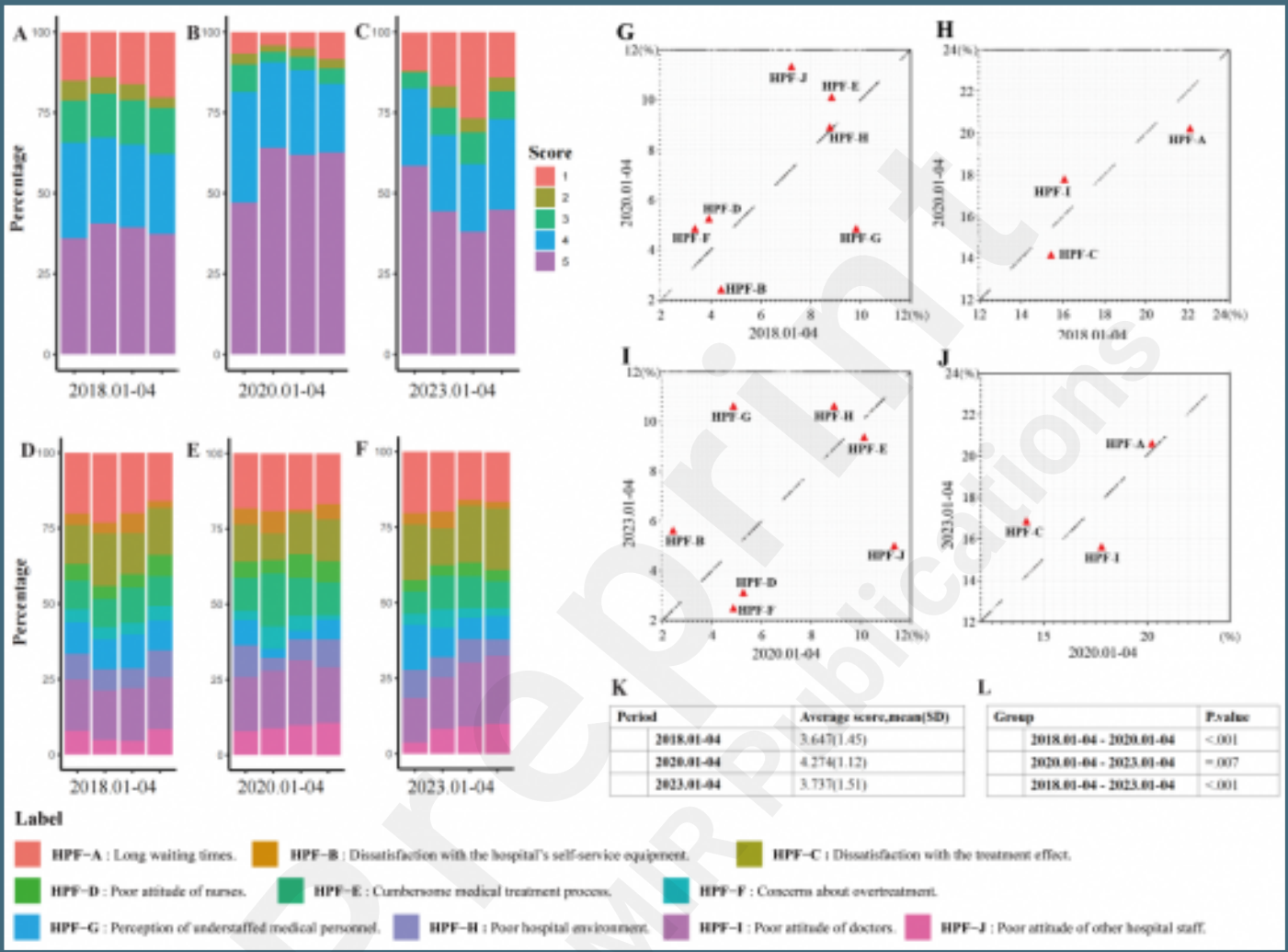
Confusion Matrix: Manual Inspection of Chat-GPT's Automatic Classification Results.



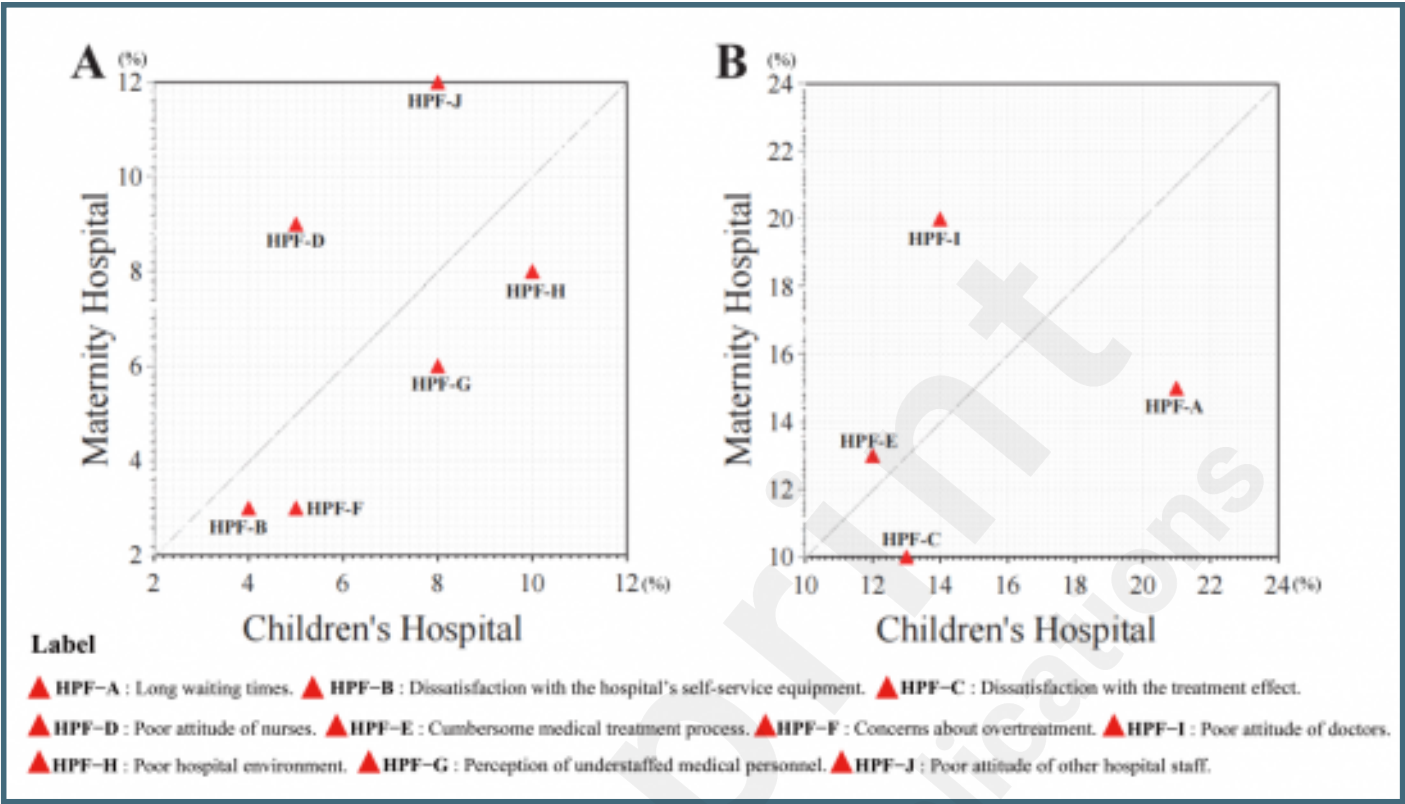
A-C: Trend chart of score changes for all hospitals, hospitals in Shanghai, and hospitals in other areas from January 2018 to August 2023.;D-E:The "M"-shaped fluctuation graph of hospitals in Shanghai, and hospitals in other areas.



A-C: Hospitals' score diagrams for the months of January to April in 2018, 2020, and 2023.;D-F: Composition chart of hospital evaluation content for January-April in 2018, 2020, and 2023.;G-H:Comparison chart of hospital evaluation content composition for January to April in 2018 and 2020.;I-J:Comparison chart of hospital evaluation content composition for January to April in 2020 and 2023.;K:Average hospital scores for January to April in 2018, 2020, and 2023.;L:P.value.



A-B: Comparison chart of the composition of evaluation content between children’s hospitals and maternity hospitals.



Multimedia Appendixes

Original Data: 30,317 pieces of hospital review information collected from the Dianping website.

URL: <http://asset.jmir.pub/assets/fc8148503949200b97885a0dfe45d29a.xlsx>

Supplementary picture.

URL: <http://asset.jmir.pub/assets/1f524a49c7863fa5142c945fd1f48794.pdf>

