

Comparing EQ-5D-3L and EQ-5D-5L in Measuring the HRQoL Burden of 4 Health Conditions in China

Yani Tan

Jinan University College of Pharmacy <https://orcid.org/0000-0002-9476-6124>

Zhen Zhang

Jinan University College of Pharmacy

Yanming Hong

Jinan University College of Pharmacy

Zhihao Yang

Jinan University College of Pharmacy

Nan Luo

National University Singapore Saw Swee Hock School of Public Health

Jie Jiang

Jinan University College of Pharmacy

Sha Li (✉ tlisha@jnu.edu.cn)

Jinan University College of Pharmacy

Research article

Keywords: EQ-5D-3L, EQ-5D-5L, health-related quality of life, disease burden, China

Posted Date: September 25th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-76958/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: This study was conducted to compare the measurement properties of both the EQ-5D-3L and EQ-5D-5L in quantifying health related quality of life (HRQoL) burden for 4 different health conditions in China.

Methods: A total of 500 participants from China were recruited to complete the 3L and 5L questionnaire via Internet. The 3L and 5L were compared in terms of distribution properties, ceiling effects and relative efficiency. Distributions of responses were also examined.

Results: Ceiling effects decreased significantly in the 5L in all domains ($P \leq 0.01$), especially in the pain/discomfort dimension (relative difference: 43.10%). Relative efficiency suggested that 5L had a higher absolute discriminatory power than the 3L version between perfectly healthy participant and 4 condition groups, especially for the HIV/AIDS group when the 3L results was not significant. 5L also had smoother and less clustered index value distributions.

Conclusions: Both the 3L and 5L are demonstrated to be valid HRQoL instruments for quantifying disease burden in China. The 5L system may be preferable to the 3L, as it demonstrated superior performance with respect to lower ceiling effect, better relative efficiency and distributions.

Background

EQ-5D has been used to measure health related quality of life (HRQoL) across the globe [1]. The EQ-5D questionnaire for data collection consists of two essential parts: a multidimensional health descriptive system and the EQ visual analogue scale (EQ-VAS). The EQ-5D descriptive system comprises five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression [2]. It has two versions, a three-level (EQ-5D-3L) and a five-level version (EQ-5D-5L). EQ-5D-3L (hereinafter 3L) was developed in 1987 and has been the most popular preference-based instrument. It is reported that the 3L has suboptimal sensitivity and suffers from ceiling effects [3]. Therefore, EQ-5D-5L (hereinafter 5L) was introduced in 2009 [4]. In total, 3L defines a total of 243 unique health states, while 5L defines 3,125 health states. The higher number of health states described by the 5L version is aimed at improving sensitivity to small differences or changes in HRQoL [5].

So far, a large number of studies have compared the 5L and 3L and many studies have reported that the 5L is superior to the 3L in terms of psychometric properties such as ceiling effect and descriptive efficiency [6, 7]. However, the evidence of how the two versions of EQ-5D performed in China is limited. In particular, there lacks study comparing the ability of 3L and 5L to quantify disease burden or differentiate between individuals with and without a particular disease or condition in China. So far, only You et al studied 4 aging-related medical conditions: metabolic syndrome (including high blood pressure, high blood sugar, etc.), arthritis, osteoporosis, and other conditions as one category. A limitation of that study is that a between-group design was used where the two versions of EQ-5D were administered to two

different groups of individuals. As a result, findings in that study might be biased due to incomparability of the study groups.

In this study, we aimed to compare the sensitivity of the two versions of EQ-5D in quantifying the HRQoL burden associated with 4 chronic conditions including chronic hepatitis B, depression, anxiety, and AIDS/HIV in China. These conditions are not aged-related and is likely to affect young people. Moreover, the HRQoL burden of these conditions are mainly in the mental or psychological domain which may be difficult for 3L to detect and necessitates a more sensitive measure.

Methods

Study design

This study utilized the data collected in a cross-sectional online survey in China. The survey was part of an international study called 'extending the QALY (E-QALY) project' [8]. The E-QALY project aims to develop a new quality of life measure. In the online survey of China, 500 respondents with and without a selected chronic condition were recruited to complete the E-QALY items, EQ-5D-3L, EQ-5D-5L and the Short Warwick-Edinburgh Mental Well-being Scale (SWEMWS). This data was collected between April and July 2019 online by Accent, an U.K online survey company. Quotas and inclusion criteria were applied to recruit a sample of 500 individuals in which there were similar numbers of individuals with generalized anxiety disorder (GAD), HIV/AIDS, chronic Hepatitis B patients, or depression, or without any of those 4 chronic conditions. The study was approved by the Ethic Committee of University of Sheffield, United Kingdom.

The online survey began by giving an outline of the research purpose. Participants were then asked to report their disease history. Eligible respondents reported their background information including education level, gender and age etc. Next, respondents were asked to respond to a battery of questionnaires including (in the order of) a subset of E-QALY items, the 3L/5L (half of the sample responded to 5L and the other half 3L), some more E-QALY items, the SWEMWS, the 5L/3L, and the EQ-VAS.

Instruments

The EQ-5D-3L and EQ-5D-5L were both preference based HRQoL instruments developed by the EuroQol Group. Both instruments have the same five health dimensions, i.e. mobility, self-care, usual activities, pain/discomfort, anxiety/depression. The difference is the 3L has three response levels (no problems, some problems, extreme problems) while the 5L has five response levels (no problems, slight problems, moderate problems, severe problems, and unable/extreme problems) for each dimension. To calculate the utility, we used the value set developed by Liu et al for the 3L [9] and the value set developed by Luo et al for the 5L [10].

Statistical analyses

We first described the characteristics of our sample and we reported 1) the distributions of the 3L and 5L utility scores by condition group; 2) the ceiling effects by condition group; 3) the differences in responses between healthy and each condition group; 4) the relative efficiency in distinguishing between healthy and each condition group. Data was analyzed using IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp (2013).

Ceiling effects were defined as the proportion of respondents indicating 'no problems' in all five dimensions [11]. Previous studies have shown that the use of 5L could reduce the ceiling effects. For this reason, we hypothesized that the 5L has lower ceiling effects and calculated the reduction of ceiling effects from 3L to 5L. We evaluated the reduction separately for each dimension and all five dimensions taken together.

Relative efficiency (RE) of the two EQ-5D index scores in differentiating between the known groups described above was evaluated using the ratio of F statistics derived from the ANOVA tests. The F-statistic was widely used to assess RE of measurement scales [12–14]. The index score with a higher F statistic would be deemed to be more efficient than its comparator since a higher value of F statistic is more likely to result in statistical significance. To understand the RE of the index scores, we compared the distributions of the responses to the EQ-5D dimensions between the healthy group and each of the condition group. Mann-Whitney test was used. For reference, we listed the median values of each dimension reported by the healthy group and 4 condition groups.

Results

In total, 500 individuals completed the online survey, including 140 healthy individuals, 122 individuals with hepatitis B, 107 with depression, 90 individuals with GAD and 101 with HIV/AIDS. Some respondents reported multiple conditions, e.g. 68 individuals reported both depression and GAD. In general, the whole study sample was young (mean age: 35.8, SD: 8.64) and well educated. The gender proportions of the five groups were generally balanced except for the group of HIV/AIDS, in which, about 87.1% of individuals were female. In terms of the age distribution, the healthy group was mostly young; the chronic hepatitis B group has more participants aged between 40 and 49; the depression and GAD groups had individuals from all four age groups and the HIV/AIDS group aged mainly from 30 to 49. Individuals with tertiary education accounted for over 80% for all four disease groups and the healthy group had more individuals with secondary education. Table 1 shows the demographic information by condition [see Additional file 1].

There were 13, 34, 28, 26 and 26 unique states reported for the healthy, chronic hepatitis B, depression, GAD and HIV/AIDS group respectively for the 3L. The corresponding numbers were 18, 43, 46, 42 and 35 for the 5L.

It is evident from table 2 that the ceiling effects of the 5L was smaller than the 3L for all dimensions and for all condition groups [see Additional file 2]. When all dimensions are considered, the number of health profiles 11111 decreased by almost 40% when reporting using the 5L. The reduction of ceiling effect is

more salient in the pain/discomfort and anxiety/depression dimensions. The most prominent difference in ceiling effects between 3L and 5L was observed in the HIV/AIDS group, i.e. all dimensions had a relative reduction of over 30%.

Table 3 shows the median responses to 3L and 5L dimensions for the 5 groups [see Additional file 3]. When measured by the 3L, the median responses of healthy and HIV/AIDS groups were 'no problems' across the five dimensions. In comparison, when measured by the 5L, while the median responses for the healthy group remained to be 'no problems', the median responses of HIV/AIDS group were all 'slight problems'. Similarly, the median responses for the chronic hepatitis B group were 'slight problems' for the last three dimensions when measured by the 5L, while the only pain/discomfort dimension had 'moderate problem' measured by the 3L. The median responses for the GAD and depression group were similar. The Mann-Whitney results were all significant at 0.01 level suggesting all 4 condition groups had a different distribution of responses against the healthy group, for both 3L and 5L.

For the 3L, the healthy group had a mean utility of 0.948 (SD: 0.104), followed by HIV/AIDS of 0.791 (SD: 0.116), chronic hepatitis B of 0.781 (SD: 0.166), depression of 0.721 (SD: 0.180) and GAD of 0.711 (SD: 0.178). In comparison, the mean utilities of the 5L were 0.947 (SD: 0.080) for the healthy group, 0.785 (SD: 0.180) for the chronic hepatitis B group, 0.779 (SD: 0.115) for the HIV/AIDS group, 0.746 (SD: 0.202) for the depression group and 0.718 (SD: 0.217) for the GAD group. As shown in Figure 1, the 5L index value had a smoother distributions, less clustering and wider range (-0.305 to 1) compared with the 3L (ranged from -0.03 to 1). An exception is the 5L utility score clustered at 0.734 for the HIV/AIDS group. This is the utility score of health state 22222, which had been reported most frequently in the HIV/AIDS group.

Table 4 shows the mean index values of each condition groups and shows the relative efficiency of the 3L and 5L between groups [see Additional file 4]. The index values of the 5L ranged from 0.711 of the GAD group to 0.948 of the healthy group. In comparison, the index values of the 3L ranged from 0.718 of the GAD group to 0.947 of the healthy group. It is clear that two versions of EQ-5D produced comparable index values for each sub-group and both versions demonstrated good known-group validity. Compared with the 5L index, the 3L index did not show a difference of statistically important in the comparison of the healthy and HIV/AIDS groups. The relative efficiency of the 5L index was higher in the comparison of the healthy group with the condition groups.

Discussion

Our study used both the 3L version and the 5L version of EQ-5D to measure the HRQoL burden of 4 chronic conditions in China and focused on comparing the measurement sensitivity of these two EQ-5D versions. In general, we found both versions of EQ-5D to be sensitive tools to quantify the HRQoL loss caused by the 4 chronic conditions. When combined with the life expectancy data, quality-adjusted life years (QALYs) can be calculated and can be used as a standard measure to reflect how a condition could affect the length and the quality of an individual, which provided a single metric to reflect disease burden.

Disease burden is typically measured using disability adjusted life years (DALYs), but QALYs can also be used and may be a better measure as it could provide individual level HRQoL based data and it is recommended measure for economic evaluations [15].

For the measurement properties, our results generally agrees with the findings reported before, that is, although both versions produced highly agreed responses, index values and both had good known-group validity, the 5L performed better in terms of ceiling effects, index value distributions and relative efficiency [6, 16, 7]. The ceiling effect is most evident for the pain/discomfort and anxiety/depression dimensions. For the HIV/AIDS group, the use of 5L largely improved the ceiling effects and provided more accurate index value for the HIV/AIDS patients. In overall, around 40% of relative reduction of ceiling effect was observed in terms of reporting health state 11111. This was larger than the numbers reported in other studies, which ranged between 6.9% and 33.7% [7, 16]. A possible explanation is the characteristics of the patients in our study was unique in a sense that many patients did not report any problems using the 3L, but reported mild problems using the 5L. This clearly shows the limitation of less cutting-off points provided by the 3L descriptive system, which was first reported by Mathieu F et al [6]. Hence, the 5L has more advantage in measuring the HRQoL of individuals with mild health conditions.

For the index value distributions, it was observed that the 3L had more gaps and clustering's than the 5L. Two major factors are in play to decide the distribution of the index value, i.e. the health profiles and characteristics of the value set [17]. Notably in the 3L, there was a large gap between the index value 1 (profile 11111) and the second highest index value 0.887 (profile 11211). In the 5L, there are 5 different profiles worse than 11111 but having an index value higher than 0.9. In terms of the clusterings, 3L resulted in more clusterings than the 5L and the reason for the clusterings of the 3L index values is due to the clusterings of the profiles [17]. In comparison, there were still some clusterings in the 5L distributions, but it is not due to the clusterings of profiles, instead, it is because more profiles were reported and some profiles have similar index values. In overall, the increased levels of the 5L defined more health states and provided more accurate index values.

The relative efficiency results favored the 5L and is in line with the study of You et al. Based on the F-ratios in Table 4, it appears that 5L is more sensitive than 3L in physical diseases (i.e. CHB and HIV/AIDS). This is mainly because in the two mental condition group (i.e. GAD and depression), the profiles were more similar, i.e. in Table 3, the median levels for both 3L and 5L were both 11122 for the mental condition group. In contrast, the median levels of the HIV/AIDS group were 22222 when used the 5L and were 11111 when used the 3L. This was also observed in the CHB group, where the median level of usual activities was slight problems in the 5L, but no problems in the 3L. The 5L could pick up more accurate problems than the 3L, and when used the 3L, respondents tent to report no problems. The minimal advantage of 5L compared to 3L in differentiating individuals with and without a mental disorder could be due to the weakness of EQ-5D in measuring mental health. There is only one item in the instrument targeting mental problems.

There are some limitations in this study. First, the study sample was young and highly educated. It should be due to the fact that old people and less educated people are less active on Internet. Therefore, findings of this study may not be generalizable to older populations. It has been reported the elderly with less education attainment may have more difficulty to use the five level descriptive system and in such occasions, the 3L may be a more feasible alternative [18]. Second, some respondents reported more than one condition, but we did not provide deep analysis about the possible effect of multi-conditions. It should be noted that our sample, which was recruited online and the health condition was self-reported. Ideally, clinical data is used to verify the presence and absence of diagnoses reported by the study subjects. Due to these limitations, the superiority of the 5L warrants further studies.

Conclusions

In this study, both EQ-5D versions could quantify the disease burden of both physical and mental diseases in terms of self-reported HRQoL. Overall, our findings favor the 5L version of EQ-5D for measuring health related quality of life burden caused by chronic disease in China. As patient-reported outcome (PRO) is increasingly being used all over the world to measure disease burden, EQ-5D can be a useful tool in this context.

Abbreviations

HRQoL: Health related quality of life

EQ-VAS: EQ visual analogue scale

3L: EQ-5D-3L

5L: EQ-5D-5L

E-QALY : Extending the QALY

SWEMWS: Short Warwick-Edinburgh Mental Well-being Scale

GAD: Generalized anxiety disorder

RE: Relative efficiency

QALYs: Quality-adjusted life years

DALYs: Disability adjusted life years

PRO: Patient-reported outcome

Declarations

(1) Ethics approval and consent to participate

The study was approved by the Ethic Committee of University of Sheffield, United Kingdom. Written, informed consent was obtained from all participants included in the study.

(2) Consent for publication

Manuscript is approved by all authors for publication.

(3) Availability of data and material

Not applicable.

(4) Competing interests

No Competing interests exist in the submission of this manuscript.

(5) Funding

We thank Natural Science Foundation of Guangdong Province (No. 2020A1515010990) to SL and Science and Technology Program of Guangzhou, China (201704020198) for funding this study.

(6) Author contributions

Yani Tan and Zhen Zhang contributed to acquisition of data. Zhihao Yang and Yanming Hong analyzed and interpreted the data. Yani Tan and Zhen Zhang draft and revise the manuscript. Sha Li, Jie Jiang and Nan Luo and contributed to conception, design and the interpretation of the results. All authors reviewed and approved the final version.

(7) Acknowledgements

Not applicable.

References

1. Devlin, N. J., & Brooks, R. EQ-5D and the EuroQol Group: Past, Present and Future. *Appl Health Econ Health Policy*. 2014;15(2):127-137. doi:10.1007/s40258-017-0310-5.

2. Szende, A., Janssen, B., & Cabases, J. M. Self-reported population health: an international perspective based on EQ-5D. New York: Springer. 2014.
3. Devlin, N. J., & Krabbe, P. F. The development of new research methods for the valuation of EQ-5D-5L. *Eur J Health Econ.* 2013;14 Suppl 1:S1-3. doi:10.1007/s10198-013-0502-3.
4. Janssen, M. F., Pickard, A. S., 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):1717-1727. doi:10.1007/s11136-012-0322-4.
5. 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). *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation.* 2011;20(10):1727-1736. doi:10.1007/s11136-011-9903-x.
6. Janssen, M. F., Bonsel, G. J., & Luo, N. (2018). 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):675-697. doi:10.1007/s40273-018-0623-8.
7. Buchholz, I., Janssen, M. F., Kohlmann, T., & Feng, Y. S. (2018). A Systematic Review of Studies Comparing the Measurement Properties of the Three-Level and Five-Level Versions of the EQ-5D. *Pharmacoeconomics.* 2018;36(6):645-661. doi:10.1007/s40273-018-0642-5.
8. SchHARR, t. U. o. S. Extending the QALY. <http://scharr.dept.shef.ac.uk/e-qaly>. Accessed 4th June 2018.
9. Liu, G. G., Wu, H., Li, M., Gao, C., & Luo, N. Chinese time trade-off values for EQ-5D health states. *Value Health.* 2014;17(5):597-604. doi:10.1016/j.jval.2014.05.007.
10. 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):662-669. doi:10.1016/j.jval.2016.11.016.
11. Terwee, C. B., Bot, S. D. M., de Boer, M. R., van der Windt, D. A. W. M., Knol, D. L., Dekker, J., et al. Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology.* 2007;60(1):34-42. doi:10.1016/j.jclinepi.2006.03.012.
12. Pan, C. W., Sun, H. P., Wang, X. Z., Ma, Q. H., Xu, Y., Luo, N., et al. The EQ-5D-5L index score is more discriminative than the EQ-5D-3L index score in diabetes patients. *Quality of Life Research.* 2015;24(7):1767-1774. doi:10.1007/s11136-014-0902-6.
13. Luo, N., Johnson, J. A., Shaw, J. W., & Coons, S. J. Relative efficiency of the EQ-5D, HUI2, and HUI3 index scores in measuring health burden of chronic medical conditions in a population health survey in the United States. *Med Care.* 2009;47(1):53-60. doi:10.1097/MLR.0b013e31817d92f8.
14. Vickrey, B. G., Hays, R. D., Genovese, B. J., Myers, L. W., & Ellison, G. W. Comparison of a generic to disease-targeted health-related quality-of-life measures for multiple sclerosis. *J Clin Epidemiol.* 1997;50(5):557-569.
15. Kennedy-Martin, M., Slaap, B., Herdman, M., van Reenen, M., Kennedy-Martin, T., Greiner, W., et al. Which multi-attribute utility instruments are recommended for use in cost-utility analysis? A review of

national health technology assessment (HTA) guidelines. The European Journal of Health Economics. 2020. doi:10.1007/s10198-020-01195-8.

16. You, R., Liu, J., Yang, Z., Pan, C., Ma, Q., & Luo, N. Comparing the performance of the EQ-5D-3 L and the EQ-5D-5 L in an elderly Chinese population. Health Qual Life Outcomes. 2020;18(1):97. doi:10.1186/s12955-020-01324-0.
17. Feng, Y., Devlin, N., Bateman, A., Zamora, B., & Parkin, D. Distribution of the EQ-5D-5L Profiles and Values in Three Patient Groups. Value in Health. 2019;22(3):355-361. doi:<https://doi.org/10.1016/j.jval.2018.08.012>.
18. Yang, F., Jiang, S., He, X.-n., Li, H.-c., Wu, H.-y., Zhang, T.-t., et al. Do Rural Residents in China Understand EQ-5D-5L as Intended? Evidence From a Qualitative Study. Pharmacoeconomics - Open. 2020. doi:10.1007/s41669-020-00212-z.

Figures

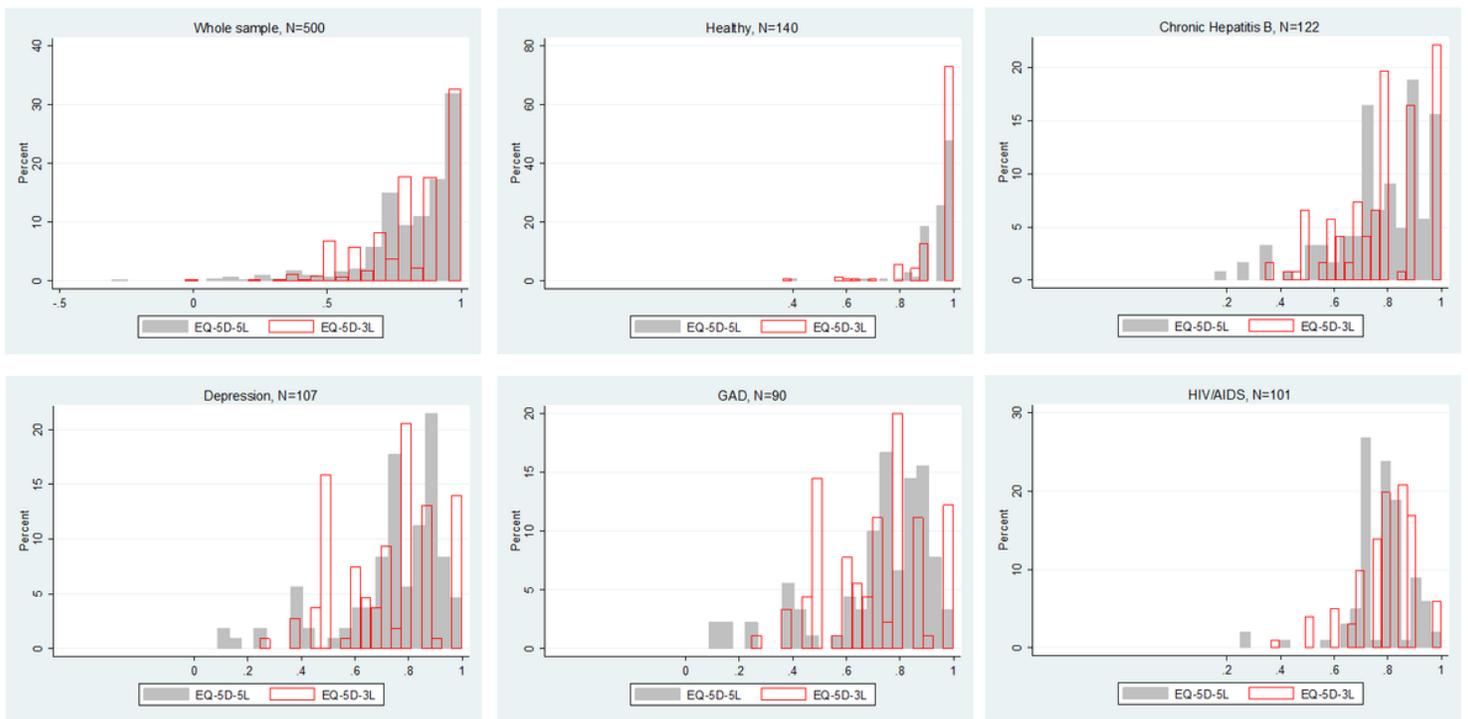


Figure 1

Distributions of utility score for the whole sample and by condition groups

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supplement1.xlsx](#)

- [supplement2.xlsx](#)
- [supplement3.xlsx](#)
- [supplement4.xlsx](#)