BMC Psychology

https://doi.org/10.1186/s40359-026-05428-y

Article in Press

Trends in WHOQOL-BREF and EUROHIS-QOL-8 health-related quality of life scores among the Thai population from the 1990s to 2023: an application of Rasch analysis

Received: 10 October 2025

Accepted: 20 August 2026

Published online: 26 August 2026

Cite this article as: Kangwanrattanakul K., Medvedev O.N. & Krägeloh C.U. Trends in WHOQOL-BREF and EUROHIS-QOL-8 health-related quality of life scores among the Thai population from the 1990s to 2023: an application of Rasch analysis. *BMC Psychol* (2026). https://doi.org/10.1186/s40359-026-05428-y

Krittaphas Kangwanrattanakul, Oleg N. Medvedev & Christian U. Krägeloh

Trends in WHOQOL-BREF and EUROHIS-QOL-8 health-related quality of life scores among the Thai population from the 1990s to 2023: an application of Rasch analysis

Krittaphas Kangwanrattanakul, PhD[1, *], Oleg N. Medvedev, PhD[2], Christian U. Krägeloh, PhD[3]

[1]*Division of Social and Administrative Pharmacy, Faculty of Pharmaceutical Sciences, Burapha University, 169 Long-Hard Bangsaen Rd., Mueang, Chonburi 20131, Thailand*

[2] *School of Psychological and Social Sciences, University of Waikato, Hamilton, New Zealand*

[3] *Department of Psychology and Neuroscience, Auckland University of Technology, Auckland, New Zealand*

*****Corresponding author

E-mail: krittaphas@buu.ac.th, krittaphas@go.buu.ac.th

Krittaphas Kangwanrattanakul, ORCID: 0000-0002-4921-8508

Oleg N. Medvedev, ORCID: 0000-0002-2167-5002

Christain U. Krägeloh, ORCID: 0000-0002-7298-0736

Abstract

Background The World Health Organization Quality of Life-brief (WHOQOL-BREF) and the EUROHIS-QOL-8, an abridged version of WHOQOL-BREF, tools have been used widely in Thailand to assess health-related quality of life (HRQoL) in the general population. The availability of data spanning multiple time points now presents an opportunity to explore long-term trends in HRQoL scores and examine whether Thai respondents' interpretations of questionnaire items have changed over time. Therefore, the primary objective of this study was to examine long-term trends of WHOQOL-BREF scores collected at three time points. Analyses of differential item functioning (DIF) were also performed to investigate whether Thai individuals interpreted and endorsed WHOQOL-BREF and EUROHIS-QOL-8 consistently across time to support the validity of these comparisons.

Methods This study conducted Rasch analysis on small, randomly selected Bangkok-only subsamples (n=150 per time point) from three datasets (n = 300 in 1990s, n = 1200 pre-COVID 2019, and n = 2000 post-COVID 2023) using the WHOQOL-BREF and EUROHIS-QOL-8 instruments. DIF explored temporal changes in item interpretation across age, sex, and dataset.

Results Mean HRQOL scores derived from four WHOQOL-BREF domain and EUROHIS-QOL-8 increased across the three time points; however, social and environmental domains did not show any statistically significant difference (p > 0.05). DIF revealed that physical and psychological domains were more difficult to score high for the

samples from the 1990s compared to those from the two time periods, whereas social and environmental domains displayed an opposite effect.

Conclusions Although these findings highlight the increasing HRQoL trends in Bangkok and suggest the importance of considering societal changes when interpreting long-term HRQoL data, an investigation with a nationally representative sample of Thai population is warranted for future studies to confirm these findings.

Keywords: WHOQOL-BREF, Thai population, Rasch analysis, Health-related quality of life, DIF analysis

1. Introduction

Health-related quality of life (HRQoL) is a multidimensional concept affected by physical and psychological states (1). HRQoL measures have become a major outcome variable to assess the impact and effectiveness of healthcare interventions, and it can also be used to measure overall health of general population in national health surveys (2). Given the demand for HRQoL measurement, generic and condition-specific instruments have been employed to measure HRQoL levels for various purposes (3).

One of the most widely used HRQoL measures are those developed by the World Health Organization Quality of Life (WHOQOL) Group. The original WHOQOL-100 has 100 items covering a range of health facets related to physical, psychological, level of independence, social relationships, environment, and spirituality domains (1). It is considered to have strong cross-cultural validity due to its development in a collaboration of field centers in 14 countries (1).

Due to the length and associated response when completing the WHOQOL-100, an abbreviated version, the WHOQOL-BREF, was subsequently developed, which includes 26 items covering diverse four health domains such as physical, psychological, social, and environmental domains (4, 5). The WHOQOL-BREF has been shown to be a valid and reliable instrument in a wide range of populations (5-7).

As the WHOQOL tool was developed with the involvement of a Thai field center, a Thai version of the WHOQOL-BREF has been available since the 1990s (8). Classical test theory approaches for psychometric testing revealed that it is valid and reliable to measure HRQoL levels among the general population (9), elderly samples (10), university students (11), and those with health conditions such as cancer (12) and human immunodeficiency virus/acquired immunodeficiency syndrome (13). A recent Rasch analysis study (14, 15) also confirmed that the Thai WHOQOL-BREF is a valid instrument for the general Thai population both for total scores and individual subscales (16).

The Rasch model assumes that a measurement scale functions consistently across different groups being assessed (17, 18). Differential item functioning (DIF) helps evaluate how individuals from different sociodemographic subgroups (such as age, sex) respond to specific HRQoL items. Through DIF analysis, researchers can identify whether subgroups interpret and respond to items differently, with significant differences determined through analysis of variance (ANOVA) with Bonferroni correction (19).

Thailand is an upper-middle income country with a population of approximately 65.9 million (20, 21). With 94% of the population adhering to Buddhism (22), this high level of religious homogeneity sets Thailand apart, particularly from Western countries, and may contribute to a distinct perspective on HRQoL items (1, 23). The COVID-19 pandemic during 2019-2022 also had a significant impact on health and Thai health policy, prompting the government to implement various infection control measures (24). Many previous studies suggested that both physical and psychological health could be diminished after the pandemic compared to before (25-28). Additionally, Thailand shifts to face its newer challenge for the increased growth of aging population and non-communicable diseases (29). As a result, the above-mentioned changes in public health indicators, economic development, and responses to the COVID-19 pandemic may be expected to have affected both actual HRQoL levels as well as how individuals in Thailand perceive HRQoL items over the past three decades.

To facilitate faster assessment of HRQoL and to reduce response burden, several abridged versions of the WHOQOL have been developed, namely the 26-item WHOQOL-BREF, the 8-item EUROHIS-QOL-8, and the 5-item WHOQOL-5, which have been used wide range of population across the globe including Thailand (16, 30-32). For the Thai general population, however, previous Rasch analysis only supported the use of the WHOQOL-BREF and EUROHIS-QOL-8, while it revealed that the WHOQOL-5 lacks satisfactory reliability (16). In order to explore trends of HRQoL in the Thai population over time, the present study compared scores on these two instruments using datasets collected at three time points: (a) the original WHOQOL study conducted in the early 1990s, (b) data collected prior to the COVID-19 pandemic in 2019, and (c) data collected after the pandemic in 2023. Nevertheless, meaningful WHOQOL-BREF score comparisons require evidence to support that the questionnaire measures the same underlying construct consistently across three time points. Therefore, Rasch analysis was performed to evaluate the measurement invariance of the WHOQOL-BREF and EUROHIS-QOL-8 across these three time points, while DIF analysis could reveal whether each WHOQOL-BREF and EUROHIS-QOL-8 item functioned differently over time.

2. Methods

2.1 Dataset and study samples

This study was a secondary data analysis comparing quality of life measurements across three time periods in Thailand using data from three previous studies: 1) The WHOQOL development study conducted in the 1990s, which used the WHOQOL-100 (1), 2) a general population study conducted in 2019 (Pre-COVID-19 pandemic), for which WHOQOL-BREF data are available (33), and 3) a study about population norm measurement of health-related quality of life in the general Thai population in 2023 (Post-COVID-19 pandemic) (34), for which WHOQOL-BREF data are available. All study participants gave their informed consent to participate before the commencement of each study.

The WHOQOL-100 dataset collected samples from Bangkok, Thailand using a quota sampling with regard to some sociodemographic factors: age (50% > 45, 50% ≤ 45), sex (50% male, 50% female), health status (50 healthy samples, 250 samples with impaired health status). Some adjustments to the samples of Studies 2 and 3 were required due to differences in sampling approaches. While the WHOQOL-100 dataset was collected with samples in Bangkok, the other two datasets employed a four-stage stratified random sampling to select provinces, districts, sub-districts, and villages for data collection. To avoid confounding due to potential differences between rural and urban populations, the datasets for Studies 2 and 3 were modified to exclude participants not residing in Bangkok. All surveys had collected data through paper-based face-to-face interviews with convenient samples at their household.

Age and sex were the only common demographic information across three datasets. There were some variations of samples' age across these three datasets because the pre-COVID-19 dataset recruited individuals aged from 20 to 70 years, while the other datasets recruited participants equal to and above the age of 18 years. To create comparable sub-samples, we randomly selected 150 participants aged between 20 and 70 years residing in Bangkok from each dataset in order to minimize the variation of the samples' characteristics across these three datasets, resulting in 450 participants included in this analysis. This sample size is considered optimal for the purposes of Rasch analysis (35, 36).

2.2 Instruments

2.2.1 WHOQOL-BREF

The WHOQOL-BREF has 26 items categorized into four dimensions: physical health (7 items), psychological health (6 items), social relationships (3 items), and environment (8 items). There are two additional items measuring general health and overall quality of life which are not assigned to any of the four domains. For each item, participants rate their HRQoL during the past 2 weeks with a 5-point Likert scale. WHOQOL scores can be computed as domain scores derived from summating each item score belonging to each dimension, where higher scores indicate a better HRQoL. As with other WHOQOL-BREF versions, the Thai version contains three negatively worded items that require reverse scoring before domain score computation. However, these are presented in a different order, with Item 2 expressing the facet *pain*, Item 9 *negative feelings*, and Item 11 *medication*. It should be noted that the WHOQOL-BREF items for Sample 1 were extracted from the WHOQOL-100, while the other two surveys employed the WHOQOL-BREF for HRQoL measurement in 2019 and 2023.

2.2.2 EUROHIS-QOL-8

The EUROHIS-QOL-8 was extracted from the corresponding WHOQOL-BREF items to develop a shorter version of WHOQOL-BREF (30). Many previous studies have employed this approach for psychometric validation purpose to ensure the cross-comparison between the WHOQOL-BREF and its shorter version without imposing a respondent's burden from administering two questionnaires concurrently (16, 32). It has eight items sourced from the WHOQOL-BREF, namely satisfaction with health (Item 1), energy (Item 3), self-esteem (Item 7), daily life activities (Item 10), personal relationships (Item 13), home environment (Item 16), financial support (Item 17), and overall QOL (Item 26). Response options and recall period follow those of WHOQOL-BREF questionnaire. Each item contributes to the total score, with higher scores indicate better health status. For Sample 1, EUROHIS-QOL-8 scores were calculated using items extracted from the WHOQOL-100, and for Samples 2 and 3, items were extracted from the WHOQOL-BREF.

2.3 Data Analyses

None of these participants in the combined sample of 450 participants had missing values. To compare scores across time points, means and standard deviations were calculated for each of the items of the 26-item WHOQOL-BREF and 8-item EUROHIS-QOL-8. Ordinal logistic regression was employed to compare the scores from dimension- and item-level of both instruments across the three time points. Prior to running ordinal logistic regression, proportional odds assumptions were tested and were not violated ($p > 0.05$), indicating that predictor effects were stable across

response levels. Then, an omnibus test was performed where item- or domain-level score served as the dependent variable, while time point (1990s, Pre-COVID-19 pandemic, and Post-COVID-19 pandemic) was included as an independent variable along with age and sex as covariates to avoid potential confounding effects when running each ordinal logistic regression. If an omnibus test was significant, a post-hoc comparison was subsequently performed to compare the mean scores of dimension- and item-level of both instruments using Bonferroni adjustment.

Prior to performing DIF analysis, we investigated whether Rasch model fits were adequate for each of the domains as well as the unidimensional EUROHIS-QOL-8 profile. A nonsignificant Rasch model fits indicate all items functioned consistently and measured a construct of interest, HRQoL, with an acceptable psychometric performance. A fit was considered adequate if the chi-square value for item-trait interaction was not significant. If data did not fit the Rasch model, a subtest approach was used by merging items that exhibited high residual correlations or shared similar content. The scale was considered unidimensional if either: 1) fewer than 5% of t-tests comparing the most dissimilar items (identified through principal component analysis of residuals as loading high or low on the first principal component) were statistically significant, or 2) the lower bound of the confidence interval for the percentage of significant t-tests fell below the 5% threshold (37). Person Separation Index (PSI) was used to define the scale reliability where a value ≥ 0.70 indicates satisfactory reliability (38).

DIF analysis explored potential differences in item response pattern across the three datasets, (WHOQOL-100, Pre-COVID-19, and Post-COVID-19 datasets) and it was performed to explore the response pattern by age (20-34, 35-49, and 50-70 years) and sex (male/female). While ANOVA with Bonferroni correction ($p < 0.05$) was used to highlight potential DIF (19). Furthermore, we identified significant DIF whether it was uniform DIF, indicating a consistent difference in response to WHOQOL-BREF domains/items and EUROHIS-QOL-8 items across all class intervals, or non-uniform DIF, indicating that the difference in responses varied across class intervals of different levels of latent trait. Uniform-DIF was identified when WHOQOL-BREF domain/items and EUROHIS-QOL-8 items showed statistical significance after adjusting with Bonferroni correction ($p < 0.05$) for dataset, while a non-uniform DIF was identified when demonstrating statistical significance after adjusting with Bonferroni correction ($p < 0.05$) for interactions between dataset and class interval. The magnitude of DIF was quantified for all flagged items by splitting each affected item by dataset and reporting the difference in item location estimates (in logits) between groups. Following the ETS classification (39, 40), DIF magnitudes below 0.43 logits were classified as negligible, 0.43–0.64 as slight to moderate, and ≥ 0.64 as moderate to large. This corresponds closely to the

widely used 0.5-logit criterion (41), which approximates at least slight to moderate DIF magnitude based on ETS classification. The uniform DIF significance was also assessed using the z statistic, calculated as the difference in item locations divided by the pooled standard error and compared with the critical value of 1.96 (41). Conversely, the effect size of non-uniform DIF was quantified by computing the Partial Eta-squared (η_p^2) using the Sum of Squares (SS) interaction divided by SS (interaction) + SS (error) from the ANOVA output (42), and the magnitude of the effect size was interpreted using the conventional benchmarks of $\eta_p^2 =$ 0.01 (small), 0.06 (medium), and 0.14 (large) (43, 44). Readers who are unfamiliar with the Rasch methodology and its interpretations may refer to two recent tutorial-style publications that provide detailed, non-technical background information (45-47). Descriptive statistics were performed using STATA 17 (StataCorp LLC, College Station, TX, USA), while Rasch analysis was performed using RUMM 2030 (48).

3. Results

Table 1 shows descriptive statistics of the three samples' demographic characteristics. The distribution of sex was similar across the three datasets ($\chi^2[2] = 3.06$, $p = 0.22$), with a slightly higher proportion of female respondents (62.7%) for the WHOQOL-100 dataset, compared to 54.0% for both pre-and post-COVID-19 datasets. However, the age difference across the three samples was significant (F [2, 447] = 15.03, $p < 0.05$), whereby the respondents in the WHOQOL-100 dataset were significantly younger compared to those of the other two datasets.

Table 2 shows the overall Rasch model fit statistics for the 24-item WHOQOL-BREF, four-domain WHOQOL-BREF, and 8-item EUROHIS-QOL-8. The baseline model for the WHOQOL-BREF, treating all 24 items as a unidimensional scale, did not achieve adequate model fit ($\chi^2[72] = 294.20$ $p < 0.01$). When the 24-item WHOQOL-BREF was categorized into four domain subtests, a non-significant fit to the Rasch model was obtained ($\chi^2[12] = 19.81$, $p = 0.07$). Additionally, unidimensionality was confirmed, and reliability was satisfactory (PSI = 0.85). The EUROHIS-QOL-8 also achieved an adequate fit ($\chi^2[24] = 33.76$, $p = 0.09$) without the need for modifications. There was no evidence against unidimensionality, and reliability was satisfactory (PSI = 0.76). These findings indicated that all four WHOQOL-BREF domains and EUROHIS-QOL-8 items measured an intended construct of interest (HRQoL) with satisfactory reliability.

The mean scores by dataset for each of the WHOQOL-BREF items, the four domain WHOQOL-BREF scores, and the EUROHIS-QOL-8 scores are shown in Table 3. Controlling for age and sex, there was a significant increasing trend in the mean scores from the 1990s dataset to the post-COVID-19 dataset for almost all WHOQOL-BREF items except for sleep (Q4), activities of daily living (Q10), self-esteem (Q7), personal relationship (Q13), social support (Q14), and security (Q15).

Although the post-COVID-19 generally had the highest mean scores among the three datasets (Table 3), a pairwise post-hoc comparison showed that some WHOQOL-BREF items did not show statistically significant different mean scores between pre- and post-COVID-19 datasets (all $p > 0.05$) including energy (Q3), positive feeling (Q5), concentration (Q6), bodily image (Q8), home environment (Q16), work capacity (Q12), financial support (Q17), health care (Q18), leisure activity (Q20), physical environment (Q21), and overall QoL (Q26). As also shown in Table 3, the total domain scores showed a similar trend with all domain scores being lowest for Sample 1 (1990s) and highest for Sample 3 (post-COVID-19 dataset). However, this difference was not significant for the social and environmental domains according to the omnibus test (all $p > 0.05$). For the psychological domain, the post-hoc test revealed no significant difference between Samples 2 and 3 ($p = 0.16$). For the EUROHIS-QOL-8, Sample 1 scores were also lowest and Sample 3 scores highest. While the omnibus test was significant, the only significant difference in the post-hoc tests was between Samples 1 and 3 ($p < 0.01$).

Because the Rasch analysis showed that the higher-order WHOQOL-BREF fit with the domain subtests met the assumptions of the Rasch model (Table 2), DIF analysis for the four-domain WHOQOL-BREF could be performed to explore potential differences in domain subtest performance across the three datasets (Figure 1). The ANOVA output showed that there was no significant DIF by sex and age. However, significant uniform DIF by dataset (all $p < 0.05$) was detected for all of the WHOQOL-BREF domain subtests. Both physical health and environmental dimensions showed moderate to large DIF magnitude, while the psychological health and social dimensions had negligible DIF magnitude. As shown in Table 4, the physical health and environmental dimension displayed uniform DIF by dataset of 0.833 logits (SE = 0.040, z = 20.983, p < .001) and of 0.691 logits (SE = 0.041, z = 17.016, p < .001), respectively. The psychological health domain demonstrated uniform DIF by dataset of 0.331 logits (SE = 0.046, z = 7.134, p < .001); however, its magnitude was negligible (ETS category A) reflecting the sensitivity of significance tests to sample size; interpretation was therefore based on the magnitude classification.

Regarding the DIF direction, both the physical and psychological domains were more difficult to endorse for the participants from the 1990s than those pre- and post-COVID-19. Conversely, the social and environmental domains were easier to endorse for the participants in the 1990s than those from the other two time periods. Because the Rasch fits for the individual domain analyses required the use of subtests, no meaningful DIF analysis by item could be conducted. However, as shown in Table 2, the EUROHIS-QOL-8 achieved the commendable fit to the Rasch model without the need for subtests. For that reason, DIF analysis was performed for each of the 8 items of the EUROHIS-QOL-8 by dataset. Significant DIF was detected for the items on general health (Q1), personal relationships (Q13), and home environment (Q16) (all $p < 0.05$). According to the ANOVA with Bonferroni adjustment at $p < 0.05$, these three items exhibited moderate to large uniform-DIF by dataset. As also shown in Table 4, Q1, Q13 and Q16 displayed uniform DIF of 0.944 logits (SE = 0.134, z = 7.052, p <.001), 1.034 logits (SE=0.162, z = 6.367, p <.001), and 0.806 logits (SE = 0.156, z = 5.155, p < 0.001), respectively. Regarding the DIF direction, Q1 was more difficult to score high for the participants from the 1990s compared to the other two datasets, as indicated by the DIF graph (Fig. 2a). In contrast to the item on general health (Q1), personal relationships (Q13) and home environment (Q16) were easier to endorse for the participants from the 1990s than those from the other two datasets (Fig. 2e and Fig. 2f). Nevertheless, a financial support item (Q17) displayed non-uniform DIF across cohorts ($F(6, 432) = 4.71$, $p < .001$) with moderate effect size (partial $\eta^2 = .061$). DIF direction demonstrated the relative ordering of cohorts changing across class intervals where the 1990s cohort showed the lowest endorsement at the lowest level of the trait but exceeded the PreCOVID cohort at the second-class interval, with the later cohorts higher at the upper intervals (Fig. 2g).

It should be noted that no significant DIF against sex and age was detected for all 8-item EUROHIS-QOL-8 indicating that the items functioned equally across these demographic factors and the item response pattern was stable across different levels of the latent trait

4. Discussion

The present study investigated changes in HRQoL among the general Thai samples residing in Bangkok across three time points (1990s, Pre- and Post-COVID-19 pandemic). Such information is valuable to illustrate to what extent Thai society has changed over the last 30 years and also as a result of the COVID-19 pandemic. Overall, most of the WHOQOL-BREF items showed an increasing trend in their mean scores across these three time points.

Similar to a previous Rasch analysis on a Thai sample (16), the present study showed an acceptable fit to the Rasch model, unidimensionality, and a satisfactory reliability for the general Thai population across three time points. Furthermore, a shorter version, EUROHIS-QOL-8, continued to achieve a commendable fit to the Rasch model and unidimensionality with an acceptable reliability. These indicate that both four WHOQOL-BREF dimensions and EUROHIS-QOL-8 items functioned consistently and measured HRQoL level with an acceptable psychometric performance across three time points.

Although both the physical health and psychological health domains exhibited uniform DIF, only the physical health domain showed a moderate to large DIF magnitude, whereas the psychological health domain showed negligible DIF according to the ETS classification and Draba's logit rule. When looking into the graphical displays for DIF directions, they showed that both physical and psychological domains were more difficult to endorse for the samples in the 1990s compared to the other two datasets, resulting in lower mean scores for both dimensions in the 1990s than the other two datasets. This is consistent with findings from similar previous studies in Hong Kong (49) and Zhejiang in China (50), which showed a significant rise in both physical and mental health dimensions from the 1990s to 2022, although those two dimensions were assessed by the SF-12v2 for the Hong Kong study (49) and by a newly developed HRQoL questionnaire for Zhejiang. Conversely, a study in Tianjin, China (51), revealed a downward trend in both health dimensions from 2008 to 2020, as determined by the EQ-5D questionnaire. Notably, our HRQoL analyses only controlled for two sociodemographic factors (age and sex), which means that its upward trend might have been also affected by other factors. Furthermore, another possible reason for the upward trend in HRQoL level could have been the provision of universal health coverage for the general Thai population, and promotion of healthy behaviors such as annual check-ups among women. The improved quality of the Thai health care system also appears to be linked to increased life expectancy and reduced mortality from communicable diseases (52), resulting in improved health status of general Thai people since 2002. However, this mixed finding with the previous studies should be further reinvestigated using various HRQoL instruments.

Regarding the social domain, it demonstrated a uniform DIF; however, its magnitude was negligible. The DIF direction from the graphical display showed that it was easier to endorse for the samples in 1990s than the other two datasets although its trend of mean scores was decreasing without any significant difference across these three time points ($p = 0.07$). Interestingly, when looking in more detail at the social domain items of the WHOQOL-BREF, the decreased mean scores might be due to the sexual activity item (Q25) because the 1990s

sample showed significantly lower HRQoL scores compared to post-COVID-19. Notably, for another item belonging to the social domain, personal relationship (Q13), respondents from the 1990s reported higher mean scores than those from the two other time points (Pre- and Post-COVID-19). The lower endorsement of the personal relationship (Q13) item might reflect societal changes in which interpersonal relationships might have deteriorated in the recent cohorts, as the Thai society has gradually become a more independent or individualistic society (53). Furthermore, increased urbanization and a shift from extended to nuclear families may have reduced personal interactions with people in the vicinity including family members and close social networks. Notably, this analysis was performed with individuals residing in Bangkok; therefore, these findings could reflect the characteristics of Thai urban society which may have become more Westernized (54). These explanations are also confirmed by the DIF graph for this item from the EUROHIS-QOL-8 (Fig.2e), according to which a higher probability of endorsement of this item was observed for the samples from the 1990s as opposed to the more recent cohorts.

With regard to the environmental domain, its mean scores for the samples from the 1990s were lower than those of the other two time points without statistically significant differences across these three time points. However, it was easier to endorse higher scores on this domain for the sample from the 1990s compared to those from the pre- and post-COVID-19 datasets. Within the environmental domain, *home environment* (Q16) was rated highest in the 1990s compared to those from other two time points. We reasoned that rapid industrial development, and urban expansion may have led to a number of environmental problems including air pollution, noise, traffic congestion, and environmental degradation around residential areas which might have resulted in decreased satisfaction with their home environment (55-57). Furthermore, the rapid growth of urbanization accompanied by increased population density, smaller living space, and widened disparities of housing styles could have resulted in less favorable perception of home environments for individuals in more recent cohorts (58).

Nevertheless, those observed patterns of WHOQOL-BREF and EUROHIS-QOL-8 item interpretation and endorsement explained by various factors including individualization, changes in family structure, urbanization, environmental conditions, and housing characteristics might be a possible contextual explanation from the previous literature. Notably, these factors were not directly assessed in this present study because of the limited sociodemographic covariates and cross-sectional nature of the data. Future research incorporating these variables in a longitudinal design is therefore required to highlight any causal effects.

Unlike some previous studies (59, 60), our study revealed that the mean scores for the four-domain WHOQOL-BREF and EUROHIS-QOL-8 from the Post-COVID-19 dataset were higher than those of the pre COVID-19 dataset. Specifically, physical, psychological, and environmental health scores increased across the three time points, from the 1990s to the post-COVID-19 period in 2023, whereas the social health score fluctuated over time. However, statistically significant increases were observed only for physical and psychological health scores, while the changes in social and environmental health scores did not reach statistical significance. These findings should be interpreted with caution because we did not collect the infection status of the individuals and other pandemic-related factors from the Post-COVID-19 dataset. Furthermore, this analysis was performed with samples from Bangkok. Therefore, future studies should reinvestigate the effect of the COVID-19 pandemic influencing HRQoL scores with larger, nationally representative samples of the general Thai population to confirm these findings.

It is important to note that the present study also has several limitations. First, all participants were from one city (Bangkok) with an imbalanced proportion of samples' age across these three datasets. This was due to the fact that the sample from the 1990s was collected in Bangkok only, which meant that location-matched participants needed to be selected for the comparative Samples 2 and 3. Second, we selected sub-samples from each cohort in Bangkok to ensure comparability for Rasch and DIF analysis across three time points. However, restricting the study to Bangkok and using relatively small subsamples may have limited the representativeness of the findings and their generalizability to the broader Thai population. Rasch and DIF analysis are required to reinvestigate with a larger and representative of general Thai population to confirm these findings across multiple time points in future studies. Third, there were only two sociodemographic factors could be reported in this study because the WHOQOL-100 dataset from the 1990s collected only age and sex of the samples as sociodemographic factors, which means that it was not possible to investigate the potential effects of other sociodemographic factors affecting the HRQoL levels across these three datasets. Fourth, we did not collect the infectious status and other pandemic-related factors, so we could not investigate the causal relationship between those factors and the HRQoL scores. Future study is greatly required to collect those factors and investigate the causal relationship with HRQoL scores. Fifth, this study used the latest dataset collected in 2023 because it was the most recent data when performing data analysis. It is relatively outdated in the current time and might not fully reflect the current Thai health profile. The updated analysis using the more recent available data of health profile is therefore encouraged. Finally, the EUROHIS-QOL-8 items were

extracted from the corresponding WHOQOL-BREF items rather than using the stand-alone EUROHIS-QOL-8 instrument. Although the items were identical for the shorter version, response patterns might have been altered by completing the longer WHOQOL-BREF due to different contexts and respondent's fatigue. Therefore, a stand-alone EUROHIS-QOL-8 should be administered with the Thai samples to investigate the measurement properties across multiple cohorts in future studies.

5. Conclusions

The HRQoL levels measured by both four-domain WHOQOL-BREF and EUROHIS-QOL-8 revealed the increasing trend of HRQoL mean scores was observed among the general Thai samples over the period from 1990s to 2023. However, the social and environmental domains did not show statistically significance differences among three time points. The DIF analysis also showed that both physical and psychological domains were more difficult to endorse for the samples in the 1990s, while social and environmental domains showed otherwise. These findings suggest that improvements in physical and psychological health can contribute to enhanced overall HRQoL level, while monitoring social and environmental health is still required. However, a broader policy implication should wait until the confirmation from future studies using nationally representative samples of the general Thai population.

List of abbreviations

ANOVA Analysis of Variance

COVID-19 Coronavirus disease starting in 2019

DIF Differential item functioning

EQ-5D EuroQol five-dimension

HRQoL Health-related quality of life

PSI Person Separation Index

SF-12v2 The short form health survey-version 2

WHOQOL-BREF World Health Organization Quality of Life-brief

Declarations

Ethics approval and consent to participate

The protocol and data collection process for the three datasets was approved by the ethical committees: Burapha University Institutional Review Board (IRB 108/2562 and IRB1-031/2566) for pre and post COVID-19 datasets. All

participants gave their informed consent to participate before the study commenced. Furthermore, it adhered to the principles outlined in the Declaration of Helsinki.

Consent for publication

Not Applicable

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Competing interests

The authors declare no competing interests.

Funding

Both Pre COVID-19 and Post COVID-19 datasets were funded by the Research Grant of Burapha University through National Research Council of Thailand under Grant No. Rx2/2562 and Rx6/2566, respectively.

Author contributions

K.K., O.M., and C.K. designed the study. K.K. performed the data collection and formal analysis, while C.K. rechecked the results and interpretation of the formal analysis. K.K. prepared the first draft of the manuscript; O.M. and C.K. edited the manuscript. The final manuscript was reviewed and approved by all authors. The content has not been published and under consideration elsewhere.

Acknowledgements

We would like to express our sincere gratitude to the director of Suan Prung Psychiatric Hospital's director for their valuable permission to use the Thai WHOQOL-BREF for data collection of the two datasets (Pre and Post COVID-19 pandemic). We also would like to thank all well-trained interviewers for facilitating all data collection with all convenient samples. Special thanks go to all participants involved in those three datasets for their valuable time to participate in this study. The Thai WHOQOL-100 data from the 1990s is jointly owned by the WHOQOL Group collaborators in the countries and WHO, Geneva, and was shared with researchers based at the New Zealand WHOQOL Centre. The WHOQOL-100 dataset used in these studies will not be made publicly available. Requests to access these data should be directed to Dr. Shekhar Saxena, Division of Mental Health and Substance Abuse, World Health Organization, Geneva (saxenas@who.int).

References

1. The World Health Organization Quality of Life Assessment (WHOQOL): development and general psychometric properties. Soc Sci Med. 1998;46(12):1569-85.

2. Phyo AZZ, Freak-Poli R, Craig H, Gasevic D, Stocks NP, Gonzalez-Chica DA, et al. Quality of life and mortality in the general population: a systematic review and meta-analysis. BMC Public Health. 2020;20(1):1596.

3. Churruca K, Pomare C, Ellis LA, Long JC, Henderson SB, Murphy LED, et al. Patient-reported outcome measures (PROMs): A review of generic and condition-specific measures and a discussion of trends and issues. Health Expect. 2021;24(4):1015-24.

4. Development of the World Health Organization WHOQOL-BREF quality of life assessment. The WHOQOL Group. Psychol Med. 1998;28(3):551-8.

5. Skevington SM, Lotfy M, O'Connell KA. The World Health Organization's WHOQOL-BREF quality of life assessment: psychometric properties and results of the international field trial. A report from the WHOQOL group. Qual Life Res. 2004;13(2):299-310.

6. Kalfoss MH, Reidunsdatter RJ, Klöckner CA, Nilsen M. Validation of the WHOQOL-Bref: psychometric properties and normative data for the Norwegian general population. Health Qual Life Outcomes. 2021;19(1):13.

7. Goes M, Lopes M, Marôco J, Oliveira H, Fonseca C. Psychometric properties of the WHOQOL-BREF(PT) in a sample of elderly citizens. Health Qual Life Outcomes. 2021;19(1):146.

8. Department of Mental Health . Quality of life indicators of the World Health Organization Thai version (WHOQOL-bref-Thai). https://www.dmh.go.th/test/whoqol/. Accessed Aug 28 2025.

9. Mahatnirundkul S. Comparison of the WHOQOL-100 and the WHOQOL-BREF (26 items). J Ment Health Thai. 1998;5:4-15.

10. Taboonpong S, Suttharangsee W, Chailangka P. Evaluating psychometric properties of WHO quality of life questionnaire in Thai elderly. J Gerontol Geriatric Med. 2001;2:6-12.

11. Li K, Kay NS, Nokkaew N. The Performance of the World Health Organization's WHOQOL-BREF in Assessing the Quality of Life of Thai College Students. Social Indicators Research. 2009;90(3):489-501.

12. Phungrassami T, Katikarn R, Watanaarepornchai S, Sangtawan D. Quality of life assessment in radiotherapy patients by WHOQOL-BREF-THAI: a feasibility study. J Med Assoc Thai. 2004;87(12):1459-65.

13. Sakthong P, Schommer JC, Gross CR, Sakulbumrungsil R, Prasithsirikul W. Psychometric properties of WHOQOL-BREF-THAI in patients with HIV/AIDS. J Med Assoc Thai. 2007;90(11):2449-60.

14. Kaplan RM, Hays RD. Health-Related Quality of Life Measurement in Public Health. Annu Rev Public Health. 2022;43:355-73.

15. Medvedev ON, Krägeloh CU. Rasch Measurement Model. In: Medvedev ON, Krägeloh CU, Siegert RJ, Singh NN, editors. Handbook of Assessment in Mindfulness Research. Cham: Springer International Publishing; 2022. p. 1-18.

16. Kangwanrattanakul K, Krägeloh CU. Psychometric evaluation of the WHOQOL-BREF and its shorter versions for general Thai population: confirmatory factor analysis and Rasch analysis. Qual Life Res. 2024;33(2):335-48.

17. Tennant A, Pallant J. DIF matters: A practical approach to test if Differential Item Functioning makes a difference. Rasch Measurement Transactions. 2007;20:1082-4.

18. Smith RM. Fit analysis in latent trait measurement models. J Appl Meas. 2000;1(2):199-218.

19. Andrich D, Hagquist C. Real and Artificial Differential Item Functioning in Polytomous Items. Educ Psychol Meas. 2015;75(2):185-207.

20. The World Bank Data. GNI per capita, Atlas method (current US$) 2021. https://data.worldbank.org/indicator/NY.GNP.PCAP.CD. Accessed 14 March 2023.

21. National Statistical Office. Number of Population from Registration, Male, Female and Households by Region and Province: 2021-2024.https://www.nso.go.th/nsoweb/nso/statistics_and_indicators?impt_branch=300. Accessed 28 Aug 2025.

22. Winzer L, Samutachak B, Gray RS. Religiosity, spirituality, and happiness in Thailand from the perspective of Buddhism. Journal of Population and Social Studies. 2018;26(4):332-43.

23. Kleinman A, Eisenberg L, Good B. Culture, illness, and care: clinical lessons from anthropologic and cross-cultural research. Annals of internal medicine. 1978;88(2):251-8.

24. Yorsaeng R, Suntronwong N, Thongpan I, Chuchaona W, Lestari FB, Pasittungkul S, et al. The impact of COVID-19 and control measures on public health in Thailand, 2020. PeerJ. 2022;10:e12960.

25. Ishikawa H, Kato M, Kiuchi T. Declines in health literacy and health-related quality of life during the COVID-19 pandemic: a longitudinal study of the Japanese general population. BMC Public Health. 2021;21(1):2180.

26. Pieh C, Budimir S, Probst T. The effect of age, gender, income, work, and physical activity on mental health during coronavirus disease (COVID-19) lockdown in Austria. J Psychosom Res. 2020;136:110186.

27. Smith P, De Pauw R, Van Cauteren D, Demarest S, Drieskens S, Cornelissen L, et al. Post COVID-19 condition and health-related quality of life: a longitudinal cohort study in the Belgian adult population. BMC Public Health. 2023;23(1):1433.

28. Malik P, Patel K, Pinto C, Jaiswal R, Tirupathi R, Pillai S, et al. Post-acute COVID-19 syndrome (PCS) and health-related quality of life (HRQoL)-A systematic review and meta-analysis. J Med Virol. 2022;94(1):253-62.

29. Potempa K, Rajataramya B, Singha-Dong N, Furspan P, Kahle E, Stephenson R. Thailand's Challenges of Achieving Health Equity in the Era of Non-Communicable Disease. Pac Rim Int J Nurs Res Thail. 2022;26(2):187-97.

30. Schmidt S, Mühlan H, Power M. The EUROHIS-QOL 8-item index: psychometric results of a cross-cultural field study. Eur J Public Health. 2006;16(4):420-8.

31. Geyh S, Fellinghauer BA, Kirchberger I, Post MW. Cross-cultural validity of four quality of life scales in persons with spinal cord injury. Health Qual Life Outcomes. 2010;8:94.

32. Balalla SK, Medvedev ON, Siegert RJ, Krägeloh CU. Validation of the WHOQOL-BREF and Shorter Versions Using Rasch Analysis in Traumatic Brain Injury and Orthopedic Populations. Arch Phys Med Rehabil. 2019;100(10):1853-62.

33. Kangwanrattanakul K, Parmontree P. Psychometric properties comparison between EQ-5D-5L and EQ-5D-3L in the general Thai population. Qual Life Res. 2020;29(12):3407-17.

34. Kangwanrattanakul K, Krägeloh CU. EQ-5D-3L and EQ-5D-5L population norms for Thailand. BMC Public Health. 2024;24(1):1108.

35. Linacre J. Sample size and item calibration stability. Rasch Mes Trans. 1994;7:328.

36. Hagell P, Westergren A. Sample Size and Statistical Conclusions from Tests of Fit to the Rasch Model According to the Rasch Unidimensional Measurement Model (Rumm) Program in Health Outcome Measurement. J Appl Meas. 2016;17(4):416-31.

37. Smith EV, Jr. Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. J Appl Meas. 2002;3(2):205-31.

38. Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? Arthritis Rheum. 2007;57(8):1358-62.

39. Seo W. Psychometric Evaluation of the Social Anxiety Scale for Adolescents in a School-Based Sample of Korean Youths: A Rasch Analysis. Asian Nurs Res (Korean Soc Nurs Sci). 2025.

40. Zwick R, Thayer DT, Lewis C. An empirical Bayes approach to Mantel‐Haenszel DIF analysis. Journal of Educational Measurement. 1999;36(1):1-28.

41. Draba R. The identification and interpretation of item bias (Research Memorandum No. 25). Chicago, IL: University of Chicago. 1977.

42. Mordkoff JT. A simple method for removing bias from a popular measure of standardized effect size: Adjusted partial eta squared. Advances in Methods and Practices in Psychological Science. 2019;2(3):228-32.

43. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.

44. Richardson JT. Eta squared and partial eta squared as measures of effect size in educational research. Educational research review. 2011;6(2):135-47.

45. Pierobon A, Krägeloh C. Rasch analysis in the development of self-reported outcome measures to assess physical function. Musculoskelet Sci Pract. 2026;83:103544.

46. Adu P, Bartholomew E, Medvedev ON. Rasch Methodology for Assessment in International Contexts. In: Krägeloh CU, Alyami M, Medvedev ON, editors. International Handbook of Behavioral Health Assessment. Cham: Springer Nature Switzerland; 2023. p. 1-30.

47. Tristan A. An adjustment for sample size in DIF analysis. Rasch Measurement Transactions. 2006;20(3):1070-1.

48. Andrich D, Sheridan B, Luo G. RUMM 2030 (Beta Version for Windows) Perth. Western Australia: RUMM Laboratory Pty Ltd. 2009.

49. Tang EHM, Wong CKH, Bedford LE, Yu EYT, Tse ETY, Dong W, et al. Trend in health-related quality of life and health utility and their decrements due to non-communicable diseases and risk factors: analysis of four population-based surveys between 1998 and 2015. Qual Life Res. 2020;29(11):2921-34.

50. Li N, Xu C, Ma G, Yu M, Wang H, Liu X. Trends and Causes of Educational Inequality in Health-Related Quality of Life in China During the COVID-19 Pandemic. Applied Research in Quality of Life. 2024.

51. Wang D, Xie S, Wu J, Sun B. The trend in quality of life of Chinese population: analysis based on population health surveys from 2008 to 2020. BMC Public Health. 2023;23(1):167.

52. Tangcharoensathien V, Witthayapipopsakul W, Panichkriangkrai W, Patcharanarumol W, Mills A. Health systems development in Thailand: a solid platform for successful implementation of universal health coverage. Lancet. 2018;391(10126):1205-23.

53. Wiwattanapantuwong J, Tuicomepee A, Suttiwan P, Watakakosol R, Ben-Ezra M, Goodwin R. For better or worse: Relationship change in Thailand during COVID-19. PLoS One. 2022;17(5):e0264614.

54. Pinyuchon M, Gray LA. Understanding Thai Families: A Cultural Context for Therapists Using a Structural Approach. Contemporary Family Therapy. 1997;19(2):209-28.

55. Chuersuwan N, Nimrat S, Lekphet S, Kerdkumrai T. Levels and major sources of PM2.5 and PM10 in Bangkok Metropolitan Region. Environment International. 2008;34(5):671-7.

56. Vichit-Vadakan N, Vajanapoom N. Health impact from air pollution in Thailand: current and future challenges. Environ Health Perspect. 2011;119(5):A197-8.

57. Komin S. A social analysis of the environmental problems in Thailand. Asia's environmental crisis: Routledge; 2019. p. 257-74.

58. Friel S, McMichael AJ, Kjellstrom T, Prapamontol T. Housing and health transition in Thailand. Rev Environ Health. 2004;19(3-4):311-27.

59. Nshimirimana DA, Kokonya D, Gitaka J, Wesonga B, Mativo JN, Rukanikigitero JMV. Impact of COVID-19 on health-related quality of life in the general population: A systematic review and meta-analysis. PLOS Glob Public Health. 2023;3(10):e0002137.

60. Xu F, Brodszky V. The impact of COVID-19 on health-related quality of life: a systematic review and evidence-based recommendations. Discover Psychology. 2024;4(1):90.

Table 1 Sample characteristics

Characteristics	**WHOQOL-100/ 1990s (*n*=150)**	**Pre COVID-19 (*n*=150)**	**Post COVID-19 (*n*=150)**	***p*-value**
Sex, n (%)				0.22
Male	56 (37.3)	69 (46.0)	69 (46.0)	
Female	94 (62.7)	81 (54.0)	81 (54.0)	

Age group, n(%)							<0.01	
20-34		70 (46.7)	42 (28.0)		38 (25.3)			
Analyses	**Measures**	**Item fit residual (mean±*SD*)**	**Person fit residual (mean±*SD*)**	**Goodness of fit**		**PSI**	**Significant *t*-test (Unidimensionality)**	
				χ^2	***p*-value**		**%**	**Lower bound CI**
WHOQOL-BREF								
1	24-item WHOQOL-BREF baseline	0.30±2.33	−0.49±2.00	294.20 (72)	<0.01	0.91	7.78	5.76
2	4-domain subtests	0.49±1.72	−0.43±1.08	19.81 (12)	0.07	0.85	6.22	4.21
WHOQOL-BREF domains								
3	Physical baseline (7 items)	0.42±1.69	−0.51±1.46	63.27 (42)	0.02	0.72	1.56	−0.46
4	Physical with 2 subtests	0.04±1.16	−0.56±1.35	23.89 (30)	0.78	0.77	1.12	−0.90
5	Psychological baseline (6 items)	0.17±1.33	−0.52±1.24	69.66 (36)	<0.01	0.68	4.67	2.65
6	Psychological with 1 subtest	−0.35±2.97	−0.49±0.98	28.78 (24)	0.23	0.73	4.02	2.00
7	Social baseline (3 items)	0.23±1.00	−0.62±1.20	27.95 (6)	<0.01	0.39	2.00	−0.01
8	Social with 1 subtest	−0.44±3.34	−0.61±0.90	9.75 (10)	0.46	0.54	5.35	3.33
9	Environment baseline (8 items)	−0.15±1.23	−0.50±1.35	84.55 (48)	<0.01	0.80	7.57	5.56
10	Environment with 1 subtest	−0.89±4.10	−0.49±1.08	42.09 (30)	0.07	0.82	3.80	1.78
EUROHISQOL-8								
11	Initial/Final	−0.09±1.29	−0.50±1.40	33.76 (24)	0.09	0.76	4.68	1.11
35-49		46 (30.7)	57 (38.0)		43 (28.7)			
50-70		34 (22.7)	51 (34.0)		69 (46.0)			
Age (Mean, [SD])		37.8 (12.5)	43.3 (13.2)		46.2 (14.6)		<0.01	

Table 2 Summary of overall Rasch model fit statistics for the WHOQOL-BREF and EUROHIS-QOL-8 items across three datasets

Table 3 Mean scores of 26- item WHOQOL-BREF, 4-domain WHOQOL-BREF, and EUROHIS-QOL-8 for three datasets.

Measures	Mean scores (Mean±SD)			
	WHOQOL-100 (A)	Pre COVID-19 (B)	Post COVID-19 (C)	*p*-value
1. Thai WHOQOL-BREF items				
1.1 Physical health				
Pain (Q2)	3.05±1.28	3.70±0.93	4.07±0.79	Omnibus <0.001††
				A vs B < 0.001**
				A vs C < 0.001**
				B vs C = 0.001*
Energy (Q3)	3.64±1.06	3.73±0.62	3.83±0.77	Omnibus = 0.018†
				A vs B = 0.191
				A vs C =0.018
				B vs C = 0.214
Sleep (Q4)	3.58±0.97	3.55±0.87	3.54±0.89	Omnibus = 0.723
Activities of daily living (Q10)	3.76±0.91	3.69±0.68	3.84±0.71	Omnibus = 0.287
Dependence on medical aids (Q11)	3.03±1.24	4.13±0.91	4.43±0.81	Omnibus <0.001††
				A vs B < 0.001**
				A vs C < 0.001**
				B vs C < 0.001**
Work capacity (Q12)	3.37±0.99	3.59±0.72	3.73±0.72	Omnibus <0.001††
				A vs B = 0.006*
				A vs C < 0.001**
				B vs C = 0.181
Mobility (Q24)	3.55±0.94	3.87±0.69	4.14±0.68	Omnibus <0.001††
				A vs B < 0.001**
				A vs C < 0.001**
				B vs C < 0.001**
1.2 Psychological health				
Positive feeling (Q5)	2.97±0.91	3.70±0.70	3.71±0.70	Omnibus <0.001††
				A vs B < 0.001**
				A vs C < 0.001**
				B vs C = 0.987
Concentration (Q6)	3.06±0.84	3.71±0.70	3.71±0.72	Omnibus <0.001††
				A vs B < 0.001**
				A vs C < 0.001**
				B vs C = 0.762
Self-esteem (Q7)	3.62±0.97	3.74±0.69	3.81±0.71	Omnibus = 0.052
Bodily-image (Q8)	3.55±1.03	3.76±0.75	3.79±0.71	Omnibus = 0.017†
				A vs B = 0.029
				A vs C = 0.014*
				B vs C = 0.736

Table 3 Mean scores of 26- item WHOQOL-BREF, 4-domain WHOQOL-BREF, and EUROHIS-QOL-8 for three datasets (cont.)

Table 3 Mean scores of 26- item WHOQOL-BREF, 4-domain WHOQOL-BREF, and EUROHIS-QOL-8 for three

Measures	**Mean scores (Mean±*SD*)**			
	WHOQOL-100 (A)	Pre COVID-19 (B)	Post COVID-19 (C)	*p*-value
Negative feeling (Q9)	3.26±1.21	4.05±0.96	4.37±0.74	Omnibus <0.001††
				A vs B < 0.001**
				A vs C < 0.001**
				B vs C = 0.007
Personal belief (Q23)	3.95±0.89	3.87±0.69	4.14±0.71	Omnibus = 0.009†
				A vs B = 0.445
				A vs C < 0.010*
				B vs C <0.001**
1.3 Social relationship				
Personal relationship (Q13)	3.90±0.76	3.80±0.70	3.77±0.73	Omnibus = 0.370
Social support (Q14)	3.59±0.84	3.57±0.77	3.60±0.72	Omnibus = 0.677
Sexual activity (Q25)	3.36±0.87	3.25±0.90	3.97±0.80	Omnibus <0.001††
				A vs B = 0.576
				A vs C < 0.001**
				B vs C < 0.001**
1.4 Environmental health				
Security (Q15)	3.59±0.84	3.55±0.65	3.55±0.73	Omnibus = 0.358
Home environment (Q16)	3.86±0.85	3.53±0.67	3.60±0.73	Omnibus =0.001††
				A vs B < 0.001**
				A vs C = 0.001*
				B vs C =0.753
Financial support (Q17)	2.97±0.96	3.15±0.80	3.21±0.64	Omnibus =0.002†
				A vs B = 0.017*
				A vs C = 0.002*
				B vs C =0.435
Health care (Q18)	3.58±0.85	3.12±0.97	3.30±0.66	Omnibus = 0.001††
				A vs B < 0.001**
				A vs C < 0.001**
				B vs C = 0.151
Accessibility of needed information (Q19)	3.11±0.92	3.54±0.66	3.35±0.71	Omnibus = 0.005†
				A vs B < 0.001**
				A vs C = 0.002*
				B vs C = 0.015*
Leisure activity (Q20)	2.82±0.94	3.23±0.90	3.31±0.85	Omnibus <0.001††
				A vs B < 0.001**
				A vs C < 0.001**
				B vs C = 0.601

datasets (cont.)

Measures	Mean scores (Mean±*SD*)			
	WHOQOL-100 (A)	Pre COVID-19 (B)	Post COVID-19 (C)	*p*-value
Physical environment (Q21)	3.40±0.83	3.40±0.69	3.62±0.76	Omnibus =0.025† A vs B = 0.865 A vs C = 0.028 B vs C =0.032
Transport (Q22)	3.50±0.77	3.57±0.74	3.87±0.78	Omnibus < 0.001†† A vs B = 0.280 A vs C < 0.001** B vs C = 0.001*
1.5 General health and overall quality of life items				
General health (Q1)	2.87±1.13	3.53±0.76	3.79±0.76	Omnibus < 0.001†† A vs B < 0.001** A vs C < 0.001** B vs C = 0.005*
Overall QOL (Q26)	3.25±0.65	3.43±0.65	3.51±0.71	Omnibus = 0.002†† A vs B = 0.021 A vs C = 0.002* B vs C = 0.352
2. Four-domain totals				
Physical	23.99±4.33	26.26±3.45	27.57±3.60	Omnibus <0.001†† A vs B < 0.001** A vs C < 0.001** B vs C = 0.013*
Psychological	20.40±3.40	22.82±3.21	23.52±3.08	Omnibus <0.001†† A vs B < 0.001** A vs C < 0.001** B vs C = 0.160
Social	10.85±1.80	10.62±1.72	11.34±1.74	Omnibus = 0.070
Environmental	26.91±4.24	27.09±4.14	27.81±4.24	Omnibus =0.080
3. EUROHIS-QOL-8	27.86±4.19	28.59±3.75	29.35±4.00	Omnibus = 0.001†† A vs B = 0.116 A vs C = 0.002* B vs C = 0.109

For Omnibus tests: †$p < 0.05$, ††$p < 0.01$; for post-hoc tests, a Bonferroni correction was applied: *$p < 0.017$ (0.05/3), **$p < 0.001$ (0.01/3)

Table 4 Differential item functioning magnitude and interpretation for WHOQOL-BREF domain and EUROHIS-QOL-8 items by dataset

Measures	Location Group 1 (dataset1990s)	SE	Location Group 2 (dataset 2019& dataset2023)	SE	Absolute magnitude	SE (DIF)	Z-score	Draba's logit rule	ETS* categories	DIF** direction
WHOQOL-BREF domain										
Physical Health	0.488	0.030	-0.345	0.026	0.833	0.040	20.983	Exceeded	C	More difficult
Psychological Health	0.548	0.037	0.217	0.028	0.331	0.046	7.134	No	A	More difficult
Social Health	-0.288	0.057	-0.339	0.043	0.051	0.071	0.714	No	A	Easier
Environment Health	-0.049	0.032	-0.740	0.025	0.691	0.041	17.016	Exceeded	C	Easier
EUROHIS-QOL-8 items										
Satisfied with health (Q1)	1.610	0.100	0.666	0.089	0.944	0.134	7.052	Exceeded	C	More difficult
Personal relationships (Q13)	-1.540	0.131	-0.506	0.096	1.034	0.162	6.367	Exceeded	C	Easier
Home environment (Q16)	-1.245	0.121	-0.439	0.099	0.806	0.156	5.155	Exceeded	C	Easier

*A: Negligible, B: Slight to moderate, C: Moderate to large

**DIF direction for WHOQOL-BREF domain and EUROHIS-QOL-8 items in the 1990s cohort relative to the 2019 and 2023 coho

a

b

c

d

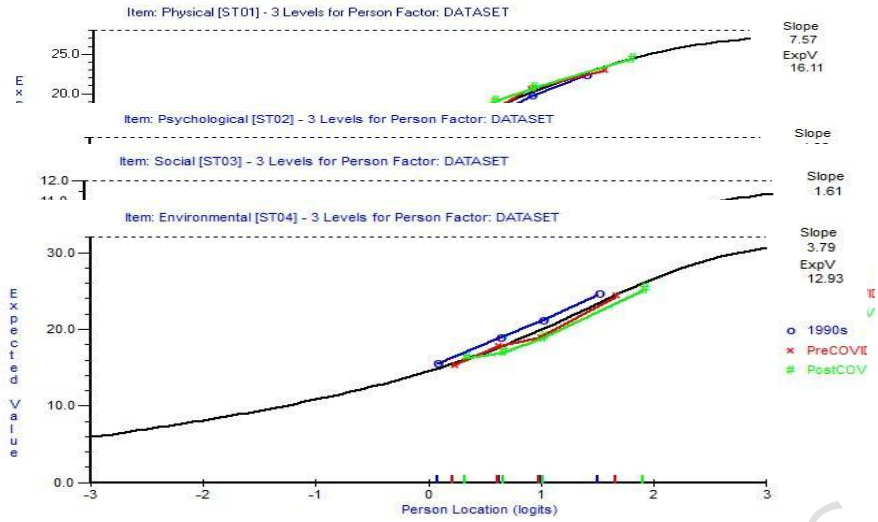

Figure 1 Graphical inspection of differential item functioning for (**a**) Physical domain, (**b**) Psychological domain, (**c**) Social domain, and (**d**) Environmental domain

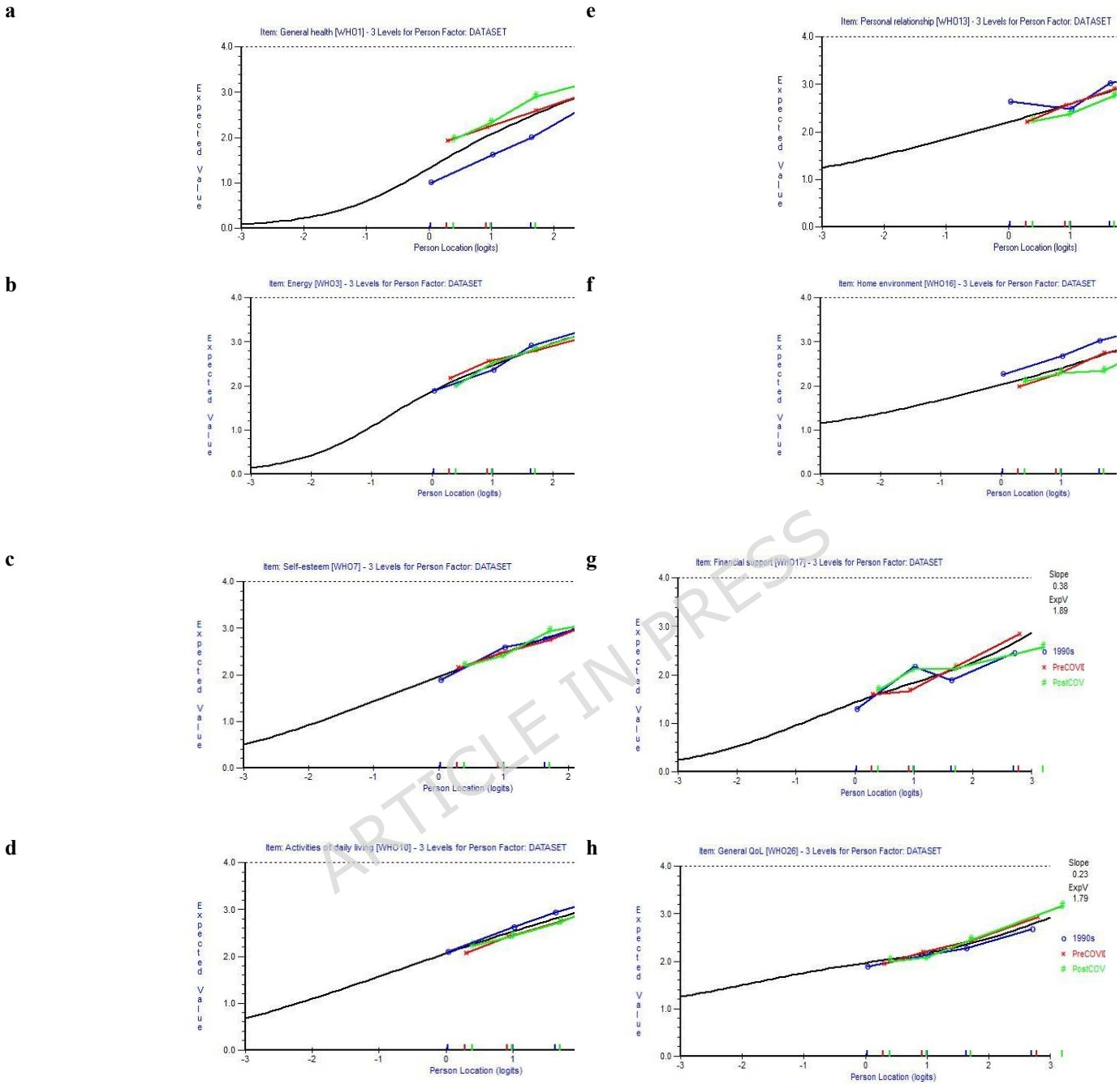

Figure 2 Graphical inspection of differential item functioning for EUROHIS-QOL-8 items (**a**) General health, (**b**) Energy, (**c**) Self-esteem, (**d**) Activities of daily living, (**e**) Personal relationships, (**f**) Home environment, (**g**) Financial support, and (**h**) General quality of life