

This is a repository copy of *Do Rural Residents in China Understand EQ-5D-5L as Intended? Evidence From a Qualitative Study*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/159550/>

Version: Published Version

Article:

Yang, Fan orcid.org/0000-0003-4689-265X, Jiang, Shan, He, Xiao-ning et al. (4 more authors) (2020) Do Rural Residents in China Understand EQ-5D-5L as Intended? Evidence From a Qualitative Study. *PharmacoEconomics - Open*. ISSN 2509-4254

<https://doi.org/10.1007/s41669-020-00212-z>

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial (CC BY-NC) licence. This licence allows you to remix, tweak, and build upon this work non-commercially, and any new works must also acknowledge the authors and be non-commercial. You don't have to license any derivative works on the same terms. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



Do Rural Residents in China Understand EQ-5D-5L as Intended? Evidence From a Qualitative Study

Fan Yang¹ · Shan Jiang² · Xiao-ning He³ · Hong-chao Li⁴ · Hong-yan Wu⁵ · Tian-tian Zhang⁶ · Jing Wu³ 

© The Author(s) 2020

Abstract

Background The 5-level EQ-5D (EQ-5D-5L) has been increasingly used in China to measure the health status of the general population and patients. However, its content validity among rural residents in China has not been formally evaluated. This qualitative study aims to assess the content validity of EQ-5D-5L among rural Chinese.

Methods Participants were recruited from four regions (North, South, East and West) across China. Eligible participants were those living in the rural area in last three years and making a living by agricultural operations. Semi-structured interviews were conducted. Interview transcripts were analysed to assess the comprehensibility, relevance, clarity and comprehensiveness.

Results Sixty-two participants were included, comparable to the national figures regarding age, sex and education. For comprehensibility, participants could understand the ‘mobility’, ‘self-care’ and ‘usual activities’ domains well, but some reported confusions in ‘pain/discomfort’ ($n = 42$) and ‘anxiety/depression’ ($n = 35$). Some also reported difficulties in understanding anxiety ($n = 6$) and depression ($n = 9$), possibly due to the formal wording used. For relevance, all domains were reported as health-related and participants’ responses were based on their own health. For clarity, all could distinguish the five levels, but suggestions on reducing response levels and alternative wording for ‘slight’ were raised. For comprehensiveness, two aspects (fatigue/energy and appetite) were raised beyond the EQ-5D-5L domains. The ‘mobility’ domain was selected as the most important and ‘anxiety/depression’ as the least important.

Conclusion Rural Chinese reported problems on the content validity of Chinese EQ-5D-5L. It might be sensible to consider some revisions to make it more understandable for rural residents.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s41669-020-00212-z>) contains supplementary material, which is available to authorized users.

✉ Jing Wu
jingwu@tju.edu.cn

- ¹ Centre for Health Economics, University of York, York, UK
- ² School of Population and Public Health, University of British Columbia, Vancouver, BC, Canada
- ³ School of Pharmaceutical Science and Technology, Tianjin University, Tianjin, China
- ⁴ School of International Pharmaceutical Business, China Pharmaceutical University, Nanjing, China
- ⁵ School of Medicine and Health Management, Guizhou Medical University, Guiyang, China
- ⁶ College of Pharmacy, Jinan University, Guangzhou, China

Key Points for Decision Makers

Generally, the content validity of the Chinese version of EQ-5D-5L is satisfactory in rural Chinese.

The ‘pain/discomfort’ and ‘anxiety/depression’ domains may be subject to poor comprehensibility and there is a slight lack of clarity regarding the response levels.

It might be sensible to discuss how to improve the current Chinese version to make it more understandable for rural residents.

1 Introduction

The EQ-5D has been the most commonly used instrument for valuing health since its development in the 1980s [1]. The health utilities generated from EQ-5D provide a way to estimate quality-adjusted life-years (QALYs) for use

in cost-effectiveness analysis (CEA) to inform resource allocation decisions in health care. EQ-5D has also been used in clinical trials, observational studies, population surveys and routine data collection in health care systems as a patient-reported outcome measure [1]. It includes five domains—mobility, self-care, usual activities, pain/discomfort and anxiety/depression, with multiple response levels for each domain that allow respondents to self-rate their health by choosing the appropriate level [2]. Currently, there are two versions available, the earlier EQ-5D-3L with three levels for each domain (i.e., no, some and extreme problems) and the newer EQ-5D-5L with five levels for each domain (i.e., no, slight, moderate, severe and extreme problems). The EQ-5D-5L was developed to address the presence of ceiling effects and poor sensitivity associated with the EQ-5D-3L [3–6] and it has been increasingly used in literature [7, 8].

In China, the EQ-5D-3L has been widely used to measure the health status of the general population [9–11] and patients [12–14]. The Chinese version of EQ-5D-5L became available in 2012 [15] and it has been used in several studies [12, 16, 17]. In light of the improved measurement properties [6], EQ-5D-5L is expected to be used more often in future studies. Both EQ-5D versions have been recommended by China Guidelines for Pharmacoeconomic Evaluations as a tool for conducting health technology assessment [18].

However, the Chinese version of EQ-5D-5L was developed through a response scaling exercise, among 50 participants recruited from downtown Beijing [15]. As commented by the authors, the small sample only included urban residents [15], so its validity in rural areas may be limited. Rural areas account for 90% of mainland China and rural residents comprise 40% of the Chinese population [19]. In China, rural residents and urban residents differ significantly in many aspects. For example, according to the 2010 population census, only 10% of rural residents received the senior secondary level or higher education while the figure for urban residents was 46% [20]. There was a higher proportion of people over 65 years of age among rural residents compared with those living in urban cities (rural: 10.1% and urban: 7.7%). Rural residents also have remarkably lower incomes [21] and inferior medical health services [22, 23] than their counterparts living in urban areas. In view of these huge differences, particular attention should be paid to the rural residents. To the best of our knowledge, the content validity of EQ-5D-5L or the earlier EQ-5D-3L has not previously been evaluated in rural residents in China. Therefore, we conducted this qualitative study to assess its content validity among rural Chinese.

2 Methods

2.1 Study Participants

Four regions (North, South, East and West) across China were selected on the basis of geographic location involving the cities of Tianjin (municipality), Guangzhou (Guangdong province), Nanjing (Jiangsu province) and Guiyang (Guizhou province) and then one county (rural area) from each region was selected for data collection. Participants from each county were recruited through convenience sampling. Rural residents who had lived in the county for the last 3 years and made a living by agricultural operations were eligible.

To obtain a representative sample, we used the national figures on the population distribution [20] and sought advice from an expert panel. Healthcare researchers from four universities in the four cities (Tianjin, Guangzhou, Nanjing and Guiyang) with sufficient survey experience in rural China were invited by the lead authors via personal contact and a total of 15 members were included in the expert panel. Teleconferences were held to achieve consensus between members. It was agreed that age, sex and education level should be taken into account, but not socioeconomic status considering that the survey was self-reported and the responses to socioeconomic status may not be reliable. Additionally, we did not include disease status as there is no clear evidence showing that the disease status could affect respondents' understanding of the EQ-5D-5L and chronic disease is usually under-diagnosed among rural Chinese, so respondents may not be able to report accurate information. Disease status was not considered in quota sampling in the previous EQ-5D-5L valuation studies such as in the US [24] and in China [25].

In line with the recommendation for qualitative research that sample sizes of between 5 and 15 are typical [26], the target sample size for each county was 15. A total of eight interviewers were involved in the data collection, with two in each county. They were the junior members of the expert panel who were postgraduate students with experience in conducting surveys. Before the data collection, the interviewers received training from the team leaders and conducted pilot interviews. Training included briefing on the questionnaire, addressing queries and interviewers conducting pilot interviews with others. The feedback on the interviewing process was collected and then used to improve the interviews in the main data collection. Data were collected from September to October 2018 and involved semi-structured, one-on-one, face-to-face interviews, administered by trained interviewers using a paper-and-pencil questionnaire. The interviews were conducted at the participant's home or other locations that allowed quietness and privacy.

Participants provided written informed consent prior to the interview. During data collection, all participants were allocated with one study ID and their responses to the questions were recorded with this study ID only. No personal information was collected to protect privacy and we recorded the participants' age as age group (18–35 years/36–55 years/56 years and above). Ethical approval was obtained by the Safety and Ethics Committee of the School of Pharmaceutical Science and Technology (SPST2018-03) in Tianjin University.

2.2 Survey Instrument

The survey instrument consisted of two tasks. The first task presented a cognitive debriefing procedure [27] to elicit participants' feedback on the content of the EQ-5D-5L. In cognitive debriefing, each domain was assessed in terms of comprehensibility, relevance, clarity and comprehensiveness.

Comprehensibility was evaluated by examining whether the questions could be understood by participants as intended. A series of questions were asked about their interpretations of the questions, e.g., “*could you give an example of usual activities*”. For the two composite domains (pain/discomfort, anxiety/depression), one previous study reported the significant effect of splitting the anxiety/depression domain on self-reported health and health states valuation [28], and therefore, in this study, participants were asked to explain their interpretations of both terms and we also asked them whether there were differences between the two terms.

Relevance was assessed by examining whether the domains of EQ-5D were health-related and whether participants' response could reflect their health condition using the questions “*is the domain relevant to your health*” and “*do you choose the response level based on your health*”.

Clarity was assessed for the response levels. Participants were asked whether they could distinguish the five response levels, and then if they could give examples of each response level; for example, “*could you give an example of moderate problems with mobility*”. We also explored the potential benefits or challenges of the increasing number of levels in the EQ-5D-5L relative to EQ-5D-3L. Participants were asked to recall all the responses levels after reading the question. We hypothesised that being unable to recall all five levels might indicate that the extra levels have posed additional cognitive challenges on participants to distinguish between all the options and this may further bias their response.

Comprehensiveness was assessed by examining whether there were important aspects of health not covered by the five domains of EQ-5D-5L using an open-ended question. The five domains in EQ-5D were developed based on European populations, so it may not be able to fully capture what

matters for the Chinese population, especially for rural residents.

In the first task, participants were asked to choose a response level to each domain of the EQ-5D-5L according to their own health condition before and after debriefing, to check whether the debriefing procedure helped them in understanding and answering the questions. The second survey task required participants to select the most and least important domains related to their health among the five EQ-5D-5L domains. This task could also reflect their understanding of the concepts. Details of the instrument are available in the Appendix (see electronic supplementary material).

2.3 Analysis

All interviews were audio-recorded and transcribed, and then the transcripts were analysed using QSR NVivo 12 software following the content analysis framework [29].

3 Results

3.1 Sample Description

A total of sixty-two participants were recruited, with 27.4% aged 18–35 years, 51.6% aged 36–55 years and 21.0% aged 56 years and above. Among all participants, 51.6% were female; 25.8% participants had primary education or lower, 50.0% had junior secondary education and 24.2% had senior secondary education or higher (Table 1). Participants from different regions were slightly different in terms of age, gender and education level (Table 1).

3.2 Debriefing

Before the debriefing task, the mean (standard deviation) EQ-5D-5L score was 0.939 (0.088) using the Chinese EQ-5D-5L value set [25] (Table 1 and Fig. 1a). Participants from different regions reported slightly different pre-debriefing EQ-5D-5L scores with people from the North reporting the highest while those from the West reported the lowest (North: 0.981; South: 0.948; East: 0.925 and West: 0.897) (Table 1 and Fig. 1c).

The debriefing task took 38.8 min to complete on average and participants from the West spent the longest time (North: 32.6; South: 32.4; East: 41.0 and West: 49.9) (Table 1). After debriefing, five (8%) participants (4 from the West and 1 from the East) changed their responses, and the changes were in both directions of the response level. Two changed their response for ‘anxiety/depression’ from ‘no’ to ‘slight’ (Fig. 2a), one changed the response for ‘self-care’ from ‘no’

Table 1 Participants' characteristics

Characteristics	National figure*	Total (n = 62)	North (n = 17)	South (n = 15)	East (n = 15)	West (n = 15)
Age group						
18–35 years	37.4%	17 (27.4%)	5 (29.4%)	5 (33.3%)	4 (26.7%)	3 (20.0%)
36–55 years	41.6%	32 (51.6%)	9 (52.9%)	7 (46.7%)	7 (46.7%)	9 (60.0%)
56 years and above	21.1%	13 (21.0%)	3 (17.6%)	3 (20.0%)	4 (26.7%)	3 (20.0%)
Female, %	48.7%	32 (51.6%)	9 (52.9%)	8 (53.3%)	8 (53.3%)	7 (46.7%)
Education level						
Primary level or lower	26.7%	16 (25.8%)	4 (23.5%)	5 (33.3%)	4 (26.7%)	3 (20.0%)
Junior secondary level	44.3%	31 (50.0%)	9 (52.9%)	6 (40.0%)	7 (46.7%)	9 (60.0%)
Senior secondary level or higher	29.0%	15 (24.2%)	4 (23.5%)	4 (26.7%)	4 (26.7%)	3 (20.0%)
Interviewing time (minutes), mean (SD)		38.8 (14.7)	32.6 (10.2)	32.4 (11.1)	41.0 (16.5)	49.9 (14.3)
Pre-debriefing EQ-5D-5L score, mean (SD)		0.939 (0.088)	0.981 (0.033)	0.948 (0.053)	0.925 (0.116)	0.897 (0.109)
Post-debriefing EQ-5D-5L score, mean (SD)		0.939 (0.083)	0.981 (0.033)	0.948 (0.053)	0.918 (0.122)	0.907 (0.088)

*Source: The 2010 population census of the People's Republic of China; Available from: <https://www.stats.gov.cn/tjsj/pcsj/rkpc/6rp/indexch.htm>
SD standard deviation

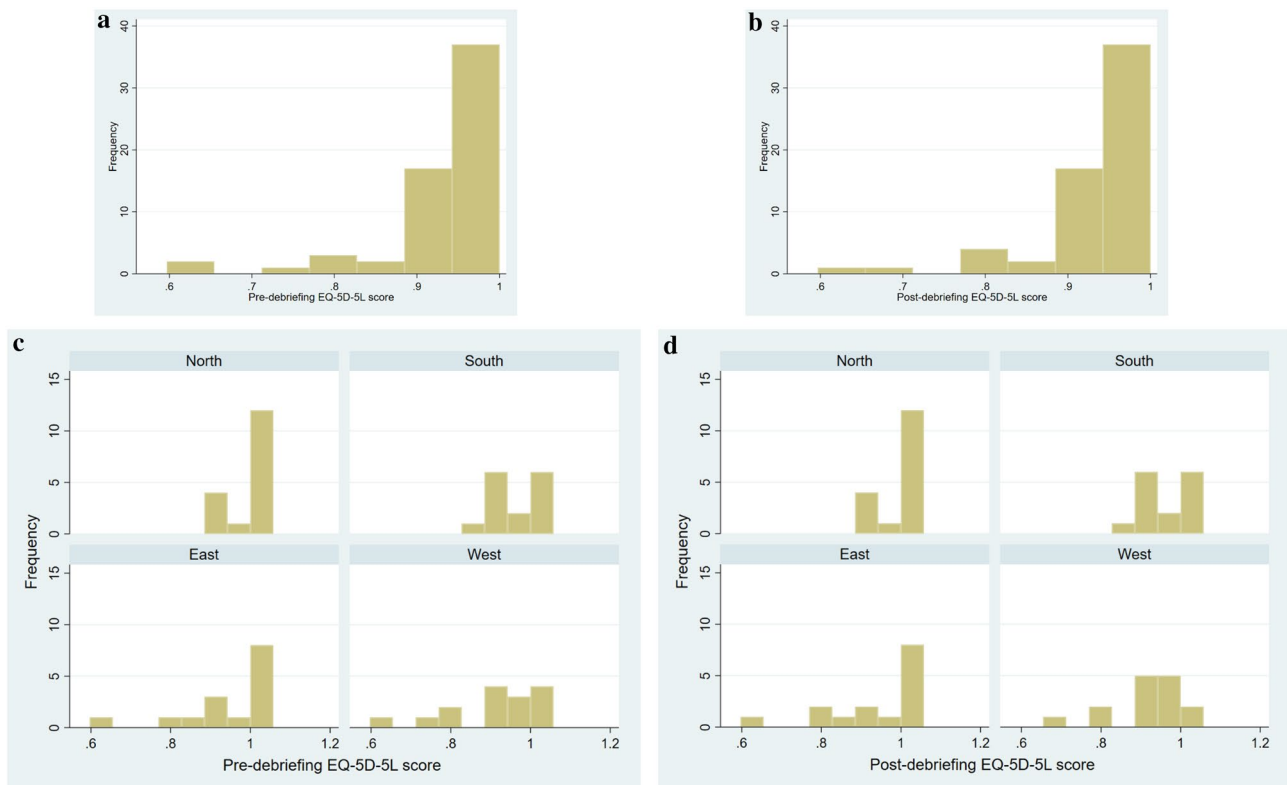


Fig. 1 EQ-5D-5L scores pre- and post-debriefing. **a** EQ-5D-5L overall score (pre-debriefing). **b** EQ-5D-5L overall score (post-debriefing). **c** EQ-5D-5L score by region (pre-debriefing). **d** EQ-5D-5L score by region (post-debriefing)

to 'slight' and the response for 'pain/discomfort' from 'moderate' to 'slight' (Fig. 2b), one changed the responses for 'anxiety/depression' from 'severe' to 'no' (Fig. 2c) and one changed the response for 'mobility' from 'no' to 'slight' and the response for 'anxiety/depression' from 'no' to 'slight'

(Fig. 2d). After debriefing, the overall EQ-5D-5L score was identical but the score of participants from the East dropped from 0.925 to 0.918 while that of patients from the West increased from 0.897 to 0.907 (Table 1).

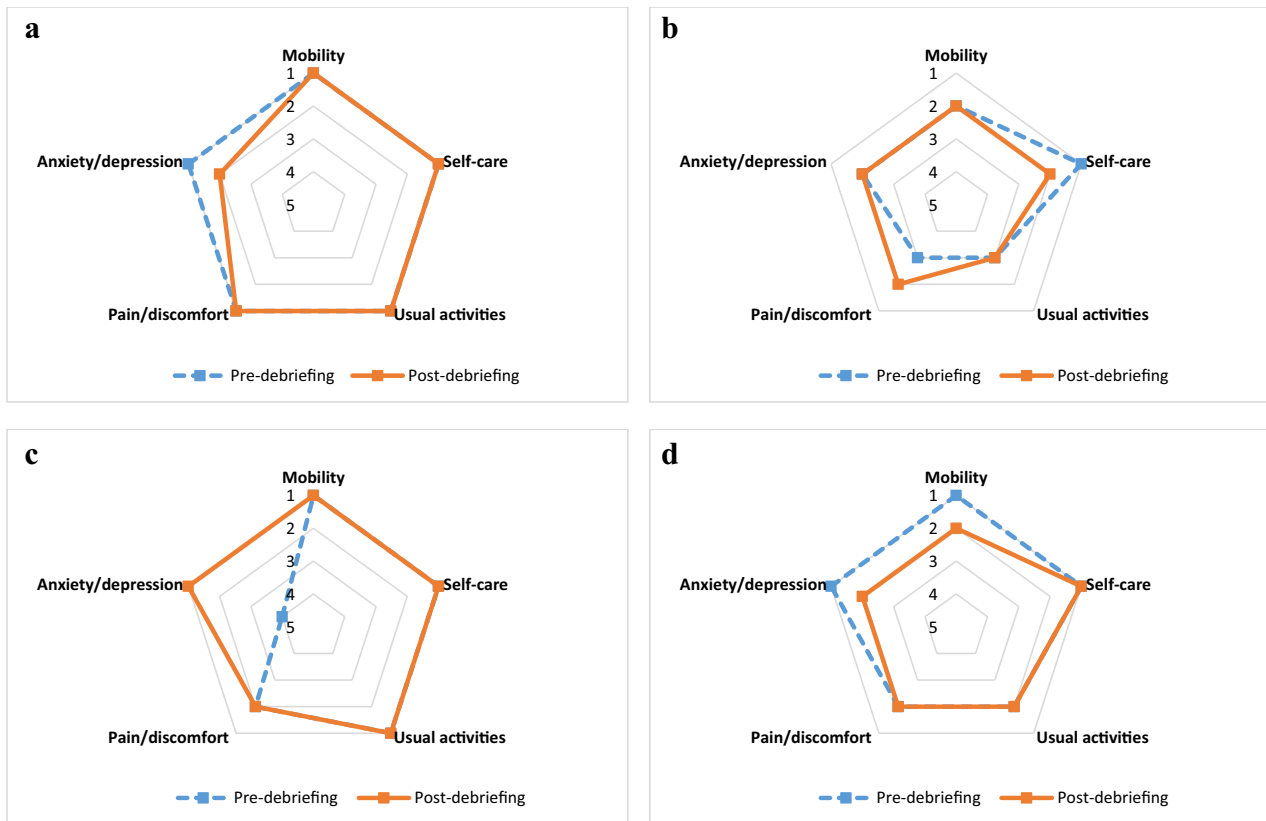


Fig. 2 Changes of responses to the EQ-5D-5L pre- and post-debriefing

3.3 Comprehensibility

For the ‘mobility’, ‘self-care’ and ‘usual activities’ domains, all participants reported that they could understand the domains well by giving their interpretations.

- “Mobility is the physical activity, such as do house work and go to the farmland.” (18–35 years, male).
- “Self-care means that as an adult, I could handle my own life well, not like a child.” (18–35 years, female).
- “Usual activities include going to community centres to meet friends, grocery shopping in markets, etc.” (36–55 years, female).

For the ‘pain/discomfort’ domain, all participants could give their interpretations, but when asked the differences between the two terms, forty-two (67.7%) participants indicated that although the two terms shared similar meaning, there were differences between ‘pain’ (Chinese: 疼痛) and ‘discomfort’ (Chinese: 不舒服). These differences were mainly related to the severity level and the range that each term covers.

- “They are different. If you have pain, you also have discomfort, but discomfort does not necessarily relate to pain, for example, it can be sore waist or bloating.” (36–55 years, female).
- “Discomfort can also be emotional, for example, unhappy, but without pain.” (18–35 years, male).
- “Pain relates to physical feelings, but discomfort covers a wider range, including emotions.” (36–55 years, male).

For the ‘anxiety/depression’ domain, six (10.0%) participants could not understand ‘anxiety’ (Chinese: 焦虑) and nine (14.5%) participants mentioned they did not understand ‘depression’ (Chinese: 沮丧). This was mainly due to the wording used being too formal for them to understand. Two (3.2%) participants suggested some revisions to make the question more straightforward, such as using the terms ‘unhappiness’ or ‘sadness’ (Chinese: 不开心/难过).

- “Just ask it simply, like whether you are happy or not.” (36–55 years, female).
- “I like the question asking whether I am in a good mood rather than anxious or depressed.” (18–35 years, male).

Similar to the ‘pain/discomfort’ domain, thirty-five (56.5%) participants reported differences between ‘anxiety’ and ‘depression’. These differences may confuse them when trying to select the appropriate response level.

“They (anxiety and depression) are different. Anxiety means there is something that puzzles you and you could not stop thinking about it. Depression is in a bad mood, similar to sadness.” (18–35 years, male).

“Anxiety is worry, mainly emotionally, but depression means lack of vitality.” (18–35 years, female).

3.4 Relevance

All participants reported that the five domains were relevant to their health and they chose the response level based on their own health, supporting the use of EQ-5D-5L to measure health status.

“Yes, the domains are related to my health and I can understand them. I chose my answer based on my health.” (18–35 years, male).

3.5 Clarity

When asked whether they could distinguish the five response levels, all participants gave positive answers. When asked to recall the levels, thirty-eight (61.3%) participants could not recall all levels and the majority of them could only recall four levels.

“I only remember some of the items, not all.” (18–35 years, female).

“Eh, no problems at all, moderate problems, severe problems, like these.” (36–55 years, female).

“No problems, severe problems, moderate problems. Could not recall others.” (56 years and above, male).

Five (8.1%) participants suggested reducing the response levels to three or four.

“Three should represent almost all the conditions, that is, no, some and severe.” (18–35 years, male).

“I like four categories, no, some, moderate and severe.” (36–55 years, female).

Furthermore, three (4.8%) participants suggested that the Chinese wording of ‘some’ (一些) should be used instead of ‘only a little’ (一点) because ‘only a little’ in Chinese could not accurately describe the health state between ‘no problem’ and ‘moderate’.

“I always have only slight problems, such as joint pain, but I report no problems. The ‘only a little’ seems useless to me. If my pain gets worse, I think it would

be ‘some’ problems, but not as severe as ‘moderate’.” (36–55 years, male).

“I don’t understand ‘only a little’ and we don’t use it here (the region), ‘some’ may be better.” (36–55 years, female).

3.6 Comprehensiveness

When asked about aspects of health not covered by the five domains of EQ-5D, fatigue and appetite were raised by some participants ($n=4$, 6.5%).

“Fatigue should be included. If I feel fatigue, I could not do the daily activities as usual.” (56 years and above, male).

“Appetite is important for my health.” (56 years and above, female).

3.7 Importance

In the task of relative importance assessment, all participants provided their answers to the question of the most important domain but two participants did not answer the question of the least important one. The item ‘mobility’ was selected as the most important by 24 (38.7%) participants while ‘anxiety/depression’ was selected as the least important by 37 (59.7%) participants (Table 2).

4 Discussion

Through a qualitative study, we assessed the content validity of the Chinese version of EQ-5D-5L among rural Chinese. Some shortcomings were observed, which may affect its appropriateness among these people.

Good comprehensibility was observed for the ‘mobility’, ‘self-care’ and ‘usual activities’ domains, but the ‘pain/discomfort’ and ‘anxiety/depression’ domains may be subject to poor comprehensibility among rural Chinese. First, the majority of participants reported differences between the two terms in these two composite domains, ‘pain/discomfort’ and ‘anxiety/depression’. As respondents commented,

Table 2 Most and least important domains among EQ-5D-5L

EQ-5D domain	Most important	Least important
Mobility	24 (38.7%)	4 (6.5%)
Self-care	11 (17.7%)	6 (9.7%)
Usual activities	15 (24.2%)	7 (11.3%)
Pain/discomfort	6 (9.7%)	6 (9.7%)
Anxiety/depression	6 (9.7%)	37 (59.7%)
Missing	-	2 (3.2%)

for the ‘pain/discomfort’ domain, the Chinese wording for ‘discomfort’ includes the emotions, but ‘pain’ in Chinese usually refers to physical feelings. The differences may confuse respondents when asked to select the appropriate level. For example, if the respondent is sad or angry, he/she may choose ‘slight’ level when considering discomfort, while the response may be ‘no problems’ when considering ‘pain’. These perceived differences were also reported for the ‘anxiety/depression’ domain and may cause similar problems. Furthermore, one previous study conducted in the UK has shown that splitting the ‘anxiety/depression’ terms would significantly affect the responses [28]. Thus, the comprehensibility of the composite terms might be worth further investigation.

Second, culture can affect people’s perceptions of health [30, 31] and there are cultural differences in understanding the same domain [32]. Taking ‘pain’ as an example, culture plays a significant role in pain perceptions, behaviours and expressions [33]. In Chinese culture, people believe that pain is an essential element of life, a ‘trial’ or a ‘sacrifice’ [34]. Thus, when a person is experiencing pain, he/she tends to endure it until the pain becomes unbearable [34]. Due to these differences, the health concepts that the instrument developed based on European populations intends to capture may not be fully captured when the same instrument is applied among Chinese. Cultural differences should be considered when adapting existing instruments in other cultural and language settings. To our knowledge, there is no study investigating the differences in understanding the same EQ-5D-5L questions between populations from different countries, but previous studies comparing the EQ-5D-5L value sets across countries have reported remarkable differences, indicating that for the same hypothetical health state, people from different cultures may select different response levels [35, 36]. One possible reason for these differences might be that they understand the same EQ-5D-5L domain differently, as illustrated in this study. Nevertheless, this hypothesis needs to be explored further in future studies.

Third, the formal wording used in the ‘anxiety/depression’ domain may also explain the poor comprehensibility. Some respondents reported difficulties in understanding the two terms ‘anxiety’ (Chinese: 焦虑) and ‘depression’ (Chinese: 沮丧). These two Chinese words are seldom used in daily conversations of rural residents, which would make them difficult to understand. Therefore, it might be worth considering using alternative words, which are more commonly used in daily life, such as ‘难过’ (‘sadness’ in Chinese) or ‘不开心’ (‘unhappiness’ in Chinese), as suggested by some participants.

As for the response levels in EQ-5D-5L, all participants reported that they could distinguish between the levels, but most could not recall all. The social desirability bias, which refers to the tendency of respondents to give socially

desirable responses instead of choosing responses that are reflective of their true feelings [37], may explain why they all reported no difficulties in differentiating levels. Being unable to recall all levels might indicate that it would be challenging for them to choose one appropriate level from five and this may further decrease the validity of EQ-5D-5L among them. Some preliminary research also found that for rural residents, EQ-5D-3L seems to perform better than EQ-5D-5L [38], but which version is more suitable for rural residents needs further investigation. Therefore, future studies assessing the health-related quality of life in China using the EQ-5D-5L instrument should keep this in mind. Revision to the level wording has been suggested by some participants. Considering the original response scaling exercise was conducted in Northern China and there are differences in the oral languages used in different regions, revisions for the wording of levels might be worth consideration.

We also identified two health-related aspects beyond the EQ-5D-5L domains, ‘fatigue/energy’ and ‘appetite’. Fatigue is found to be associated with health-related quality of life [39, 40] and the ‘fatigue/energy’ item has been constructed in the most widely used quality-of-life instrument, Short Form 36 (SF-36) [41]. More importantly, most rural residents in China rely on physical labour to make a living, and thus the energy level is important, not only to their health, but to their life. Regarding ‘appetite’, Chinese believe that ‘*the human is the iron, the food is a steel*’, so the appetite is an indicator of the health and such finding is not surprising. These results could inform future research on ‘bolt-on’ dimensions to the EQ-5D. The bolt-on approach aims to address the fact that the current five domains cover only a limited range of health-related quality of life, and it is important to prioritise the domains that are important to the target population.

It is found that rural residents in China put the highest importance on ‘mobility’ and least importance on ‘anxiety/depression’. As discussed earlier, physical labour is their main source of income, so it is in line with our expectation that the physical health-related domain would be considered important. For the least important domain, research showed that the rural residents in China lacked mental health knowledge [42], and as a result, they may not be aware of the importance of mental health. The formal words used in this domain add more barriers to the understanding, so it is not surprising that this was considered as the least important.

Compared with the published qualitative studies of EQ-5D [43, 44], this study, for the first time, evaluated the content validity of EQ-5D-5L among rural Chinese. The results presented here provide some insights on the large-scale application of EQ-5D-5L in China. This study could also inform researchers that urban–rural differences should be noted when measuring health-related quality of life using patient-reported outcome measures. There are

several limitations to this study. First, the respondents were recruited from only four regions across China, which could not fully represent the rural Chinese. There are large differences in culture, oral language and health knowledge between different regions. As the results showed, participants from different regions reported different pre- and post-debriefing EQ-5D-5L scores. These four counties were near the capital of the province or municipality, so for people living in more remote rural areas, there may be more problems in understanding the EQ-5D-5L. Second, data saturation was not adequately considered when designing this study, so it is likely that other shortcomings about the EQ-5D-5L may be missed. Third, the response rate was not recorded in this study, so there might be sampling bias that would affect the quality of the survey. Given these, the results should be seen as indicative. Future studies including participants from more regions to form a more representative sample would be preferred.

5 Conclusions

The content validity of the Chinese version of the EQ-5D-5L may be not satisfactory in rural residents as ‘pain/discomfort’ and ‘anxiety/depression’ domains are subject to poor comprehensibility. There are potential translation inaccuracies in domains and levels. It is suggested that future EQ-5D-5L-based studies implemented in rural China should keep these problems in mind. It might be sensible to discuss how to improve the current Chinese version and make it more understandable for rural residents.

Acknowledgements We would like to thank the members of the expert panel (Ms Yi Xiao, Ms Meng-nan Li, Ms Tang-qin Zhang and Ms Yifan Cai from Guiyang; Ms Zhuo-ru Liang and Mr Xu-dong Chen from Guangzhou; Mr Shi-tong Xie, Ms Meng-qian Zhang and Mr Li-dan Cao from Tianjin and Dr Wen-xi Tang from Nanjing), all interviewers and all participants who made this study possible.

Authors’ contribution FY, SJ and JW designed the study and supervised data collection. FY analysed the data and drafted the manuscript. SJ and JW revised the manuscript. XH, HL, HW and TZ coordinated the data collection at each study site and commented on the manuscript. All authors revised the manuscript and gave final approval of the submitted version.

Compliance with Ethical Standards

Funding This study was funded by National Natural Science Foundation of China (Nos. 71673197, 71804122 & 71463007).

Ethical Approval Ethical approval was obtained by the Safety and Ethics Committee of the School of Pharmaceutical Science and Technology in Tianjin University.

Conflict of interest Fan Yang, Shan Jiang, Xiao-ning He, Hong-chao Li, Hong-yan Wu, Tian-tian Zhang and Jing Wu have no conflicts of interest.

Data Availability Statement The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

References

1. Devlin NJ, Brooks R. EQ-5D and the EuroQol group: past, present and future. *Appl Health Econ Health Policy*. 2017;15(2):127–37.
2. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol group. *Ann Med*. 2001;33(5):337–43.
3. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res*. 2011;20(10):1727–36.
4. Feng Y, Devlin N, Herdman M. Assessing the health of the general population in England: how do the three- and five-level versions of EQ-5D compare? *Health Qual Life Outcomes*. 2015;21(13):171.
5. Janssen MF, Pickard AS, Golicki D, Gudex C, Niewada M, Scalone L, et al. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. *Qual Life Res*. 2013;22(7):1717–27.
6. Janssen MF, Bonsel GJ, Luo N. Is EQ-5D-5L better than EQ-5D-3L? A head-to-head comparison of descriptive systems and value sets from seven countries. *Pharmacoeconomics*. 2018;36(6):675–97.
7. Yang F, Griva K, Lau T, Vathsala A, Lee E, Ng HJ, et al. Health-related quality of life of Asian patients with end-stage renal disease (ESRD) in Singapore. *Qual Life Res*. 2015;24(9):2163–71.
8. Liang Z, Zhang T, Lin T, Liu L, Wang B, Fu AZ, et al. Health-related quality of life among rural men and women with hypertension: assessment by the EQ-5D-5L in Jiangsu, China. *Qual Life Res*. 2019;28:2069–80.
9. Sun S, Chen J, Johannesson M, Kind P, Xu L, Zhang Y, et al. Population health status in China: EQ-5D results, by age, sex and socio-economic status, from the National Health Services Survey 2008. *Qual Life Res*. 2011;20(3):309–20.
10. Huang W, Yu H, Liu C, Liu G, Wu Q, Zhou J, et al. Assessing Health-related quality of life of chinese adults in Heilongjiang using EQ-5D-3L. *Int J Environ Res Public Health*. 2017;14(3):224.
11. Pan CW, Ma Q, Sun HP, Xu Y, Luo N, Wang P. Tea consumption and health-related quality of life in older adults. *J Nutr Health Aging*. 2017;21(5):480–6.

12. Pan CW, Sun HP, Zhou HJ, Ma Q, Xu Y, Luo N, et al. Valuing health-related quality of life in type 2 diabetes patients in China. *Med Decis Mak.* 2016;36(2):234–41.
13. Deng X, Dong P, Zhang L, Tian D, Zhang L, Zhang W, et al. Health-related quality of life in residents aged 18 years and older with and without disease: findings from the First Provincial Health Services Survey of Hunan, China. *BMJ Open.* 2017;7(9):e015880.
14. Zhou T, Guan H, Yao J, Xiong X, Ma A. The quality of life in Chinese population with chronic non-communicable diseases according to EQ-5D-3L: a systematic review. *Qual Life Res.* 2018;27(11):2799–814.
15. Luo N, Li M, Liu GG, Lloyd A, de Charro F, Herdman M. Developing the Chinese version of the new 5-level EQ-5D descriptive system: the response scaling approach. *Qual Life Res.* 2013;22(4):885–90.
16. Liu L, Li S, Wang M, Chen G. Comparison of EQ-5D-5L health state utilities using four country-specific tariffs on a breast cancer patient sample in mainland China. *Patient Prefer Adherence.* 2017;11:1049–56.
17. Liu L, Li S, Zhao Y, Zhang J, Chen G. Health state utilities and subjective well-being among psoriasis vulgaris patients in mainland China. *Qual Life Res.* 2018;27(5):1323–33.
18. China Guidelines for Pharmacoeconomic Evaluations (2011). Available from <https://tools.ispor.org/PEguidelines/countrydet.asp?c=28&t=4>
19. National Bureau of Statistics. Statistical bulletin on national economic and social development in 2018. Available from: https://www.stats.gov.cn/tjsj/zxfb/201902/t20190228_1651265.html
20. The 2010 population census of the People's Republic of China. Available from: <https://www.stats.gov.cn/tjsj/pcsj/rkpc/6rp/indexch.htm>
21. Molero-Simarro R. Inequality in China revisited. The effect of functional distribution of income on urban top incomes, the urban-rural gap and the Gini index, 1978–2015. *China Econ Rev.* 2017;42:101–17.
22. Xu HW, Zhang WJ, Gu LN, Qu ZY, Sa ZH, Zhang XL, et al. Aging village doctors in five counties in rural China: situation and implications. *Hum Resour Health.* 2014;28:12.
23. Zhou XD, Li L, Hesketh T. Health system reform in rural China: voices of healthworkers and service-users. *Soc Sci Med.* 2014;117:134–41.
24. Pickard AS, Law EH, Jiang R, Pullenayegum E, Shaw JW, Xie F, et al. United States valuation of EQ-5D-5L health states using an international protocol. *Value Health.* 2019;22(8):931–41.
25. Luo N, Liu G, Li M, Guan H, Jin X, Rand-Hendriksen K. Estimating an EQ-5D-5L value set for China. *Value Health.* 2017;20(4):662–9.
26. Willis GB. Analyzing and documenting cognitive interview results. *Cognit Interview.* 2005;151:174.
27. Patrick DL, Burke LB, Gwaltney CJ, Leidy NK, Martin ML, Molsen E, 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 2—assessing respondent understanding. *Value Health.* 2011;14(8):978–88.
28. Tsuchiya A, Bansback N, Hole AR, Mulhern B. Manipulating the 5 dimensions of the EuroQol instrument: the effects on self-reporting actual health and valuing hypothetical health states. *Med Decis Making.* 2019;39(4):379–92.
29. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol.* 2006;3(2):77–101.
30. Kagawa-Singer M, Padilla GV, Ashing-Giwa K. Health-related quality of life and culture. *Semin Oncol Nurs.* 2010;26(1):59–67.
31. Collings JA. International differences in psychosocial well-being: a comparative study of adults with epilepsy in three countries. *Seizure.* 1994;3(3):183–90.
32. Molzahn AE, Kalfoss M, Schick Makaroff K, Skevington SM. Comparing the importance of different aspects of quality of life to older adults across diverse cultures. *Age Ageing.* 2011;40(2):192–9.
33. Tung WC, Li ZZ. Pain beliefs and behaviors among Chinese. *Home Health Care Man.* 2015;27(2):95–7.
34. Chen LM, Miaskowski C, Dodd M, Pantilat S. Concepts within the Chinese culture that influence the cancer pain experience. *Cancer Nurs.* 2008;31(2):103–8.
35. Gerlinger C, Bamber L, Leverkus F, Schwenke C, Haberland C, Schmidt G, et al. Comparing the EQ-5D-5L utility index based on value sets of different countries: impact on the interpretation of clinical study results. *BMC Res Notes.* 2019;12(1):18.
36. Wang P, Liu GG, Jo MW, Purba FD, Yang Z, Gandhi M, et al. Valuation of EQ-5D-5L health states: a comparison of seven Asian populations. *Expert Rev Pharmacoecon Outcomes Res.* 2019;19(4):445–51.
37. Middleton KL, Jones JL. Socially desirable response sets: The impact of country culture. *Psychol Market.* 2000;17(2):149–63.
38. Liang ZZT, Lin T, Liu L, Wang B, Fu A, Wang X, Xu X, Luo N, Jiang J. Health-related quality of life among rural Chinese men and women with hypertension: assessment by the EQ-5D-5L. *Guangzhou: EuroQol Asian Academy Meeting; 2019.*
39. Jelsness-Jorgensen LP, Bernklev T, Henriksen M, Torp R, Moum BA. Chronic fatigue is associated with impaired health-related quality of life in inflammatory bowel disease. *Aliment Pharmacol Ther.* 2011;33(1):106–14.
40. Michielsen HJ, Drent M, Peros-Golubicic T, De Vries J. Fatigue is associated with quality of life in sarcoidosis patients. *Chest.* 2006;130(4):989–94.
41. Saris-Baglama R, Dewey C, Chisholm G, Plumb E, Kosinski M, Bjorner J, et al. Quality metric health outcomes scoring software 20: User's Guide. 2007.
42. Yu Y, Liu ZW, Hu M, Liu HM, Yang JP, Zhou L, et al. Mental health help-seeking intentions and preferences of rural Chinese adults. *PLoS ONE.* 2015;10(11):e0141889.
43. Yang Z, Feng Z, Busschbach J, Stolk E, Luo N. How prevalent are implausible EQ-5D-5L health states and how do they affect valuation? A study combining quantitative and qualitative evidence. *Value Health.* 2019;22(7):829–36.
44. Jahagirdar D, Kroll T, Ritchie K, Wyke S. Using patient reported outcome measures in health services: a qualitative study on including people with low literacy skills and learning disabilities. *BMC Health Serv Res.* 2012;26(12):431.