

Quality of asynchronous webchats vs in-person consultations for postpartum depression in China: a cross-sectional, mixed methods study using standardized patients



Wenjie Gong,^{a,b,c,e,*} Lu Liu,^{a,e} Xiaoyu Li,^{a,e} Eric D. Caine,^c Jingcheng Shi,^d Zhen Zeng,^a and Kar Keung Cheng^b

^aHER Team and Department of Maternal and Child Health, Xiangya School of Public Health, Central South University, Changsha 410078, China

^bInstitute of Applied Health Research, University of Birmingham, Birmingham B15 2TT, United Kingdom

^cDepartment of Psychiatry, University of Rochester Medical Center, Rochester, NY

^dDepartment of Epidemiology and Health Statistics, Xiangya School of Public Health, Central South University, Changsha, China



Summary

Background Prompt professional care for postpartum depression (PPD) is difficult to obtain in China. Though online consultations improve accessibility and reduce stigma, the quality of services compared to in-person consultations is unclear.

Methods Five trained, undisclosed “standardized patients” (SPs) made “asynchronous webchats” visits and in-person visits with psychiatrists. Visits were made to 85 psychiatrists who were based in 69 hospitals in ten provincial capital cities. The care between online and in-person consultations with the same psychiatrist was compared, including diagnosis, guideline adherence, and patient-centeredness. False discovery rate (FDR) was used to adjust p values. Third visits using asynchronous webchats were made to psychiatrists who offered discrepant diagnoses. Thematic content analysis was used for the discrepancies.

Findings The proportion of diagnostic accuracy was lower for online than in-person visits (76.5% [65/85] vs 91.8% [78/85]; $p_{FDR} = 0.0066$), as were the proportions of completing questions involving clinical history (16.6% vs 42.7%; $p_{FDR} < 0.0001$), and management decisions (16.2% vs 27.5%; $p_{FDR} < 0.0001$) consistent with recommended guidelines. Patient-centeredness was lower online than in-person ($p_{FDR} < 0.0001$). Fifteen of 16 psychiatrists completed third visits, most of them considered lack of nonverbal information online as a key barrier.

Interpretation Online consultations using asynchronous webchats were inferior to in-person consultations, with respect to diagnostic accuracy, adherence to recommended clinical guidelines, and patient-centeredness. To fully realise the potential benefits of online consultations and to prevent safety issues, there is an urgent need for major improvement in the quality and oversight of these consultations.

Funding China Medical Board, National Natural Science Foundation of China, and Swiss Agency for Development and Cooperation Global Cooperation Department.

Copyright © 2024 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Keywords: Telemedicine; Postpartum depression; Standardized patients; Quality of Health Care; Comparative Study

Introduction

Postpartum depression (PPD) adversely affects both mothers and their offspring, with an estimated prevalence of 17.7% (95% confidence intervals [CI] 16.6–18.8%) globally and 14.8% (95% CI 13.1–16.6%) in China.^{1,2} Among Chinese women, concerns about

stigma and the practice of “doing-the-month” (a traditional ritual performed during the 42 days after childbirth that includes not leaving the house and not seeing strangers when practiced strictly) may deter some women with PPD from seeking healthcare support delivered face to face.³ Online consultations therefore

*Corresponding author. Xiangya School of Public Health, Central South University, 172 Tongzipo Road, Yuelu District, Changsha, Hunan 410078, China.

E-mail addresses: gongwenjie@csu.edu.cn (W. Gong), luna_liulu@163.com (L. Liu), 216911035@csu.edu.cn (X. Li), eric_caine@urmc.rochester.edu (E.D. Caine), jingzhengs@126.com (J. Shi), zhenzeng@csu.edu.cn (Z. Zeng), k.k.cheng@bham.ac.uk (K.K. Cheng).

^eWenjie Gong, Lu Liu and Xiaoyu Li contributed equally to the manuscript as joint first authors.

Research in context

Evidence before this study

We searched CNKI, and WANFANG, WEB OF SCIENCE, Cochrane Library, PubMed, Embase, and Ovid for publications in English or Chinese, from database inception until August 1, 2022 that reported comparisons of the quality of mental health consultations online and in-person. Search terms are shown in the [Appendix A](#). We identified 30 relevant studies, seventeen of which were conducted in the United States and none in China; sixteen of which involved online consultations via telephone, fourteen involved videos, and none involved asynchronous webchat. Twenty-five of them suggested no significant differences between online and in-person diagnoses, ten reported that patients were generally satisfied with online consultations, and two suggested that online consultations boosted efficiency and reduced costs. In all previous studies, however, real patients were assessed using randomized assignments or repeated measurements, rather than standardized patients (SPs). Only one article compared satisfaction of patients with online and in-person consultations with the same clinician, and both the clinician and patients were aware that it was a test.

Added value of this study

This is the first reported study using SPs presenting mental health symptoms to compare the quality of care of online and in-person consultations conducted by the same psychiatrists. We found that in China, though asynchronous webchat was the predominant mode of online consultations offered by direct-to-consumer telemedicine platforms, they were inferior to in-person consultations for diagnostic accuracy, adherence to guidelines, and patient-centeredness among women who presented with symptoms consistent with PPD.

Implications of all the available evidence

The use of asynchronous webchats for consultations with women with PPD is problematic. As they constitute the overwhelming majority of activities on telemedicine platforms, there is an urgent need for major improvement in their quality in order to prevent safety issues and to realize the potential benefits of such consultations in China.

represent potential solutions. As there are seldom needs for physical examination, mental health consultations may be particularly suitable online.⁴

In China, the National Health Commission issued a plan in 2020 that included boosting professional support for women during pregnancy and childbirth through online consultations, a policy that has been accelerated by the COVID-19 pandemic.⁵ “Asynchronous webchat” has become the most common mode, in contrast with the dominance of real-time telephone and video consultation in most other countries.⁶ It is a text-based online mode initiated by patients and health professionals can chat with patients after accepting the requests for consultations, which do not require the two parties to be present and texting simultaneously: patients and health professionals can start, pause, and resume a chat whenever they find convenient. They are generally limited to 3–30 rounds of conversations over 24–48 h. In 2021, one leading direct-to-consumer telemedicine (DTCT) platforms provider reported that consultations via asynchronous webchats accounted for 98.1% of all online activities on its platform.⁷

A recent meta-analysis revealed no significant difference in assessment between video and in-person consultations for mental and behavioral related problems.⁸ Previous research also reported high patient satisfaction with online consultations using video or telephone, as well as being more convenient and less costly.^{9–11} However, we know of no studies that used “standardized patients” (SPs, who are actors trained to simulate patients expressing defined sets of symptoms)

to compare the quality of different modes of consultations. Instead, they involved physicians assessing real patients online or/and in-person, with potential participant biases arising-a “Hawthorne effect”.¹² Undisclosed SPs could overcome the aforementioned problems.

We aimed to compare the quality of care for PPD in asynchronous webchat (online) and in-person consultations, and to explore potential reasons for any inconsistency between online and in-person diagnoses.

Methods

Study design and participants

This was a cross-sectional, paired design study using mixed-methods, registered at the Chinese Clinical Trial Registry (ChiCTR2100045018). All SPs provided written informed consent. No identifiable data of platforms or psychiatrists will be disclosed.

We selected the top ten DTCT platforms ([Appendix Table B.1](#)) with the highest active user engagement, all offering one-on-one consultations for various medical conditions, including depression. Additionally, we included one platform exclusively for mental health conditions due to its significant concentration of specialists in the management of depression. We selected psychiatrists using the following criteria: 1) practicing in provincial capital cities; 2) having at least one online login record on platforms within 6 months of the study period; 3) offering depression-related consultations via asynchronous webchats; 4) working in a hospital department where SPs could have in-person visits. In a

small pilot study, we estimated there might be a 15% difference in diagnostic accuracy between online and in-person visits. In a paired design, a minimum of 57 pairs of online and in-person visits were required.¹³

Procedures

A script describing a woman with signs and symptoms indicative of PPD was developed according to a real PPD case. This was reviewed by a multidisciplinary expert panel (Appendix C). The script was accompanied by a quality checklist (Table C.1 and Table C.2 in Appendix) established in accordance with four guidelines.^{14–17} The standardized patients (SPs) must be childbearing-aged women, in good health and had no mental health symptoms or had history of depression. Five SPs were recruited, and trained for 7 days to consistently simulate the standardized characteristics of our case presentation (Box 1). In both settings, SPs were trained to provide standardized responses to the same questions. The validation processes of PPD case and SPs were detailed in our prior studies, and SPs have been validated (Appendix D) as a quality assessment tool in multiple dimensions.^{18,19} SPs conducted online and in-person visits between June 2021 and April 2022; their status as participants in the research study was not disclosed to the psychiatrists. After each visit, data were recorded by SPs within 30 min and were audited by trained supervisors within 12 h. Cross-quality control checks were conducted among SPs (Appendix D), discrepancies in performance or completion of forms were confirmed with the researcher (LL). The researcher (LL) provided immediate feedback on the quality of SPs. After being secondarily verified by the researcher (LL), the data from the quality checklist were used for results analysis. To mitigate potential carry-over effects, a minimum of 1 week was required between visits to the same psychiatrist, and the order of the two visits was random. If the diagnosis of ‘depression’ or ‘postpartum depression’

was made explicitly at the online or in-person visit, the label of ‘correct diagnosis’ was given to the consultation.

SPs (who remained undisclosed) subsequently conducted a further asynchronous webchat with each of the psychiatrists who made a correct diagnosis in-person but not online to ask four open-ended questions: Q1) I have been to your clinic before, but I found that the diagnoses you gave me in-person differed from the one you provided me online. May I ask why? Q2) Would a telephone or video consultation have been better than an asynchronous webchat for my initial consultation? What about the comparison of telephone/video consultations with in-person visits? Q3) Do you think it is possible to achieve the same quality online and in-person? If the psychiatrist responds that it is not possible, or that online consultation is inferior to in-person, then ask: What do you think the difference is (between the two settings)? Q4) For you, what is the role of online consultations?

Outcomes

We adopted the Institute of Medicine (IOM) definition of quality framework to measure the quality of care (Table 1).²⁰ The primary outcome indicators were 1) diagnosis, and 2) guideline adherence. The secondary outcome indicators were 1) time, 2) cost, and 3) patient-centeredness, measured by the Chinese Version of the Patient Perception of Patient-Centeredness Questionnaire (PPPC-CN).²¹ We also collected basic characteristics of the psychiatrists (sex, job title) and the online visits (including the limits on the number of ‘rounds’ of the asynchronous webchats), and in-person visits (including names of the cities, levels of the hospitals, and basic features of the hospitals).

Data analyses

We used McNemar’s test to compare the proportions of correct diagnoses between the two settings; Wilcoxon signed-rank test for differences in guideline adherence,

Box 1.

Postpartum depression case characteristics

Background: young female (age 23–33), Chinese Han ethnicity, bachelor’s degree, married, worked as a secretary.

Chief complaint: sleep difficulties, appetite loss, and bad mood, lasting a month.

Medical history: at 39 weeks of gestation, the patient gave birth to a healthy baby girl with a history of lateral episiotomy. Two weeks after delivery, the patient began to experience sleep difficulties, decreased appetite, moodiness, and decreased energy. She occasionally cried alone, and felt that life was meaningless, but did not have any suicidal ideation. She has not sought medical help previously. She did not take any medications or experience any traumatic events. The patient had no history of depression and no family history of mental illness.

Disguise/Role-playing Key Points:

- (1) Dressing: Wearing casual and loose-fitting clothes for comfort and warmth. Do not wear any jewelry and are not adorned with makeup.
- (2) Gait: Walking with a slow pace, slightly protruding the belly, resembling remnants of the pregnancy gait shortly after delivery.
- (3) Facial expression: Tired, with a somewhat indifferent gaze.
- (4) Knowledge: Being familiar with perinatal knowledge and the developmental status of infants.

Indicators of IOM	Outcomes	Indicators	Definitions of indicators	Coding	Data measurement
Primary outcomes					
Effectiveness & safety	Diagnosis	Diagnostic accuracy	Diagnoses made by the psychiatrist included "postpartum depression or depression" (correct diagnosis)	0: no 1: yes	Quality checklist for postpartum depression
		Diagnostic Consistency	The diagnoses made by the same psychiatrist online and in-person were concordant (both correct or neither)	0: no 1: yes	Quality checklist for postpartum depression
	Guideline adherence	% of completion of questions involving clinical history recommended by guidelines	Number of questions involving clinical history asked by the psychiatrist during the consultation/all of questions involving clinical history recommended by guidelines ^a	0-1	Quality checklist for postpartum depression
		% of completion of management decisions recommended by guidelines	Number of management decisions made by the psychiatrist during the consultation/all of management decisions recommended by guidelines ^b	0-1	Quality checklist for postpartum depression
Secondary outcomes					
Timeliness	Time	Length of waiting time Online	From the time the SP sent a request for a consultation to the time the psychiatrist first responded; In-person: From the time SP arrived at the clinic to the time when the consultation began	Minutes	Questionnaire for the basic characteristics of visits
		Duration of consultation ^c	From the end of waiting time to the end of consultation	Minutes	Questionnaire for the basic characteristics of visits
Efficiency	Cost	Consultation fee	Online: online consultation fees per package; In-person: appointment fee	CNY ^d	Questionnaire for the basic characteristics of visits
		Overall fee of visit	In addition to the consultation fee, the total medical expenses throughout the visit, including tests and managements. Expenses for transportation and accommodation for in-person visits were not included	CNY ^d	Questionnaire for the basic characteristics of visits
Patient-centredness	Patient-centredness	Scores of PPPC-CN ^e	Mean scores of the overall items and each of the four sub-dimensions of the scale	1-4	PPPC-CN

^aQuestions involving clinical history recommended by guidelines: A total of 23 items (see [Appendix](#)), developed in accordance with guidelines and validated by the multidisciplinary expert panel.
^bManagement decisions recommended by guidelines: A total of five items (see [Appendix](#)), developed in accordance with guidelines and validated by the multidisciplinary expert panel. In this study, an additional management decision online - "referral to in-person" - was collected and reported in the "Comparison of guideline adherence" section of the Result. Only the five management decisions recommended by the guidelines were used for comparative analysis.
^cDuration of consultation: To minimize the effect of the SPs' response speed, the duration between the psychiatrist's question and the next SPs' response was subtracted from the duration of consultation.
^dCNY: Chines Yuan, 1 CNY = 0.15 US dollar.
^ePPPC-CN: The overall Cronbach's alpha value for the 21 items was 0.922. Higher scores indicate less patient-centeredness. Four sub-dimensions of PPPC-CN are: "exploring illness experience", "finding common ground", "understanding the whole person", and "enhancing the patient-clinician relationship".

Table 1: Quality outcome indicators.

cost, and time; and paired-t test for differences of mean scores of PPPC-CN. Odds Ratio (OR) and Cohen's *d* value was used to evaluate the effect size of the difference in diagnostic accuracy and guideline adherence, respectively. Cohen's kappa and Lin's concordance correlation coefficient (CCC) were used to measure agreement for diagnoses and other outcomes between online and in-person consultations, respectively. Criteria for the strength of agreement proposed by Landis et al. and McBride were used.^{22,23} Considering the hypothesis that the missing data are not at random (MNAR), we did not employ imputation for data analysis, and instead conducted a sensitivity analysis. Benjamini-Hochberg false discovery rate (FDR) of 0.05 was applied for correction of p values. All analyses were completed using SPSS (version 19.0) or R (version 4.1.0). The analysis of CCC

was conducted using the DescTools R packages (version 0.99.44) and the graphs were produced using ggplot2 R package (version 3.3.5).

We used inductive thematic analysis for the qualitative data derived from the third interactions involving psychiatrists who offered discrepant diagnoses. Transcripts were imported into Nvivo (version 12.0). Codes were developed initially based on the first two transcripts of these follow-up visits and then refined by three of the authors to ensure that they were appropriate for this study's context. Two authors independently coded each transcript and reached consensus by discussion of variability in coding. Codes were categorized into themes reflecting psychiatrists' comments on the discrepancies and other general issues with online and in-person visits.

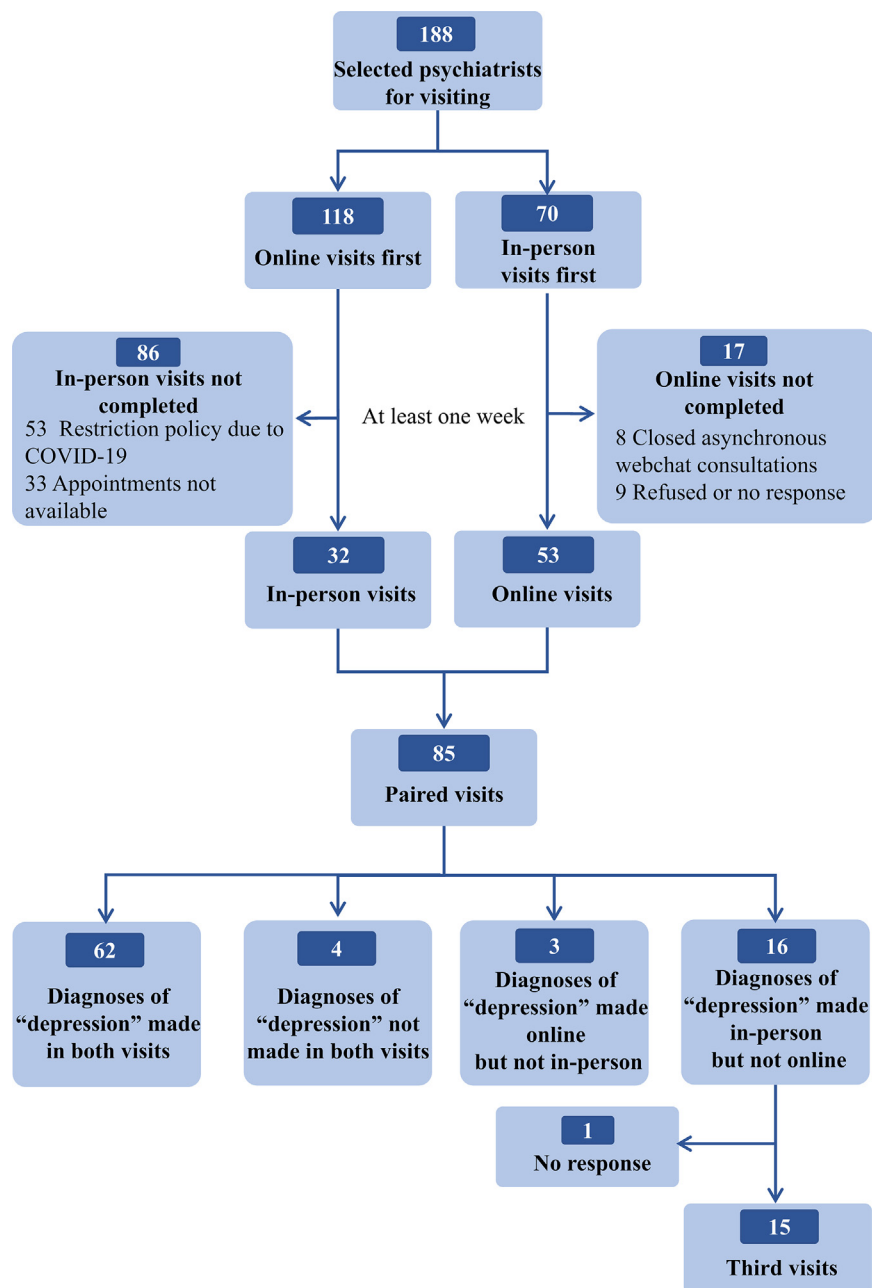


Fig. 1: The flow of visited psychiatrists through the study.

Role of the funding source

The funders of the study had no role in the study design, data collection, data analysis, data interpretation, or writing of the report.

Results

Basic characteristics of visits

A total of 171 online visits via asynchronous webchats on 11 platforms and 102 in-person visits at 83 hospitals of 10

capital cities (two each in southern, southwestern, south-eastern, and northeastern China, one in northern, and one in central) were conducted. A total of 188 psychiatrists were involved, 85 of whom were visited both online and in-person (in 69 hospitals). Due to pandemic restrictions and other factors beyond our control, we failed to conduct 17 online and 86 in-person visits (Fig. 1 and Table 2).

Five of the 11 platforms had a time limit of 24 h for the asynchronous webchats, with a limit of 48 h for the

Variables	Online visits only (n = 86)		In-person visits only (n = 17)		Paired visits (n = 85)	
	N	%	N	%	N	%
Characteristics of psychiatrists						
Sex						
Male	44	51.2	3	17.6	49	57.6
Female	42	48.8	14	82.4	36	42.4
Job title						
Attending physicians or resident physicians	41	47.7	7	41.2	35	41.2
Associate chief physicians	26	30.2	6	35.3	30	35.3
Chief physicians	19	22.4	4	23.5	20	23.5
Characteristics of online visits						
Limitation on the number of 'rounds'						
No limitation	57	66.3	NA	NA	54	63.5
3–30 rounds	29	33.7	NA	NA	31	36.5
Characteristics of in-person visits						
Names of cities						
Tier-1 cities ^a	NA	NA	6	35.3	21	24.7
Tier-2 cities	NA	NA	11	64.7	64	75.3
Levels of the hospitals						
Tier-3 hospitals	NA	NA	14	82.4	73	85.9
Tier-2 hospitals and below	NA	NA	3	17.6	12	14.1
Basic features of hospitals						
Specialist hospitals	NA	NA	5	29.4	21	24.7
General hospitals	NA	NA	12	70.6	64	75.3

^aNames of cities: according to the publication of the National Bureau of Statistics of China.

Table 2: Basic characteristics of visits.

remaining six. Five of the platforms also had limits on the number of rounds of chats between patients and doctors, ranging from 3 to 30 rounds. All 11 platforms provided online psychological tests (e.g., Patient Health Questionnaire-9), which were used to assess the mental states of SPs in 31 of 85 online visits (36.5%). Fifty-eight of the 69 hospitals visited were Tier-3 hospitals (which tended to be large hospitals). Only one out of 85 in-person visits (1.2%) involved health professionals during the mental status assessments, which were mostly completed by SP themselves on computers (52 of 85 in-person visits; 61.2%). Of note, none of the online or in-person consultations sought previous medical records

from the SPs. There was no evidence that the identity of the SPs was exposed during any of the consultations.

Comparison of diagnoses

The proportion of correct diagnoses was significantly lower following online consultation than in-person evaluation (76.5% [95% CI 67.5%–85.5%] vs 91.8% [95% CI 85.9%–97.6%]; $p_{FDR} = 0.0066$; OR = 3.429 [95% CI 1.36–8.616] (Table 3; Fig. 2). There were 66 paired visits with concordant diagnoses (including no diagnosis of “depression” in both settings) and 62 paired visits with correct diagnoses in both settings, with a Cohen’s kappa value of 0.199 (95% CI –0.026 to 0.424; $p_{FDR} = 0.029$). Sixteen psychiatrists offered diagnoses of “depression” in-person but not online, while three diagnosed “depression” online but not in-person (Fig. 1).

Of the 20 online visits and seven in-person visits where “depression” was not diagnosed, psychiatrists primarily identified “sleep problems,” “mood problems,” and “anxiety;” three online visits were completed without any diagnosis. Alternative diagnoses are shown in Table 4.

Comparison of guideline adherence

The proportion of completed questions relevant to medical history as recommended by published

	Diagnosis made in-person		Total	Statistics of McNemar’s test	p value _{FDR}
	Yes	No			
Diagnosis made online				7.579	0.0066
Yes	62	3	65		
No	16	4	20		
Total	78	7	85		

Table 3: The correct diagnoses of paired visits.

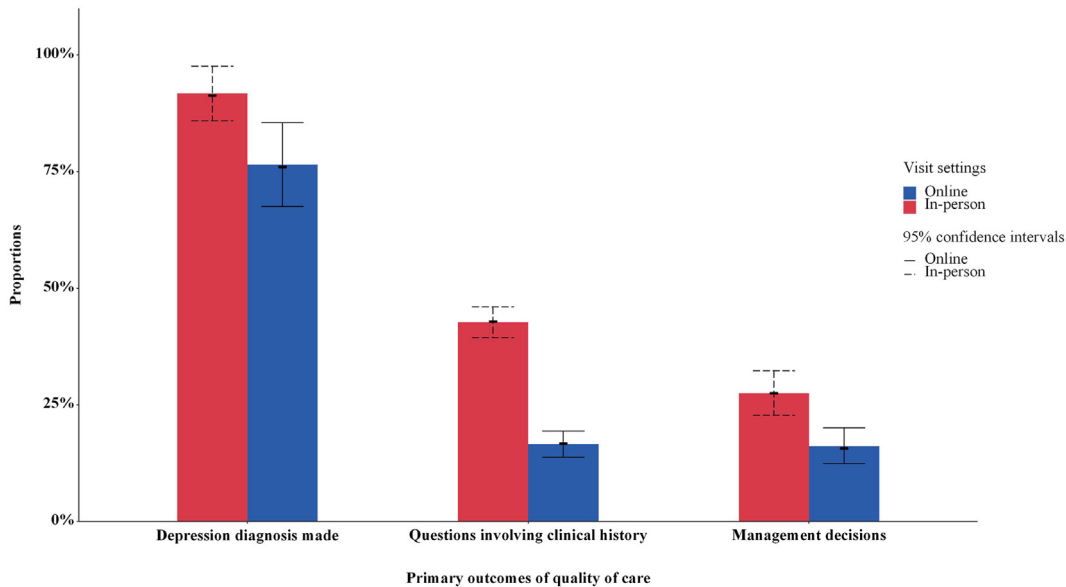


Fig. 2: The comparisons of whether the correct diagnosis was made and guideline adherence online and in-person.

guidelines was lower online than in-person (16.6% [95% CI 13.8%–19.4%] vs 42.7% [95% CI 39.4%–46.0%]; $p_{FDR} < 0.0001$); with a Cohen’s *d* value of 1.83 [95% CI 1.49–2.15] (Fig. 2). The only exception pertained to “substances or medicine use” (Fig. 3). The top three completed questions for both online and in-person was “duration of symptoms” (online 61.2% [52/85] vs in-person 94.1% [80/85]), “depressed mood” (online 51.8% [44/85] vs in-person 90.6% [77/85]), and “sleep quality” (online 45.9% [39/85] vs in-person 90.6% [77/85]). The question completed least online was “attitudes towards pregnancy” (2.4% [2/85]); for in-person, “suicidal behaviors” (8.2% [7/85]) was completed least. The four questions with the greatest discordance for completion were “sleep quality,” “domestic violence,” “depressed mood,” and “hopeless,” with proportion differences of completion of 44.7% (38/85), 40.0% (34/85), 38.8% (33/85), and 38.8% (33/85), respectively.

The proportion of completed management decisions recommended by guidelines was lower online than in-person (16.2% [95% CI 12.4%–20.1%] vs 27.5% [95%

CI 22.8%–32.3%]; $p_{FDR} < 0.0001$); with a Cohen’s *d* value of 0.56 [95% CI 0.25–0.86] (Fig. 2). This was true for all relevant items. “Health education” was the most common management decision for both online and in-person visits (55.3% [47/85] vs 34.1% [29/85]; Fig. 4). The least often completed management decision online was “further follow up” (2.4% [2/85]), vs 34.1% (29/85) in-person, whereas 64.7% (55/85) of online visits were given the “referral to in-person.” The least frequent management decision in-person was “assessment of suicidal risk or social support” (8.2% [7/85]), which was completed online in only 3.5% (3/85) of the consultations. “Further follow up” was the management decision with the greatest discordance for completion (31.8% [27/85]).

Comparison of time and cost

The length of waiting time and the duration of consultation were significantly longer online than in-person (median length 37 min [interquartile range, IQR 4.5–210.5] vs 8 min [1–20]; $p_{FDR} < 0.0001$; median duration 737 min [IQR 94.5–1535] vs 9 min [7–14]; $p_{FDR} < 0.0001$). Online consultation fees for were higher than those for in-person visits (median 30 CNY [IQR 20–80] vs 16 CNY [10–48.5]; $p_{FDR} = 0.00038$). Seventy-seven (77) mental health scales were used in 31 (36.5%) visits online, while 226 scales were used in 53 (63.4%) visits in-person. SPs (11.8%) were given prescribed medications and/or other treatments after ten in-person visits while none were prescribed online. Combining the consultation fees with tests and managements, the overall fee of the visits was cheaper online than in-person (median 30 CNY [IQR 20–80] vs 120 CNY [15–284.9]; $p_{FDR} < 0.0001$).

Diagnoses	Online (n = 20)	In-person (n = 7)
Sleep problems	11	2
Mood problems	7	4
Anxiety	6	3
Neurological problems	1	0
Not naming any diagnosis	3	0

Multiple diagnoses may be made by the same psychiatrist.

Table 4: Alternative diagnoses made by the psychiatrists.

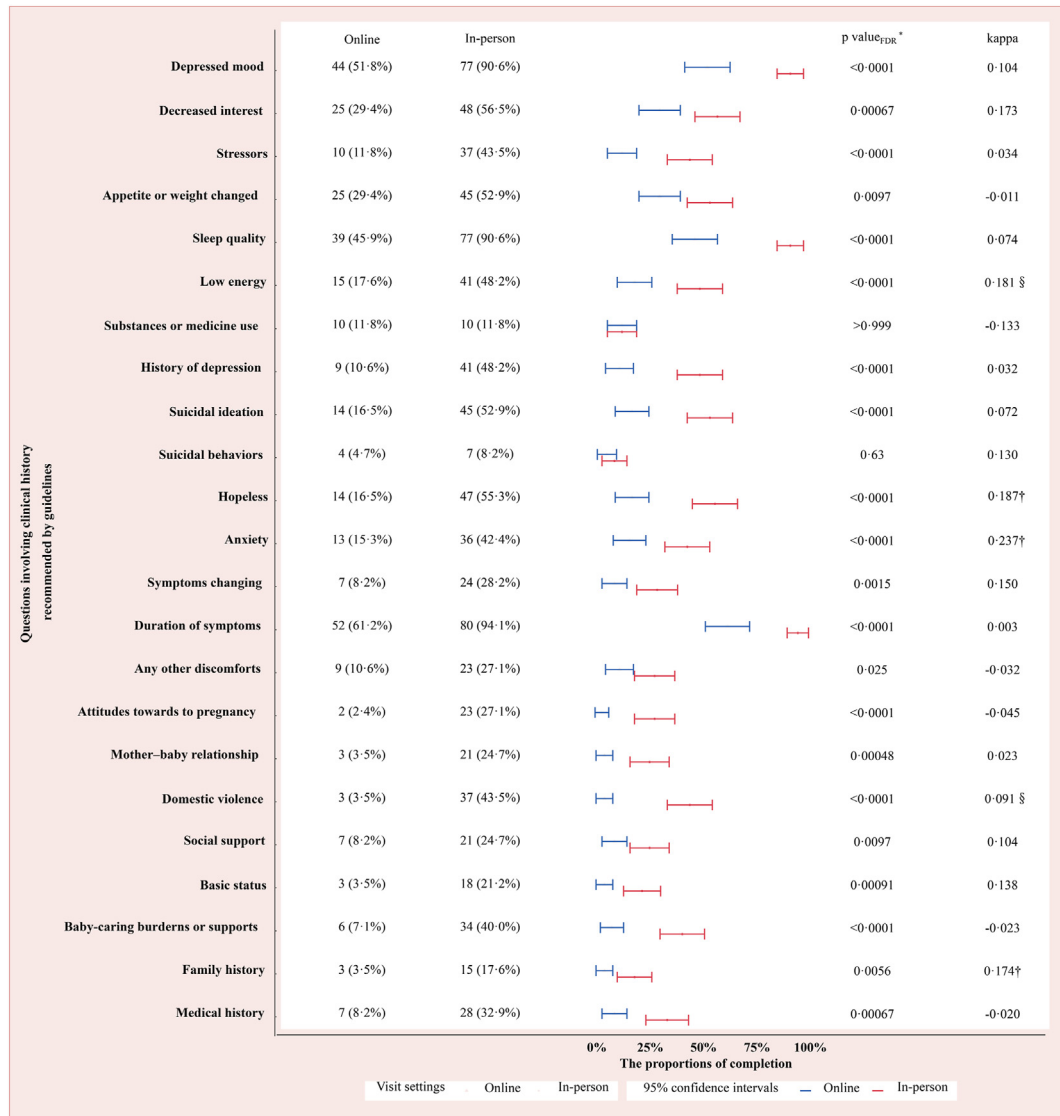


Fig. 3: The proportions of completion of questions involving clinical history recommended by guidelines online and in-person.†p value of Cohen’s kappa remained significant (<0.050) after correction of false-discovery rate method. §Unadjusted p value <0.050.

Comparison of patient-centeredness

Indicative of a lower quality of patient-centeredness online, the mean overall score of PPPC-CN for online visits and the mean overall scores of each of four sub-dimensions were higher than in-person encounters (p_{FDR} < 0.0001; Fig. 5).

Results of sensitivity analysis

No marked differences were found between the results of completed paired data and imputed data assuming missing at random, with only the significance test of difference for the two indicators - the proportion of correct diagnoses and consultation fee - were not significant under different offset parameter δ (delta), representing varying MNAR conditions (Appendix Table E.1).

The results of CCC

The CCC values for all continuous outcomes in quality of care for both online and in-person settings were below 50% (Fig. 6). The highest point estimate being the CCC value for the mean scores of PPPC-CN (28.5% [95% CI 11.3%–44.0%]); the lowest point estimate was the CCC value for the duration of waiting for consultation (-0.6% [95% CI -4.1% to 2.9%]).

Qualitative results regarding the psychiatrists who did not make the diagnosis of PPD online but did so in-person

Fifteen of 16 psychiatrists who correctly diagnosed SPs in-person but not online completed the third interview via webchats. Eleven were male; one was a resident

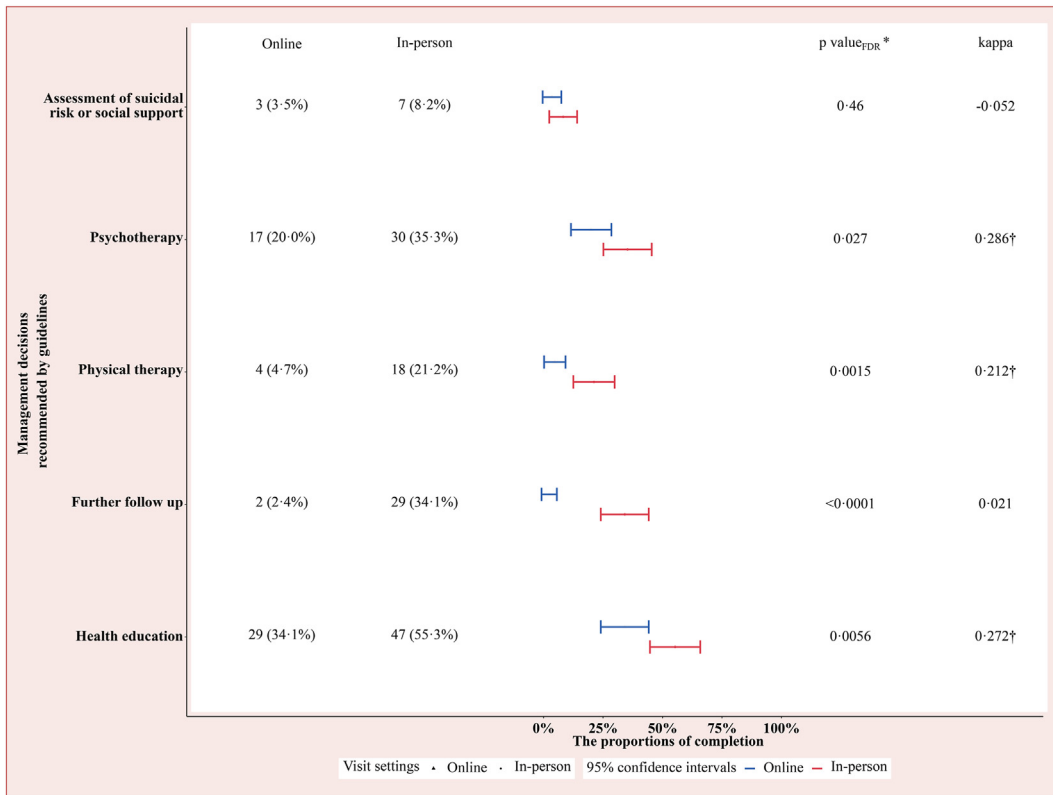


Fig. 4: The proportions of completion of management decisions recommended by guidelines online and in-person. †p value of Cohen’s kappa remained significant (<0.050) after correction of false-discovery rate method.

physician, seven were attending physicians, four were associate chief physicians, and three were chief physicians.

Thirteen of the 15 psychiatrists regarded the discrepancies between online and in-person diagnoses as a result of a lack of non-verbal information (e.g.,

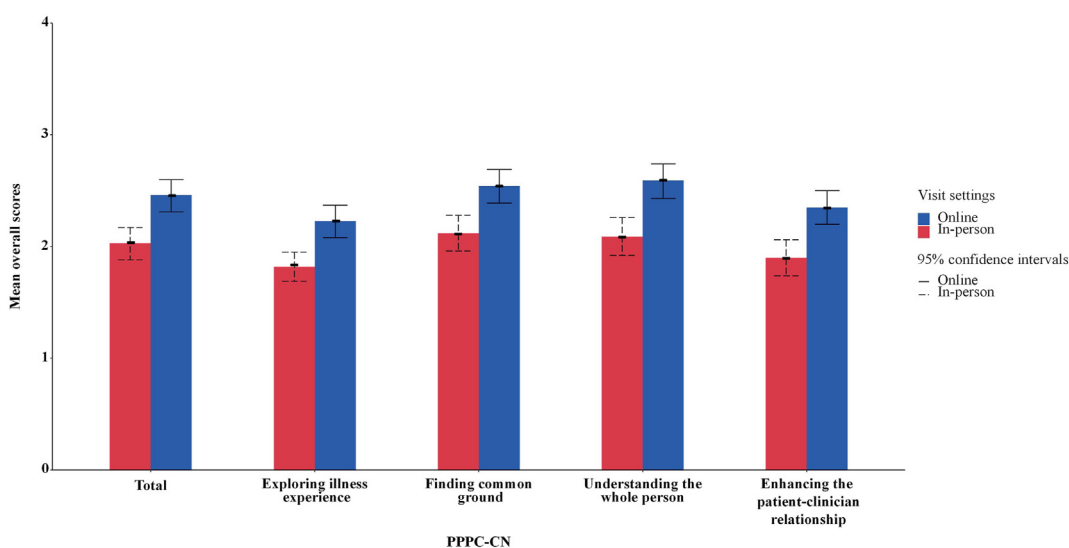


Fig. 5: The mean overall scores of PPPC-CN and of each of the four sub-dimensions for online and in-person visits.

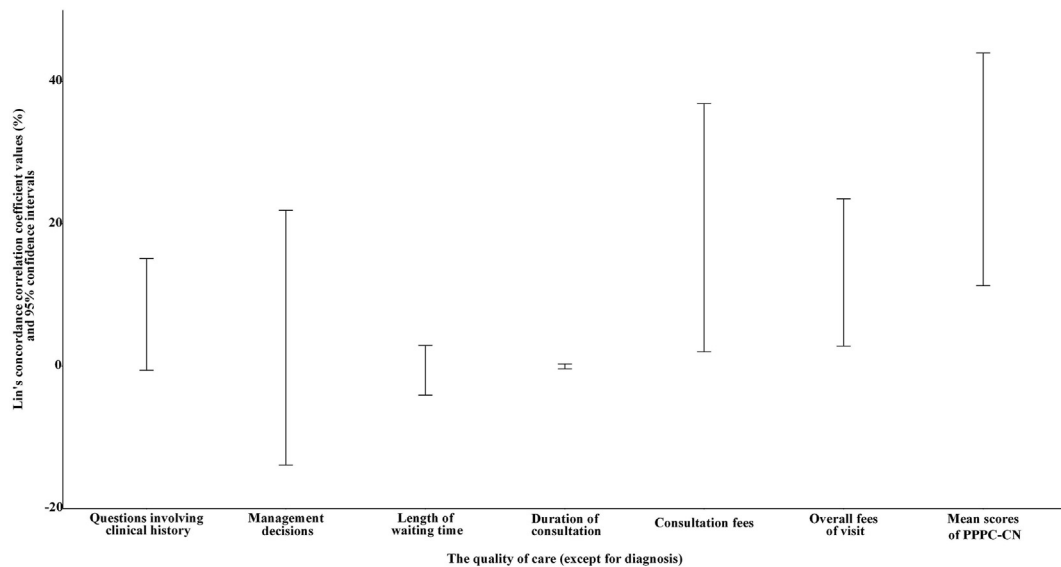


Fig. 6: The Lin's concordance correlation coefficient values and 95% confidence intervals of quality of care (except for diagnosis) online and in-person.

expressions, the tone and speed of voice) online compared to in-person.

An illustrative comment by one psychiatrist: “The observations and information online are relatively simple, resulting in conservative diagnoses.”

Another noted: “Definitely the most accurate (diagnosis) is (made) in-person consultation, because we can communicate more thoroughly and in-depth. We can see you and have a face-to-face interaction with each other, which is very important. In addition, in fact, we can also notice your look, your speech and expressions, all of which are non-verbal information that is crucial to diagnosis and assessment.”

Eight psychiatrists believed video or telephone consultations were better than asynchronous webchat consultations, though they would still be inferior to in-person evaluations.

One psychiatrist described time constraints of online consultations: “It (video or telephone consultation) is actually limited to 15 min, meaning that it will automatically cease after 15 min, which is very unfavorable for us psychiatrists.”

The majority of psychiatrists (13/15) suggested that online consultation could not replace in-person visits as the latter was more effective, although two psychiatrists offered different perspectives: “Online visits would be able to achieve similar effectiveness as in-person visits if the medical history and examination results are uploaded.”

“Most of the time, the effectiveness of online and in-person consultations is similar. They are simply distinct approaches. [...] It (the effectiveness online and in-person) is the same! The keys are clinicians' professionalism and

competence. If the clinician is responsible, he or she will be attentive whether the consultation is online or in-person!”

Twelve of the 15 psychiatrists expressed the view that online consultations served different purposes from in-person assessments. Online consultation should be used more for repeat prescriptions, follow-up or purposes such as providing only brief cognitive therapy. Ten psychiatrists mentioned that the initial diagnostic assessment should be done in-person.

“Generally speaking, first diagnoses should be made in-person. At in-person consultations, the observation of your status when you describe symptoms, in addition to the examination at the time, will be more in line with diagnostic and treatment procedures; the precision and accuracy of the diagnosis and treatment will be better. In online consultations, I cannot view your status, only the text. I primarily deliver health services in-person; online is solely used for follow up. If adjustments to management are needed after the diagnosis is made and treatment prescribed at in-person visits, but in-person consultations are inconvenient due to pandemic restrictions, you choose online consultations.”

Another stated: “Generally, I recommend in-person consultation for the first diagnosis. When things get better over time, online and in-person consultations are similar. [...] Online is mainly used for health counseling, patient management, prescription of medication, and so on. For these purposes, if it is not convenient for you to attend in person, you can consult online.”

Discussion

To the best of our knowledge, this study represents the first attempt to employ SPs to consistently evaluate the quality of care provided by the same psychiatrists in both online and in-person settings. We found that for SPs “suffering” from PPD, online consultations utilizing asynchronous webchats were inferior to in-person consultations in terms of diagnostic accuracy, adherence to established guidelines, and patient-centeredness. By incorporating a diverse cohort of 85 psychiatrists from 11 prominent DTCT platforms and 69 hospitals across ten provincial capital cities in China, results from this study give a glimpse of the current landscape of psychiatric consultations for PPD.

We found that sixteen out of 85 psychiatrists who refrained from providing an online diagnosis of PPD did so when the SPs were seen in-person. Attributing the online diagnostic errors to the competence of psychiatrists may therefore be inappropriate. Several factors may have contributed to the differences between the two modes of consultation. Firstly, nonverbal cues often play a critical role in establishing clinical diagnoses of mental disorders.²⁴ Secondly, certain psychiatrists may perceive online consultations primarily as preliminary evaluations, with the expectation of referring patients for in-person consultations if necessary, rather than making definitive diagnoses online.²⁵ For these clinicians, making a diagnosis might have been considered a secondary objective. Indeed, ten out of the sixteen psychiatrists suggested referrals for in-person visits. Their approach to online diagnosis may also stem from limited access to comprehensive information.²⁶ A majority of these sixteen psychiatrists diagnosed SPs seen online with broad terms such as “sleep problems”, “mood disorders” and “anxiety” instead of specifying “depression” as the primary diagnosis. Additionally, some psychiatrists appeared to have reservations about online consultations, perceiving them as inferior to in-person encounters, which could potentially impact their confidence and performance. However, they were included in the sample due to their apparent availability for online consultations.

The study revealed that clinical inquiries recommended in published guidelines were more consistently addressed in-person compared to online consultations. Certain inquiries, such as “suicidal ideation” and “domestic violence” were substantially less frequently explored in online settings, though in-person consultations also missed key areas of clinical investigation. Psychiatrists’ reluctance to discuss such sensitive topics through text-based asynchronous exchanges, conducted over extended periods of up to 48 h, might contribute to this discrepancy. The format of asynchronous webchats appeared to hinder the continuity, flow, coherence, and comprehensiveness of conversations, potentially leading to the omission of important but sensitive subjects.²⁷ This contradicts the commonly perceived advantages

of online consultations, which include reducing stigma and facilitating the discussion of sensitive matters.²⁸ It is noteworthy that our study did not include consultations conducted through video or telephone, with asynchronous webchat being the predominant mode employed in China at the time.

While our investigation focused exclusively on PPD, the results underscore the need for an urgent reevaluation of the appropriateness of using asynchronous webchats for initial consultation encounters. The restrictions imposed by platforms on the number of rounds and total duration of consultations warrant review. Efforts should be directed towards determining which clinical conditions or symptoms are suitable for asynchronous webchats or other forms of online consultations. Clear guidelines for referral to in-person visits and mechanisms for monitoring the quality of online consultations are imperative.²⁹ Furthermore, systematic training must be provided to healthcare providers and students to equip them better for online consultations.

This study has limitations. The data collection occurred during the COVID-19 pandemic when in-person visits were restricted at times. Additionally, we could not obtain a comprehensive sampling frame of all psychiatrists engaged in webchat consultations. No further consultations were conducted with the three psychiatrists who correctly diagnosed them online but not in-person, potentially leading to a loss of information.

In conclusion, our findings demonstrate that online consultations using asynchronous webchats for the evaluation and diagnosis of PPD are inferior to in-person consultations concerning diagnostic accuracy, adherence to guidelines, and patient-centeredness. To harness the potential benefits of online consultations and to ensure patient safety, urgent and substantial improvements in their quality are imperative.

Contributors

WJG conceived the study and designed it with KKC and EDC. LL and XYL were responsible for the revision of PPD script and the training and validation of SPs. WJG and LL were responsible for data analysis, manuscript preparation and writing. KKC and EDC critically reviewed draft manuscripts. JCS contributed to the guidance of data analysis and the revision of the manuscript. LL and ZZ provided valuable expertise regarding SPs implementation and data collecting processing. XYL was responsible for literature searching, data quality supervision and manuscript preparation.

Data sharing statement

Access to de-identified data is available on reasonable requests from the corresponding author.

Declaration of interests

All authors declare no competing interests.

Acknowledgements

This work was supported by the China Medical Board (CMB) Open Competition Program (Grant No. 20-368), National Natural Science Foundation of China (Grant No. 82273643 and No. 81973059), and

Swiss Agency for Development and Cooperation Global Cooperation Department (No. 81067392). The authors gratefully acknowledge the five SPs for participating in this study and the key advice for establishing the script and quality checklist for PPD from psychiatric specialists (Ming Cheng, Yihua Chen, Wenjuan Jiang, Mei Liao, Yongsheng Tong, Li Zhang). We are grateful for the support of Dong Xu for the design of the study and the revision of the first draft. We thank the assistance of specialists in epidemiology and health statistics (Yiyuan Cai and Jing Deng) with study design and statistical analysis, as well as the advice for the plan of data analysis from Wenjun He. We also thank the efforts of Yiqing Cai and He Zhang in information searching and schedules of visits making.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.lanwpc.2024.101053>.

References

- Nisar A, Yin J, Waqas A, et al. Prevalence of perinatal depression and its determinants in Mainland China: a systematic review and meta-analysis. *J Affect Disord.* 2020;277:1022–1037.
- Hahn-Holbrook J, Cornwell-Hinrichs T, Anaya I. Economic and health predictors of national postpartum depression prevalence: a systematic review, meta-analysis, and meta-regression of 291 studies from 56 countries. *Front Psychiatry.* 2017;8:248.
- Bina R. Predictors of postpartum depression service use: a theory-informed, integrative systematic review. *Women Birth.* 2020;33:e24–e32.
- Der-Martirosian C, Chu K, Steers WN, et al. Examining telehealth use among primary care patients, providers, and clinics during the COVID-19 pandemic. *BMC Prim Care.* 2022;23:155.
- General Office of the National Health Commission. Notice of the General Office of the National Health Commission on exploring the implementation of featured services for the prevention and treatment of depression and senile dementia. <http://www.nhc.gov.cn/jkj/s7914/202009/a63d8f82eb53451f97217bef0962b98f.shtml>; 2020. Accessed October 3, 2021.
- Carrillo de Albornoz S, Sia K-L, Harris A. The effectiveness of teleconsultations in primary care: systematic review. *Fam Pract.* 2021;39:168–182.
- The '2021 good doctor list' was released. Peoples life weekly. <https://baijiahao.baidu.com/s?id=1722806405687957748&wfr=spider&for=pc>; 2022. Accessed September 18, 2022.
- Batastini AB, Paprzycki P, Jones ACT, MacLean N. Are video-conferenced mental and behavioral health services just as good as in-person? A meta-analysis of a fast-growing practice. *Clin Psychol Rev.* 2021;83:101944.
- Chakrabarti S. Usefulness of telepsychiatry: a critical evaluation of videoconferencing-based approaches. *World J Psychiatry.* 2015; 5:286–304.
- Lyneham HJ, Rapee RM. Agreement between telephone and in-person delivery of a structured interview for anxiety disorders in children. *J Am Acad Child Adolesc Psychiatry.* 2005;44:274–282.
- Hajebi A, Motevalian A, Amin-Esmaeli M, et al. Telephone versus face-to-face administration of the structured clinical interview for diagnostic and statistical manual of mental disorders, fourth edition, for diagnosis of psychotic disorders. *Compr Psychiatry.* 2012;53:579–583.
- Leonard KL. Is patient satisfaction sensitive to changes in the quality of care? An exploitation of the Hawthorne effect. *J Health Econ.* 2008;27(2):444–459.
- Yang T. *Medical science research and design.* 2nd ed. Beijing, China: People's Medical Publishing House; 2013:154.
- National Collaborating Centre for Mental Health (UK). Antenatal and postnatal mental health: clinical management and service guidance. <https://www.nice.org.uk/guidance/cg192>; 2014. Accessed April 11, 2020.
- Matthews-Smith G. SIGN 127 management of perinatal mood disorders: a national clinical guideline. SIGN. <http://researchrepository.napier.ac.uk/id/eprint/8484>; 2012. Accessed April 11, 2020.
- Royal Australian College of General Practitioners. *Guidelines for preventive activities in general practice.* 9th ed. South Melbourne: The Royal Australian College of General Practitioners; 2016. <https://www.racgp.org.au/download/Documents/Guidelines/Redbook9/17048-Red-Book-9th-Edition.pdf>. Accessed April 11, 2020.
- Ding H, Chen L, Di X. Expert consensus on guidelines for the management of postpartum depressive disorder (based on obstetric and community physicians). *Chin J Clin Obstet Gynecol.* 2014; 15:572–676.
- Xu DR, Hu M, He W, et al. Assessing the quality of primary healthcare in seven Chinese provinces with unannounced standardised patients: protocol of a cross-sectional survey. *BMJ Open.* 2019;9:e023997.
- Xu DR, Cai Y, Wang X, et al. Improving data surveillance resilience beyond COVID-19: experiences of primary health Care Quality Cohort In China (ACACIA) using unannounced standardized patients. *Am J Public Health.* 2022;112:913–922.
- Institute of Medicine (US) Committee on Quality of Health Care in America. *Crossing the quality chasm: a new health system for the 21st century.* Washington (DC): National Academies Press (US); 2001:39–40.
- Cai YY, Guo PF, Tu J, et al. Contextualizing the revised Patient Perception of Patient-Centeredness (PPPC-R) scale in primary healthcare settings: a validity and reliability evaluation study. *BMC Prim Care.* 2024;25(1):11.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159–174.
- McBride GB. A proposal for strength-of-agreement criteria for Lin's concordance correlation coefficient. 2005 NIWA client report: HAM2005-062. <https://www.medcalc.org/download/pdf/McBride2005.pdf>; 2005. Accessed April 11, 2022.
- Hall JA, Horgan TG, Murphy NA. Nonverbal communication. *Annu Rev Psychol.* 2019;70:271–294.
- Xu D, Zhan J, Cheng T, Fu H, Yip W. Understanding online dual practice of public hospital doctors in China: a mixed-methods study. *Health Policy Plan.* 2022;37:440–451.
- He D, Gu Y, Shi Y, Wang M, Lou Z, Jin C. COVID-19 in China: the role and activities of Internet-based healthcare platforms. *Glob Health Med.* 2020;2:89–95.
- Li H, Kraut RE, Zhu H. Technical features of asynchronous and synchronous community platforms and their effects on community cohesion: a comparative study of forum-based and chat-based online mental health communities. *J Comput Mediat Commun.* 2021; 26:403–421.
- Vadlamani LN, Sharma V, Emani A, Gowda MR. Telepsychiatry and outpatient department services. *Indian J Psychol Med.* 2020;42:275–335.
- Lum E, van Galen LS, Car J. Competency-based training for entrustment in telehealth consultations. *Pediatr Clin North Am.* 2020;67:735–757.