STUDY PROTOCOL

Development and validation of the health-related quality of life instrument for Chinese infertile couples: a mixed-methods study protocol

Zhao Shi[1,2,3], Zhuxin Mao[4], Hongwei Nie[5], Ling Geng[5], Gang Chen[6] and Shunping Li[1,2,3*]

Abstract

Background: Infertility and its treatment have negative impacts on a couple's marital relationship, sexual life, psychological state and interpersonal relationships, causing personal distress. Health-related quality of life (HRQoL) has become an important component of health outcomes. HRQoL instruments developed in western culture are not always appropriate for use in China due to cultural differences. Probably due to the unique concept of fertility in China, infertility patients can be looked down upon and the family may feel shameful. This study aims to develop a HRQoL instrument for infertile couples based on the Chinese social and cultural setting.

Methods: Complementary mixed methods will be used to develop a new HRQoL instrument for Chinese infertile couples. The study consists of four stages: the first stage will involve a systematic review and qualitative interviews to construct draft candidate items. In the second stage, quantitative research [e.g., exploratory factor analysis (EFA), item response theory (IRT)] and cognitive interviews will be used for item selection. The third stage will be instrument validation, in which classical test theory (CTT) and IRT will be applied. In the final stage, the minimal clinically important difference (MCID) will be calculated by using distribution-based methods and anchor-based methods (e.g., logistic regression, receiver operating characteristic curve).

Discussion: The new HRQoL instrument for Chinese infertile couples will be developed, which will provide a standard and effective HRQoL instrument in clinical outcome assessment and health outcome measurement.

Keywords: Infertile couples, Health-related quality of life, Instrument, Validation, Chinese, Protocol

Background

Infertility is defined as the inability to conceive after at least 12 months of unprotected intercourse [1–3]. The infertility rate of couples of childbearing age in China has risen from 2.5 to 3% to around 12 to 15% in 20 years, and the number of patients has exceeded 50 million [4–6].

Infertility and its treatment have negative impacts on a couple's marital relationship, sexual life, psychological state and interpersonal relationships, causing personal distress [7–9]. With the increasing aging trend in China, infertility has also become an important public health and social problem [4, 5]. In order to raise the fertility level, the three-child policy has been implemented in China [10].

Health-related quality of life (HRQoL) is a multi-dimensional concept that represents the patient's overall perception of the impact of an illness and its treatment. It includes an individual's physical,

*Correspondence: lishunping@sdu.edu.cn
[1] Centre for Health Management and Policy Research, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan 250012, China
Full list of author information is available at the end of the article

Shi *et al. Health and Quality of Life Outcomes* (2022) 20:54

Page 2 of 8

psychological, social aspects of life and has become an important component of health outcomes [11]. Since HRQoL is a culturally relevant concept, people from different cultural backgrounds tend to have different understandings of HRQoL [12, 13]. Western HRQoL instruments are not always appropriate for use in China due to cultural differences between the East and West [13, 14]. It is worth mentioning that there are fundamental differences between Chinese and Western fertility concepts. Infertility patients in China seem to be more influenced by the traditional Chinese saying, "There are three forms of unfilial conduct, of which the worst is to have no descendants" [15], meaning that it is unacceptable for people not to have offspring. Accordingly, infertility patients can be looked down upon and their families may feel shameful [16, 17]. Women tend to be especially subject to more blame and prejudice than men in the Chinese traditional cultural setting [16].

Disease-specific HRQoL instruments have been developed among infertility patients in previous studies [9, 18, 19]. The fertility quality of life (FertiQoL) scale is currently the most commonly used disease-specific HRQoL scale for infertile couples [18, 20]. However, empirical studies have shown that some items in the FertiQoL are poorly understood in a Chinese cultural setting and irrelevant to the Chinese health service system [21]. Generic HRQoL instruments such as EuroQol-5 dimensions-5 levels (EQ-5D-5L), Medical Outcomes Study 36-Item Health Survey (SF-36), WHO Quality of Life Measure (WHOQOL-BREF) have been frequently used to assess HRQoL among infertility patients [9, 22, 23]. However, the generic measures are not able to include important disease-specific domains, which are useful to detect clinical changes specific to infertility. For example, when EQ-5D-5L was used to measure the HRQoL of infertile patients in China, it was found that the health dimensions of EQ-5D-5L were insufficient, causing a seriously high "ceiling effect" [23].

To the best of our knowledge, there was only one infertility quality of life instrument developed in China [24]. It focused on those female patients with liver depression and was based on the theory of traditional Chinese medicine, which limited their application scope. More importantly, the development of the instrument did not include qualitative interviews with infertility patients.

Objectives

This study aims to develop a de novo HRQoL instrument for infertile couples in China and evaluate its psychometric properties, with the hope of providing a validated HRQoL instrument for the Chinese population.

The specific objectives of the study are :

1. Develop a de novo HRQoL instrument for Chinese infertile couples;
2. Evaluate the reliability, validity, and sensitivity/responsiveness of the instrument;
3. Develop the minimal clinically important difference (MCID) of the instrument and provide thresholds for the interpretation of the results.

Methods

Research design

In this study, the new HRQoL instrument development process will refer to the patient-reported outcome (PRO) measures development guidelines recommended by the United States Food and Drug Administration (FDA) [11]. The item identification and domain construction of the instrument will refer to the Wilson-Cleary revised model of HRQoL theory [25, 26], a mixed methods (quantitative and qualitative) design will be used.

This study consists of four stages to be implemented sequentially. In the first stage, the item pool will be constructed using a top-down approach based on a systematic review, supplemented with a bottom-up method based on qualitative interviews between clinicians and infertility patients [27]. In the second stage of the item selection, classical test theory (CTT) [e.g., exploratory factor analysis (EFA)] and item response theory (IRT) will be used to select items [28]. Meanwhile, cognitive interviews will be used to test the rationality and comprehensiveness of the items. In the third stage of instrument validation, CTT and IRT will be used to evaluate psychometric properties [28]. In the final stage, the MCID of the HRQoL instrument for infertility patients will be developed in combination with clinical data by using anchor-based methods and distribution-based methods (e.g., logistic regression, receiver operating characteristic (ROC) curve) [28, 29]. The development process of

Fig. 1 The development process of Chinese infertile couples HRQoL instrument
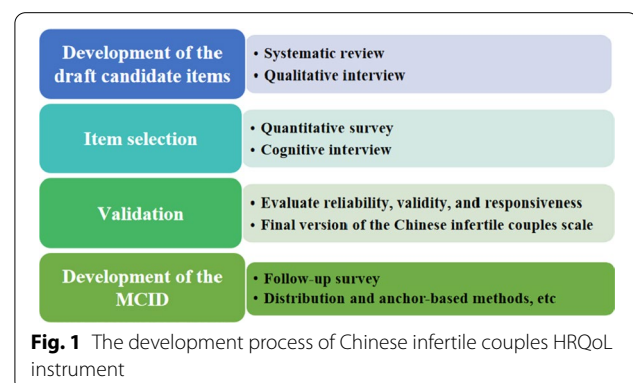
Chinese infertile couples HRQoL instrument is shown in Fig. 1.

Study samples and recruitment

One reproductive medicine center will be selected in each of the seven regions of the Northeast, Northwest, North, Central, East, Southwest and South China according to geographical location and economic development level in China (Fig. 2). Convenience sampling and purposeful sampling methods will be used to select participants, who receive different treatments, including drugs treatment, surgical treatment, artificial insemination (AI), gamete transplantation, in vitro fertilization-embryo transfer (IVF-ET).

The following inclusion criteria will be applied to determine the eligibility of participating patients: (1) being clinically diagnosed with infertility, that is the inability to conceive after at least 12 months of unprotected intercourse, including primary or secondary infertility patients, male and female infertility patients; (2) giving informed consent; (3) without cognitive impairment.

Stage 1: Draft candidate items

Systematic review

The systematic review will identify a wide scope of studies investigating the either disease-specific or generic quality of life of individuals with fertility problems. To identify such studies, eight databases will be searched, including PubMed, Web of Science, Cochrane Library, Embase, and including the most widely used database for articles written in Chinese, China National Knowledge Infrastructure (CNKI), WanFang Data, China Science and Technology Journal Database (CSTJ), and China Biology Medicine disc (CBMdisc). Key words included in the search strategy will include: infertility (sterility); patient-reported outcomes; quality of life

Fig. 2 The seven reproductive medicine centers will be selected in China

Shi *et al. Health and Quality of Life Outcomes*   (2022) 20:54

Page 4 of 8

(health related quality of life); health state preference; quality-adjusted life year, etc. Search filters will include the language of English and Chinese.

Qualitative interview

Design and sample One-to-one semi-structured in-depth interviews will be held in a private room. In order to diversify the sample, the interviewees will be selected by epidemiological characteristics, such as gender, age group, the determinism of etiology [5, 30], and social demographic characteristics, such as education level [31]. Ten patients will be initially recruited in the Hospital for Reproductive Medicine Affiliated of Shandong University for pilot testing to evaluate whether the patients can understand the key questions of the interview protocol and can respond to the questions. The interview guide will be revised accordingly based on the patients' feedback. It's estimated that a total of about 70 infertility patients will be interviewed in seven reproductive medicine centers (Fig. 2). Furthermore, it is estimated that one gynecological clinician, male clinician and psychology/nursing experts will be randomly selected to conduct qualitative interviews at each reproductive medicine center. Three experienced qualitative interviewers will be trained to be competent to undertake semi-structured interviews on clinical reproductive medicine and sensitive topics in seven reproductive medicine centers. An interview guide will be developed to ensure consistency throughout the process.

Zhihu is an open access and knowledge-sharing Chinese online community. It contains a large number of electronic diaries of patients' treatment experiences and psychological processes. Considering that fertility is a relatively private topic, the infertility patient's electronic diary recorded on *Zhihu* will be used as additional and supplementary material to qualitative interviews [32]. This will provide a more comprehensive view of the problems encountered by infertility patients. Two researchers will use a manual search to sort out and filter the HRQoL information reported by the infertility patient on *Zhihu*, and thematic analysis will be used.

Data analysis plan All interviewers will be transcribed verbatim and subject to thematic analysis based on grounded theory [33–35]. Two independent interviewers will perform the coding analysis, and discrepancies will be resolved by consensus meetings with the third interviewer to achieve consensus. The coding is mainly divided into three steps [35]. (1) Open coding: two researchers extract information from the transcript, which will be labeled and encoded respectively; (2) Axial coding: classify initial concepts according to relevance and frequency, and aggregate concepts reflecting an aspect together to develop items and domains; (3) Selective coding: the researchers constructed a HRQoL framework for infertility patients by further analyzing the logical relationships among the main items and domains, which will refer to the Wilson-Cleary revised model [26]. Finally, qualitative analysis data will be presented in the word cloud and hierarchy chart. These analyses will be performed in Chinese and then translated into English for reporting. The qualitative analysis will be conducted in MAXQDA 2020 software.

Producing draft candidate items

The items identified from the systematic review and the qualitative interview analysis will be used to construct draft candidate items, which will reflect the full range of patients' experiences and views. The statements (rows) of the interviewees and item concepts (columns) of a systematic review or qualitative interview form a frame matrix by Excel. The patients' qualitative data will be analysed to produce themes, which will be used to construct draft candidate items. The matrix will help finalise the themes [36]. We will conduct a preliminary test in a small group, with 15–20 patients who do not previously participate, and ask them to comment on the draft candidate items. The draft candidate items will be revised based on their comments to ensure that patients can understand the items and domains.

Stage 2: Item selection
Quantitative surveying

This study adopts methods of convenience sampling and purposeful sampling in quantitative surveying. The questionnaire survey will be implemented among infertile couples from seven reproductive medicine centers, and a face-to-face investigation will be conducted. According to the Graded Response Model (GRM) in IRT, the minimum sample size is calculated to be 500 [37]. We will select 100 infertility patients from each reproductive medical center, and the total number of patients being investigated is expected to be 700. The quantitative survey will consist of socio-demographic information (e.g., age, gender, occupation, annual household income, etc.), clinical characteristics (e.g., infertility duration, type of infertility, treatment history, etc.), the draft candidate items and the FertiQoL instrument.

Data analysis plan

Cronbach's α coefficient, coefficient of variation (CV), Spearman's rank correlation coefficients, EFA will be used for item selection [28]. The Cronbach's α coefficient of 0.7 or above will be considered as appropriate [28, 38]. The items of CV greater than 0.25 are considered appropriate. The strength of the correlation (r) was interpreted as follows: $r > 0.7$ indicates strong; $0.3 < r < 0.7$ indicates

Shi *et al. Health and Quality of Life Outcomes* (2022) 20:54

Page 5 of 8

moderate; $r < 0.3$ indicates weak [39]. EFA will be conducted to select items and examine domains. Factors will be retained if their eigenvalues are greater than 1.0, following the Kaiser (1960) rule [40]. The minimum factor loading of 0.4 will be considered necessary to warrant retaining a statement [41]. In the IRT, Samejima Graded Response Model, which is suitable for multiple response levels [42, 43], will be used to calculate the size of the information function of the item for selection.

Cognitive interview

Qualitative feedback regarding the final candidate items will be sought from an expert committee, which is composed of reproductive medicine clinicians, health statisticians, HRQoL researchers and infertility psychological counselling service providers. In addition, cognitive interviews will be conducted among 10–12 patients, who are not previously involved to test the rationality and comprehensiveness of the final candidate items. The patients will be recruited from outpatient clinics and inpatient wards in the Hospital for Reproductive Medicine Affiliated of Shandong University. Furthermore, convenience sampling and purposeful sampling methods will be used to select participants. The previously two experienced qualitative interviewers will conduct the cognitive interview.

Stage 3: Validation
Design and sample
The reliability, validity and sensitivity/responsiveness of the final version of the instrument will be evaluated. We will invite 10–15 experts from reproductive medicine clinical, health statisticians, HRQoL and infertility psychological counseling to evaluate the content validity of the instrument. Convenience sampling and purposeful sampling methods will be used to select participants from outpatient clinics and inpatient wards, who receive different treatments, including drugs treatment, surgical treatment, artificial insemination (AI), gamete transplantation and in vitro fertilization-embryo transfer

(IVF-ET). An expected number of 1,400 patients will be investigated in the seven reproductive medicine centers at baseline survey. The electronic or paper questionnaires will be used for face-to-face survey by trained research staff in outpatient clinics and inpatient wards. The questionnaire survey will consist of socio-demographic information (e.g., age, gender, occupation, annual household income), clinical characteristics (e.g., infertility duration, type of infertility, treatment history), the final version of the new instrument and FertiQoL instrument. To evaluate test–retest reliability, a random sample of 200 infertility patients will be selected to complete the instrument again in electronic format online 2 weeks (approximately 10–15 days) after the first test. Most infertility treatment cycles last approximately 2 months through consultation with a reproductive medicine clinician, and HRQoL is expected to change. In order to measure the responsiveness, a follow-up survey will be conducted on all patients who completed the baseline survey to finish the final instrument again after 2 months, and the online surveys will be used.

Item and psychometric analysis
Reliability   Cronbach's α coefficient will be applied to estimate the internal consistency reliability, and the intraclass correlation coefficient (ICC) will be used to estimate the test–retest reliability (Table 1). The Cronbach's α coefficient and ICC of 0.7 or above will be considered appropriate [28, 38].

Validity   The content validity will be evaluated with item-level content validity index (I-CVI) and scale-level content validity index (S-CVI) based on experts' ratings of item relevance [44, 45]. EFA and confirmatory factor analysis (CFA) will be used to show the construct validity. EFA will be used to explore the latent factor structure and account for potential correlations among factors. Moreover, CFA will be conducted to test the dimensionality of the de novo instrument and validate the hypothetical structures. FertiQoL is currently the

Table 1 Psychometric properties of the HRQoL instrument for Chinese infertile couples

Measurement property	Type	Measure
Reliability	Internal consistency	Cronbach's α
	Test-rest	Intraclass correlation coefficient (ICC) will be measured at baseline and at 2-week follow-up
	Information function	Information function of each item and domain calculated according to item response theory (IRT)
Validity	Content validity	Item-level content validity index (I-CVI)/Scale-level content validity index (S-CVI)
	Construct validity	Explore factor analysis (EFA); Confirmatory factor analysis (CFA)
	Convergent/discriminant validity	Spearman correlation coefficients
Sensitivity/responsiveness	–	Floor and ceiling effects; effect size (ES); standardized response mean (SRM); paired *t* test

Shi *et al. Health and Quality of Life Outcomes*   (2022) 20:54

Page 6 of 8

most commonly used disease-specific HRQoL scale for infertile couples [18], which will be used as a reference instrument in construct validity. The association between the Chinese infertile couples HRQoL instrument and the FertiQoL will be studied by using Spearman's rank correlation coefficients. Convergent validity will be supported if the similar attributes items correlate (r) of 0.4 or above, while weak correlations ($r \leq 0.3$) between dissimilar attributes will support discriminant validity [28].

Sensitivity/responsiveness Sensitivity is the ability to detect differences between groups, responsiveness is the ability of a scale to detect changes [28]. Cohen effect size (ES) and standardised response mean (SRM) will be calculated to evaluate sensitivity and responsiveness, and the paired t test will be used to estimate the responsiveness between baseline and 2 months follow up. The same thresholds of 0.2, 0.5 and 0.8 are commonly interpreted for both ES and SRM. An ES or SRM of 0.20 will be considered as small, a value between 0.2 and 0.5 as moderate, and one of 0.8 or greater as large [28]. Floor and ceiling effects can be used as a proxy of sensitivity in relation to how well a measure can detect changes within cross-sectional studies [46]. In this study, floor or ceiling effects will be considered as high if more than 15% of the responders achieve the lowest or highest possible score respectively [47, 48].

Stage 4: Develop the minimal clinically important difference (MCID)

MCID can be useful to specify a benchmark for interpreting mean differences. It has become a standard approach in the interpretation of clinical relevance of changes in HRQoL [49, 50]. All patients will complete the questionnaires and a 7-point global impression item as an external anchor at the baseline and 2 months follow up. The FertiQoL will be used as an external criterion against which changes in de novo infertility HRQoL instrument will be anchored and calibrated. There are a variety of methods discovering concordance to confirm the choice of an MCID [51]. In this study, we will use distribution-based methods and anchor-based methods (e.g., logistic regression, receiver operating characteristic (ROC) curve) to develop the MCID [28, 29].

Discussion

Infertility and its treatment have negative impacts on a couple's marital relationship, sexual life, psychological state and interpersonal relationships, causing personal distress [7–9]. Identifying these factors in infertility

patients, creating a good medical environment and offering diverse health education can help improve their health. The new HRQoL instrument for Chinese infertile couples will be developed, which will provide a standard and effective HRQoL instrument in clinical outcome assessment and health outcome measurement.

This study has several advantages. Firstly, the new HRQoL instrument for use among Chinese infertile couples will be developed by using a mixed method. MCID development will provide a benchmark for interpreting mean differences. Secondly, infertility patients and reproduction medicine clinicians will be surveyed from seven reproductive medicine centers to diversify the sample as much as possible, according to the geographical location and economic development level in China. It will provide a comprehensive view of problems encountered by infertility patients. Thirdly, CTT and IRT will be used for item selection and validation, which will fill in the gap that the existing infertility HRQoL instruments development were only based on CTT.

This study also has certain weaknesses. First of all, this study will investigate infertility patients treated in hospitals, and there may be some selection bias in the samples. Future studies may evaluate the psychometric properties in a larger population. In addition, some post information on *Zhihu* may have low reliability.

Shi *et al. Health and Quality of Life Outcomes*   (2022) 20:54

Page 7 of 8

Declarations

Ethics approval and consent to participate
This study has been approved by the Ethics Committee of Centre for Health Management and Policy Research, Shandong University (ECCHMPSDU20210203), and conforms to the ethics guidelines of the Declaration of Helsinki. All participants will be informed about the aim of the study and written informed consent will take from each participant.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Centre for Health Management and Policy Research, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan 250012, China. [2]NHC Key Lab of Health Economics and Policy Research, Shandong University, Jinan 250012, China. [3]Center for Health Preference Research, Shandong University, Jinan 250012, China. [4]School of Insurance, Southwestern University of Finance and Economics, Chengdu 611130, China. [5]Center for Reproductive Medicine, Cheeloo College of Medicine, Shandong University, Jinan 250001, China. [6]Centre for Health Economics, Monash Business School, Monash University, Melbourne 3145, Australia.

References
1. Zegers-Hochschild F, Adamson GD, Dyer S, et al. The international glossary on infertility and fertility care, 2017. Hum Reprod. 2017;32(9):1796–801.
2. Gurunath S, Pandian Z, Anderson RA, et al. Defining infertility—a systematic review of prevalence studies. Hum Reprod Update. 2011;17(5):575–88.
3. Zegers-Hochschild F, Adamson GD, de Mouzon J, et al. The International Committee for Monitoring Assisted Reproductive Technology (ICMART) and the World Health Organization (WHO) revised glossary on ART terminology, 2009. Hum Reprod. 2009;24(11):2683–7.
4. Qiao J, Wang Y, Li X, et al. A Lancet Commission on 70 years of women's reproductive, maternal, newborn, child, and adolescent health in China. Lancet. 2021;397(10293):2497–536.
5. Zhou Z, Zheng D, Wu H, et al. Epidemiology of infertility in China: a population-based study. BJOG. 2017;125(4):432–41.
6. The Beijing News. China's infertility rate reached 12–15 percent (in chinese). 2018. http://epaper.bjnews.com.cn/html/2018-10/30/content_736104.htm?div=-1. Accessed 20 Nov 2021.
7. Schmidt L. Psychosocial burden of infertility and assisted reproduction. Lancet. 2006;367(9508):379–80.
8. Luk BH-K, Loke AY. The impact of infertility on the psychological well-being, marital relationships, sexual relationships, and quality of life of couples: a systematic review. J Sex Marital Ther. 2015;41(6):610–25.
9. Mousavi SA, Masoumi SZ, Keramat A, et al. Assessment of questionnaires measuring quality of life in infertile couples: a systematic review. J Reprod Infertil. 2013;14(3):110–9.
10. CPC Central Committee and State Council. Decision of the CPC Central Committee and The State Council on optimizing birth policy to promote long-term balanced development of population (in chinese). 2021. http://www.gov.cn/zhengce/2021-07/20/content_5626190.htm. Accessed 20 Nov 2021.
11. FDA. Guidance for industry patient-reported outcome measures: use in medical product development to support labeling claims. 2009. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/patient-reported-outcome-measures-use-medical-product-development-support-labeling-claims. Accessed 20 Nov 2021.
12. Levesque A, Li HZ, Bohémier M. Cultural variations in health conceptions: a qualitative approach. Pimatisiwin. 2013;11(2):215–29.
13. Mao Z, Ahmed S, Graham C, et al. Similarities and differences in health-related quality-of-life concepts between the east and the west: a qualitative analysis of the content of health-related quality-of-life measures. Value Health Reg Issues. 2021;24(C):96–106.
14. Cnossen MC, Polinder S, Vos PE, et al. Comparing health-related quality of life of Dutch and Chinese patients with traumatic brain injury: do cultural differences play a role? Health Qual Life Outcomes. 2017;15(72):1–10.
15. Ni Y, Tong C, Huang L, et al. The analysis of fertility quality of life and the influencing factors of patients with repeated implantation failure. Health Qual Life Outcomes. 2021;19(32):1–10.
16. Fu B, Qin N, Cheng L, et al. Development and validation of an infertility stigma scale for Chinese women. J Psychosom Res. 2015;79(1):69–75.
17. Jing XY, Gu W, Xu XL, et al. Stigma predicting fertility quality of life among Chinese infertile women undergoing in vitrofertilization-embryo transfer. J Psychosom Obst Gyn.
18. Kitchen H, Aldhouse N, Trigg A, et al. A review of patient-reported outcome measures to assess female infertility-related quality of life. Health Qual Life Outcomes. 2017;15(86):1–12.
19. Chachamovich JR, Chachamovich E, Ezer H, et al. Investigating quality of life and health-related quality of life in infertility: a systematic review. J Psychosom Obstet Gynecol. 2010;31(2):101–10.
20. Boivin J, Takefman J, Braverman A. The fertility quality of life (FertiQoL) tool: development and general psychometric properties. Hum Reprod. 2011;26(8):2084–91.
21. Yang X. A validity and reliability study of the fertility quality of life (FertiQoL) tool in Chinese People (in chinese). Guangzhou: Southern Medical University; 2016.
22. Hubens K, Arons AMM, Krol M. Measurement and evaluation of quality of life and well-being in individuals having or having had fertility problems: a systematic review. Eur J Contracept Reprod. 2018;23(6):441–50.
23. Shi. Z, Nie. H, Geng. L, et al. Evaluating health-related quality of life and subject wellbeing among infertility patients: a cross-sectional study in eastern China. Qual Life Res. 2021;Under review
24. Wang Y. Development of quality of life scale for female liver depression infertility in traditional Chinese medicine (in chinese). Beijing: Beijing University of Chinese Medicine; 2017.
25. Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life: a conceptual model of patient outcomes. JAMA. 1995;273(1):59–65.
26. Ferrans CE, Zerwic JJ, Wilbur JE, et al. Conceptual model of health-related quality of life. J Nurs Scholarsh. 2005;37(4):336–42.
27. Patrick DL, Burke LB, Gwaltney CJ, et al. Content validity-establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: part 1-eliciting concepts for a new PRO instrument. Value Health. 2011;14(8):967–77.
28. Fayers P, Machin D. Quality of life: the assessment, analysis and interpretation of patient-reported outcomes. Hoboken: Wiley; 2016.
29. Angst F, Aeschlimann A, Angst J. The minimal clinically important difference (MCID) raised the significance of outcome effects above the statistical level, with methodological implications for future studies. J Clin Epidemiol. 2017;82:128–36.
30. Chen Z, Liu J, Huang H, et al. Guidelines for diagnosing infertility (in chinese). Chin J Obstet Gynecol. 2019;54(7):505–11.
31. National Bureau of Statistics of China. China Sixth National Census (in chinese). 2010. http://www.stats.gov.cn/tjsj/pcsj/rkpc/6rp/indexch.htm. Accessed 20 Nov 2021.
32. What does infertility feel like? (in chinese). 2021. https://www.zhihu.com/question/57975959/answers/updated. Accessed 20 Nov 2021.
33. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol. 2008;3(2):77–101.
34. Gale NK, Heath G, Cameron E, et al. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol. 2013;13(117):1–8.
35. Anselm Strauss JC. Basics of qualitative research: grounded theory procedures and techniques. Newhury Park. California, Sare Publications; 1990.
36. Muratov S, Podbielski DW, Jack SM, et al. Preference-based disease-specific health-related quality of life instrument for glaucoma: a mixed methods study protocol. BMJ Open. 2016;6(11):e012732.
37. Embretson SE, Reise SP. Item response theory for psychologists. Mahwah: Lawrence Erlbaum; 2000.

Shi *et al. Health and Quality of Life Outcomes*   (2022) 20:54

Page 8 of 8

38. Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika. 1951;16(3):297–334.
39. Ratner B. The correlation coefficient: its values range between +1/1, or do they? J Target Meas Anal Mark. 2009;17(2):139–42.
40. Kaiser HF. The application of electronic computers to factor analysis. Educ Psychol Meas. 1960;20(1):141–51.
41. Dilorio CK. Measurement in health behavior: methods for research and evaluation. London: Wiley; 2005.
42. Reeve BB, Hays RD, Bjorner JB, et al. Psychometric evaluation and calibration of health-related quality of life item banks—plans for the patient-reported outcomes measurement information system (PROMIS). Med Care. 2007;45(5):S22–31.
43. Uttaro T, Lehman A. Graded response modeling of the quality of life interview. Eval Program Plan. 1999;22(1):41–52.
44. Haynes SN, Richard DCS, Kubany ES. Content validity in psychological assessment: a functional approach to concepts and methods. Psychol Assess. 1995;7(3):238–47.
45. Polit DF, Beck CT, Owen SV. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. Res Nurs Health. 2007;30(4):459–67.
46. Franklin M, Mukuria C, Mulhern B, et al. Measuring the burden of schizophrenia using clinician and patient-reported measures: an exploratory analysis of construct validity. Patient. 2019;12(4):405–17.
47. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60(1):34–42.
48. Reuvers SHM, Korfage IJ, Scheepe JR, et al. The urinary-specific quality of life of multiple sclerosis patients: Dutch translation and validation of the SF-Qualiveen. Neurourol Urodyn. 2017;36(6):1629–35.
49. McGlothlin AE, Lewis RJ. Minimal clinically important difference defining what really matters to patients. JAMA. 2014;312(13):1342–3.
50. Guyatt GH, Osoba D, Wu AW, et al. Methods to explain the clinical significance of health status measures. Mayo Clin Proc. 2002;77(4):371–83.
51. Jayadevappa R, Cook R, Chhatre S. Minimal important difference to infer changes in health-related quality of life-a systematic review. J Clin Epidemiol. 2017;89:188–98.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.