Comparing health status in Belarus between 2001–10: a novel method using surveys with different response categories

Alena Vasianovich[1], Neil W. Scott[2], William C. S. Smith[1], Claire D. Wallace[3], Ludmila L. Shebeko[4]

1 Division of Applied Health Sciences, School of Medicine and Dentistry, College of Life Science and Medicine, University of Aberdeen, Aberdeen, UK
2 Division of Applied Health Sciences, Medical Statistics Team, School of Medicine and Dentistry, College of Life Sciences and Medicine, University of Aberdeen, Aberdeen, UK
3 Department of Sociology, School of Social Science, King's College, University of Aberdeen, Aberdeen, UK
4 Department of General and Clinical Medicine, Educational and Medical Centre, Polesski State University, Pinsk, Belarus

Correspondence: Alena Vasianovich, Division of Applied Health Sciences, School of Medicine and Dentistry, College of Life Science and Medicine, University of Aberdeen, Aberdeen, AB25 2ZD, UK, tel: +44 1224 437266; fax: +44 1224 437857; e-mail: r02av7@abdn.ac.uk

Background: Two population surveys were conducted in Belarus: The Living Conditions, Lifestyle and Health (LLH) in 2001 ($n = 2000$) and The Health in Times of Transition (HITT) in 2010 ($n = 1800$). Each survey included a question on health status. The LLH questionnaire provided a 4-point Verbal Response Scale, but the HITT questionnaire used a 5-point scale. When translated into Russian, only two response categories of these scales had identical wording. These differences made a direct comparison of self-reported health status between 2001 and 2010 difficult. **Methods:** We conducted a Health Category Response Scale (HCRS) survey in 2010 ($n = 570$) using a 100ths graduated Visual Analogue Scale (VAS) to understand how the response categories of different scales are perceived by Russian speakers. We implemented the HCRS survey's data to calculate the weighted health status (WHS) for each of the original surveys and to compare health status in Belarus between 2001 and 2010. **Results:** The WHS in Belarus showed a small, but statistically significant, improvement of 2.9 points on a 0–100 scale between 2001 and 2010 (56.2 vs. 59.1). Identical response categories were perceived differently on a 4-point and 5-point VAS. The category 'good' ('Хорошее') measured ~12 points higher, and the category 'bad/poor' ('Плохое') measured ~16 points lower, on the 4-point compared with the 5-point VAS. **Conclusion:** Our HCRS survey and novel method enabled a direct comparison of questions with different response options. When applied to the LLH and HITT projects, we concluded that health status in Belarus has improved between 2001 and 2010.

Introduction

Self-reported health status is an important indicator of people's actual health.[1,2] Individuals can describe their current health

according to how they feel, and self-reported health status has proved to be a strong predictor of mortality, morbidity and health care utilization.[1–4] Approved measures of subjective health are established in population health and lifestyle surveys[1,4,5] and in

health-related quality of life studies.[6–8] Various health measurement instruments have included questions on self-reported general health using different wording, layouts and response scales.[4,9–11] The numerical rating scales known as Visual Analogue Scales (VAS) have been widely used and evaluated to assess health status,[4,12,13] pain[14,15] and other symptoms in various studies.[16,17] Some health measurement instruments have used ordered categorical response options known as a Verbal Response Scale (VRS) or Likert scale,[9–11,18] or both.[19,20] This article presents the results of our study, the Health Category Response Scale (HCRS) survey that used a VAS to compare self-reported health status in Belarus between 2001 and 2010.

In the past decades, two European Union (EU)-funded projects (Institute for Advanced Studies, Vienna) investigated health, lifestyle and living conditions in several former Soviet Union (fSU) countries during the time of transition after dissolution of the Soviet Union in the 1990s.[9,10] These were the Living Conditions, Lifestyle and Health (LLH) project (2000–03)[21] and the Health in Times of Transition (HITT) project (2009–13), a cross-national follow-up project that continues the research of the LLH project.[22] The LLH project's population survey was conducted in 2001 in eight fSU countries (Armenia, Belarus, Georgia, Kazakhstan, Kirgizstan, Moldova, Russia and Ukraine),[21] and the HITT project's survey was conducted in 2010 in nine fSU countries (the original eight fSU countries plus Azerbaijan).[22] The design and development of each project's questionnaires have been described elsewhere.[23–27] Data from these two EU-funded projects were used to explore changes in self-reported health status in Belarus between 2001 and 2010.

Both the LLH and HITT surveys asked participants about their current health status, but the exact wording and response categories provided were not identical. In the LLH project, respondents were asked: 'How would you describe your state of health these days? Would you say it is...' with four possible response options (4-point VRS): (i) good; (ii) quite good; (iii) rather bad; and (iv) bad. In the HITT project, respondents were asked: 'In general, would you say your health is...' with five possible response options (5-point scale): (i) very good; (ii) good; (iii) fair; (iv) poor; or (v) very poor. When translated into Russian, only two of the response options from each questionnaire directly corresponded. The option 'good' from both questionnaires (the LLH and HITT) was translated as 'Хорошее'. The options 'bad' from the LLH project and 'poor' from the HITT project were translated as 'Плохое'. All other health response categories from the original projects were different in the English and Russian translations. The LLH instrument also included response options 'quite good' as 'Скорее хорошее' and 'rather bad' as 'Скорее плохое', and the HITT instrument also consisted of response options 'very good' as 'Очень хорошее', 'fair' as 'Удовлетворительное' and 'very poor' as 'Очень плохое'. As the response options used were different between these two projects, it is difficult to make a direct comparison of self-reported health status between the two time points (2001-LLH and 2010-HITT) using only data from the original EU projects.

Our study has two main objectives: (i) to devise a method to measure the strength of the Russian response categories used in the LLH and HITT projects; and (ii) to use these measurements to compare self-reported health status in Belarus between 2001 and 2010.

Methods

Our HCRS survey, based on a VAS and using self-completed questionnaires (in Russian), was carried out in three different geographical areas in Belarus: (i) Gorki (staff and patients of the outpatient clinic of the District Hospital); (ii) Mogilev (undergraduate students and staff from Mogilev State Technological University); and (iii) Pinsk (undergraduate students and staff from Polesski State University). A non-probability (convenience) sample of Russian speakers was chosen. The target group was adults aged ≥17 years, as the general age of secondary school graduates in Belarus is 16 years.

Research tool & VAS

Two HCRS questionnaires were developed: one questionnaire concerned the question on self-reported health status from the LLH project (2001) and the second one the question from the HITT project (2010). Part I of both new questionnaires included general socio-demographic questions regarding age, gender and education status of the study participants.

Both Part II and Part III of each HCRS questionnaire included a VAS (a vertical 100ths graduated ruler) with: (i) a 4-point scale for the LLH-related and (ii) a 5-point scale for the HITT-related questionnaires. Boxes with the verbal response categories (labels) from each of the original project were placed on the left side of the ruler. Figures showing the LLH-related and HITT-related VAS are available in the Supplementary Appendix (Supplementary figures S1 and S2, respectively). Participants were asked to rate verbal health response categories between 0 and 100 by drawing a line from the labelled box to the point on the VAS, where 0 represented the worst/unhealthiest grade and 100 the best/healthiest grade.

In Part II, respondents were asked about their own health status. First, they reported their current health status using either the 4-point or 5-point VRS (LLH-related or HITT-related questionnaire, respectively). Then, participants were asked to grade this on the VAS.

In Part III (the main part of each questionnaire), participants were asked to measure the strength of each Russian response category (either LLH-related or HITT-related) on a similar 0–100 VAS. The draft HCRS questionnaires (in Russian) were piloted in Belarus among a Russian-speaking population ($n = 10$). After minor amendments, the final version was administered.

Questionnaire administration

The HCRS survey took place in Belarus between September and November 2010. The LLH- and HITT-related questionnaires were allocated alternately ('quasi-randomization') and administered face-to-face. Each participant received only one version of the questionnaire: either the LLH- or the HITT-related version.

Statistical analysis

Data were analysed using PASW Statistics 18 software. The mean VAS score, standard deviation (SD) and range (minimum–maximum) of each health response category from our HCRS survey (4-point for the LLH and 5-point for the HITT) were calculated. An independent samples *t*-test or one-way analysis of variance test was conducted to compare mean VAS ratings by gender, age group and place of survey.[28,29] The weighted health status (WHS) was then calculated for the LLH and HITT projects using the following equation:[29]

$$\text{WHS} = \Big\{ [(N_1/N_t) \times \text{Mean}_1] + [(N_2/N_t) \times \text{Mean}_2]$$
$$+ \ldots\ldots + [(N_n/N_t) \times \text{Mean}_n] \Big\}$$

where:

- WHS – Weighted health status.
- $N_1 - \ldots - N_n$ – number of participants for each response category in original project ($n = 4$ for the LLH and $n = 5$ for the HITT).
- $\text{Mean}_1 - \ldots - \text{Mean}_n$ – mean rating for each response category based on VAS measurement from the HCRS survey.
- N_t – total valid number of participants in original project (LLH and HITT).

Bootstrapping methods were used to produce confidence intervals (CIs) for the WHS for the LLH and HITT projects.[30] For each project, 1000 samples without replacement were taken, the weighted (average) health status calculated for each and the 95% CI calculated using the 2.5th and 97.5th percentiles.[30] An independent *t*-test was then used to compare the mean weighted average scores in the two surveys using the bootstrapped samples.[30,31]

Ethics

Approval to conduct surveys was given by the Clinical Director of Gorki District Hospital and the Principals of both Universities. All questionnaires were treated anonymously.

Results

A total of 598 questionnaires were received of which 570 were fully completed, valid and used in the analysis. Twenty-eight questionnaires were excluded (20 were invalid and eight respondents rated thresholds between categories), giving a total of 275 (48.2%) LLH-related and 295 (51.8%) of HITT-related questionnaires (table 1).

Characteristics of the study participants

Overall, 85 out of 570 questionnaires (14.9%) were completed in Gorki, 109 (19.1%) in Mogilev and 376 (66.0%) in Pinsk (table 1). Participants in the HCRS survey were relatively young with mean age of 23.3 years [SD 10.7, range (17–82)]. This was similar for those who received our LLH-related [mean age 23.1 years, SD 10.0, range (17–77)] and HITT-related [mean age 23.5 years, SD 11.4, range (7–82)] questionnaires. For further statistical analysis, the study participants were categorized into two groups: <25 years and ≥25 years (table 1). Overall, more than half of the respondents in our survey, 349 out of 570 (61.2%) had completed secondary education, 105 out of 570 (18.4%) had incomplete higher, 61 out of 570 (10.7%) had completed higher and 38 out of 570 (6.7%) had completed secondary special education (table 1). Only a small proportion of the respondents (3.0%, 17 out of 570) had incomplete secondary education (table 1). The socio-demographic characteristics of the study participants that received LLH-related and HITT-related questionnaires (Part I) were similar according to the place of survey, age, gender and education (table 1).

Among 275 respondents who completed our LLH-related questionnaire, 73 (26.5%) reported their own health status as 'good', 171 (62.2%) as 'quite good', 29 (10.5%) as 'rather bad' and two (0.7%) reported 'bad' health status. Among 295 participants who received our HITT-related questionnaire, 17 (5.8%) reported their own health status as 'very good', 156 (52.9%) as 'good', 114 (38.6%) as 'fair', eight (2.7%) as 'poor' and nobody reported 'very poor' health status. Mean self-reported health status of those who completed the LLH-related and HITT-related questionnaires (Part II) was similar according to the VAS self-measurements (table 1).

Based on the VAS measurements of the Russian categories from the LLH and HITT projects made by participants during our HCRS survey (Part III), mean, SD and range were calculated for each health response category separately for the LLH-related (4-point) and the HITT-related (5-point) scales.

VAS ratings for the LLH-related HCRS survey

The mean rating for the 'good' health response category on the LLH-related VAS was 90.0; for 'quite good', 69.3; for 'rather bad', 37.0 and for 'bad', 13.0 (table 2). A figure demonstrating measurements of the each response category from the LLH and HITT projects is available in the Supplementary Appendix (Supplementary figure S3).

There were no statistically significant differences in mean rating for any of the health response categories (LLH) by gender or age groups (table 2). There was some evidence of a statistically significant difference in the mean score of the LLH 'good', 'quite good' and 'rather bad' health response categories by place of survey; however, the actual difference in mean scores between the groups was small (table 2).

VAS ratings for the HITT-related HCRS survey

The 'very good' health response category on the VAS (HITT-related) was rated with a mean of 95.7; 'good', 77.9; 'fair', 54.6; 'poor', 29.2 and 'very poor', 9.9 (table 3).

There was a statistically significant difference in the mean ratings for 'very poor' health response category between males and females, and for 'poor' and 'very poor' between the two age groups (>25 years and ≥25 years) (table 3). A one-way analysis of variance showed a statistically significant difference in the mean score for 'good' from the HITT health response category, but again the size of differences in mean scores between the groups was small (table 3).

Weighted health status

In the original LLH project ($n = 2000$), 87 participants (4.4%) have reported that they 'don't know' when were asked about their health status. These participants were excluded from the calculation of the weighted health status, and only the participants with valid responses ($n = 1913$) were included (table 4). In the original HITT project, all ($n = 1800$) participants reported their health status (table 4). The mean rating of each response category (from the LLH and HITT projects) based on VAS measurements from our HCRS survey was used to calculate the weighted health status (WHS) in Belarus separately for the LLH (2001) and HITT (2010) projects (table 4).

A small increase of 2.9 points on a 0–100 scale for the WHS in Belarus was found (56.2 vs. 59.1) between 2001 and 2010 (table 4). The 95% CIs were calculated for the WHS in the LLH and the HITT surveys using bootstrapping methods[30,31] (table 4). The *t*-test comparing the weighted average health status for the LLH and HITT surveys was statistically significant ($P < 0.001$).

Table 1 Characteristics of the study participants by type of questionnaire, HCRS-2010[a] survey

Characteristics	Type of questionnaire	
Socio-demographic	LLH-related $n = 275$ n (%)	HITT-related $n = 295$ n (%)
Place		
Gorki	33 (12.0)	52 (17.6)
Mogilev	54 (19.6)	55 (18.6)
Pinsk	188 (68.4)	188 (63.7)
Age		
<25 years	224 (81.5)	239 (81.0)
≥25 years	51 (18.5)	56 (19.0)
Gender		
Female	195 (70.9)	213 (72.2)
Male	80 (29.1)	82 (27.8)
Education		
Incomplete secondary	4 (1.5)	13 (4.4)
Completed secondary	170 (61.8)	179 (60.7)
Completed secondary special	18 (6.5)	20 (6.8)
Incomplete higher	54 (19.6)	51 (17.3)
Completed higher	29 (10.5)	32 (10.8)
Self-reported health status	Mean (SD) (range)	Mean (SD) (range)
0–100 VAS measurements	71.9 (15.8) (20.0–100.0)	71.7 (14.8) (20.0–100.0)

a: HCRS-2010: our Health Category Response Scale survey conducted in Belarus, 2010 ($n = 570$).

Table 2 The LLH-related (four-point) HCRS measurements (Mean, SD, range) on VAS (0–100) by gender, age and place of survey ($n = 275$)

LLH-related	VAS (0-100) measurements, Mean (SD) (range: minimum–maximum)							Total $n = 275$
	Gender		Age groups		Place of survey			
	Male $n = 80$	Female $n = 195$	<25 $n = 224$	≥25 $n = 51$	Gorki $n = 33$	Mogilev $n = 54$	Pinsk $n = 188$	
Good	89.3 (7.8) (65–100)	90.4 (8.2) (55–100)	90.5 (7.3) (65–100)	87.9 (10.7) (55–100)	84.6 (12.9) (55–100)*	92.1 (6.2) (70–100)*	90.4 (7.0) (55–100)*	90.0 (8.1) (55–100)
Quite good	68.6 (9.2) (50–90)	69.5 (9.4) (50–90)	69.2 (9.1) (50–90)	69.4 (10.4) (50–90)	67.4 (11.0) (50–90)*	73.2 (8.1) (50–90)*	68.5 (9.10) (50–90)*	69.3 (9.5) (50–90)
Rather bad	35.5 (10.9) (5–60)	37.6 (9.0) (10–60)	37.4 (8.7) (5–60)	35.3 (13.0) (5–60)	31.7 (13.5) (5–50)*	39.4 (10.0) (0–0)	37.3 (8.3) (5–60)*	37.0 (9.6) (5–60)
Bad	13.1 (8.4) (0–50)	12.9 (7.8) (0–35)	12.5 (7.6) (0–40)	14.8 (9.4) (0–50)	14.1 (9.1) (0–35)	14.1 (9.1) (0–50)	12.4 (7.4) (0–40)	13.0 (8.0) (0–50)

*Statistically significant, $P < 0.05$ (One-way analysis of variance test).

Table 3 The HITT-related (five-point) HCRS measurements (Mean, SD, range) on VAS (0–100) by gender, age and place of survey ($n = 295$)

HITT-related	VAS (0–100) measurements, Mean (SD) (range: minimum–maximum)							Total $n = 295$
	Gender		Age groups		Place of survey			
	Male $n = 82$	Female $n = 213$	<25 $n = 239$	≥25 $n = 56$	Gorki $n = 52$	Mogilev $n = 55$	Pinsk $n = 188$	
Very good	94.9 (4.9) (80–100)	96.0 (5.3) (65–100)	95.7 (5.0) (75–100)	95.6 (6.2) (65–100)	94.9 (6.7) (65–100)	95.8 (5.4) (75–100)	95.9 (4.7) (75–100)	95.7 (5.2) (65–100)
Good	77.6 (8.2) (60–95)	78.0 (7.6) (60–95)	77.5 (7.7) (60–95)	79.6 (7.8) (60–95)	79.9 (8.7) (60–95)*	76.1 (7.6) (60–90)*	77.9 (7.6) (60–95)*	77.9 (7.8) (60–95)
Fair	54.6 (9.3) (25–80)	54.6 (8.6) (30–80)	54.2 (8.4) (25–80)	56.3 (10.0) (30–80)	56.8 (11.5) (30–80)	53.3 (7.3) (40–80)	54.4 (8.2) (25–80)	54.6 (8.8) (25–80)
Poor	29.9 (9.0) (5–60)	28.9 (9.0) (5–50)	28.6 (8.6) (5–50)**	31.5 (10.4) (5–60)**	28.6 (13.7) (5–60)	29.9 (6.8) (20–50)	29.1 (7.9) (5–50)	29.2 (9.0) (5–60)
Very poor	11.6 (8.6) (0–50)**	9.2 (6.9) (0–30)**	9.2 (6.6) (0–40)**	12.9 (10.0) (0–50)**	11.6 (10.4) (0–50)	10.9 (6.8) (0–30)	9.1 (6.5) (0–40)	9.9 (7.5) (0–50)

*Statistically significant, $P < 0.05$ (One-way analysis of variance test).
**Statistically significant, $P < 0.05$ (Independent t-test).

Table 4 Weighted health status in Belarus in 2001 and 2010, based on the LLH-2001 ($n = 2000$), HITT-2010 ($n = 1800$) and HCRS-2010[a] ($n = 570$) surveys

4-point scale	LLH-2001 n (%)	HCRS-2010[a] VAS (0–100) Mean score	5-point scale	HITT-2010 n (%)	HCRS-2010[a] VAS (0–100) Mean score
Good[b]	312 (16.3)	90.0	Very good	56 (3.1)	95.7
Quite good	830 (43.4)	69.3	Good[b]	566 (31.4)	77.9
			Fair	911 (50.6)	54.6
Rather bad	487 (25.0)	37.0	Poor[c]	236 (13.1)	29.2
Bad[c]	293 (15.3)	13.0	Very poor	31 (1.7)	9.9
Valid	1931 (100.0)		Valid	1800 (100.0)	
Weighted health status-2001:		56.2	Weighted health status-2010:		59.1
95% CI*		54.8–57.1**	95% CI*		58.2–59.9**

a: HCRS-2010: our Health Category Response Scale survey conducted in Belarus, 2010 ($n = 570$).
b: 'Good' is equivalent of 'Хорошее' (Rus.).
c: 'Bad' and 'Poor' both are equivalent of 'Плохое' (Rus.).
*Using bootstrapping method.
**$P < 0.001$ (Independent t-test).

Discussion

Although both EU projects (LLH-2001 and HITT-2010) asked about self-reported health status, they used questions with slightly different wording and different verbal response categories.[9,10,21,22] We designed our study to make possible a direct comparison of health status in Belarus between 2001 (the LLH) and 2010 (the HITT). As with both projects, our survey was conducted in Belarus in Russian, and different areas of Belarus were represented in this study. The study design, size and sample were chosen because of their advantage of practicality.[28,32]

A limitation of our study when applied to the LLH and HITT projects is that we did not exactly replicate the demographics of the original studies.[21–27] There were differences in age, gender and education status between the LLH and HITT, which were population surveys, and the HCRS survey, which used a convenience sample of mainly university students. The purpose of our HCRS survey was to use a group of Russian speakers in Belarus to compare the strength of the Russian response categories used in the two original population surveys. Although, there may be differences between old and young people in how they would understand and rate common Russian words and phrases as 'very good', we do not believe these differences would be large. Furthermore, even if age differences in rating do exist in understanding and interpreting various self-reported health outcome measures,[33] we are mainly interested in the differences between the LLH and HITT categories. As the HCRS survey questionnaires were allocated using a pseudo-random method, the characteristics of those answering the two HCRS questionnaires were similar (table 1), and therefore our results should reflect differences in how these two sets of categories are rated. For this reason, we thought it unnecessary that the HCRS should be a population survey and instead placed the emphasis on obtaining a large number of ratings using a convenience sample of Russian speakers in Belarus. This approach was highly successful, and we were able to include results from >500 Russian speakers in three Belarusian cities (table 1). Despite this, it would be very interesting to adjust the results by age category using a stratified analysis. In practice, however, there are only a small number of people aged >25 years in the HCRS, and therefore such an approach would not be practical in our study, as the results for those aged >25 years would not be very accurate.

Other advantages of our HCRS survey were the ability to (i) test questions and wording from both original EU projects simultaneously; and (ii) prevent response bias, as one person would receive only one randomly allocated version of the questionnaire (the LLH- or HITT-related).

We designed a VAS 100ths graduated ruler for both the 4-point and the 5-point response categories of the original LLH and HITT projects. VASs with various layout (vertical or horizontal), different scale rating and response category scales have been applied and validated in various international studies.[34–38] We are not aware of any literature with a similar approach to that used in our study. We chose a vertical 100ths graduated ruler (VAS) to measure the strength of each response category used in the original LLH and HITT projects. First, we asked participants to rate their own health using both a VRS and a VAS, before our primary question concerning the wording of the Russian response categories from the LLH and HITT projects. This was to introduce the idea of grading and help them to understand the context of the main task.

It is interesting that the same health response categories (Russian words) were rated very differently on the LLH- and HITT-related VASs. The category 'good' ('Хорошее') measured ~12 points higher, and the category 'bad/poor' ('Плохое') measured ~16 points lower, on the 4-point compared with the 5-point VAS (tables 2 and 3). This suggests that respondents may place more emphasis on the relative position of categories and whether they are the most extreme choice, and less emphasis on the actual wording of the category.

We then calculated the WHS for both the LLH (2001) and the HITT (2010) projects to compare the health status between the two time points (table 4). The average health status in Belarus showed a small, but statistically significant, improvement of 2.9 points on a 0–100 scale between 2001 and 2010 (56.2 vs. 59.1) that may also be an important change.[39,40]

We successfully used our HCRS survey to standardize the results from the original EU-funded projects with different response categories. The importance of our study is that it demonstrates that a survey method based on a VAS can be used to measure and compare health status between various population surveys with different response categories. A similar approach could be used in other research studies to compare scales with different response categories, in longitudinal studies and perhaps to compare studies in different languages.

Conclusion

We developed a method for comparing 4-point and 5-point health response categories. When applied to the LLH (2001) and HITT (2010) projects, a direct comparison showed a perceived improvement in health status. This type of survey method based on a VAS may be useful in other quantitative studies where a comparison needs to be made between similar questions with different response category scales. Also, our HCRS survey has showed an important secondary finding of the impact of the number of response categories in population-based studies when, whatever the wording, individuals may subconsciously assume that categories are equally spaced.

Supplementary Data

Supplementary data are available at *EURPUB* online.

Acknowledgements

Funding

Key points

- The use of this novel method, the HCRS survey, successfully enabled a direct comparison between two EU-funded projects (the LLH-2001 and HITT-2010) that had measured self-reported health status differently.
- Weighted health status was calculated for Belarusian respondents for the 2001 (LLH) and 2010 (HITT) surveys. There was a difference of 2.9 in the mean score on a 0–100 graduated VAS between the two projects showing a small, but statistically significant, improvement in self-reported health status.

- Individuals rated the same health response categories very differently depending on the overall number of categories available, showing the impact of the number of response categories in population-based studies.
- This methodological approach may be useful for comparing health status with other countries that participated in the original EU projects, and can be applied to other research with different response scales or different languages and longitudinal studies.

References

1 Bowling A. *Measuring Health. A Review of Quality of Life Measurement Scales*, 3rd edn. Berkshire, England: Open University Press, 2005.

2 Rumsfeld JS. Health status and clinical practice. When will they meet? *Circulation* 2002;106:5–7.

3 Abbot PA, Turmov S, Wallace C. Health world views of post-Soviet citizens. *Soc Sci Med* 2006;62:228–38.

4 Sturgis P, Thomas R, Purdon S, et al. Comparative Review and Assessment of Key Health State Measures of the General Population. Report. London: Department of Health, 2001. Available at: http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsStatistics/DH_4008058 (02 August 2010, date last accessed).

5 Shmueli A. Socio-economic and demographic variation in health and its measures: the issue of reporting heterogeneity. *Soc Sci Med* 2003;57:125–34.

6 Shmueli A, Messika D, Murad H, Freedman L. Does greater exposure to own-health data make a difference on the visual analog scale? *Eur J Health Econ* 2008;9:63–7.

7 Badia X, Schiaffino A, Alonso J, Herdman M. Using the EuroQol 5-D in the Catalan general population: feasibility anvd construct validity. *Qual Life Res* 1998;7:311–22.

8 Stein K, Fry A, Round A, et al. What value health? A review of health state values used in early technology assessments for NICE. *Applied Health Econ Health Policy* 2005;4:219–28.

9 The *"Living conditions, Lifestyle and Health"* Project (by the European Community under the FP5 horizontal programme "Confirming the International Role of Community Research"; INCO2-Copernicus; Contract No: ICA2-2000–10031, Project No: ICA2-1999–10074), London, 2000.

10 Health in Times of Transition (HITT-CIS) Project. *Population Health and Health Policies in CIS Countries*. An EP7-HEALTH collaborative research project (2009-2013). Available at: http://www.hitt-cis.net/ (13 May 2010, date last accessed).

11 Bobak M, Pikhart H, Hertzman C, et al. Socioeconomic factors, perceiving control and self-reported health in Russia. A cross-sectional survey. *Soc Sci Med* 1998;47:269–79.

12 König HH, Bernet S, Angermeyer MC. Health status of the German population: results of a representative survey using EuroQol questionnaire. *Gesundheitswesen* 2005;67:173–82.

13 Kimura T, Hayashida K, Araki Y, et al. How valid is the self-administered visual analogue scale for assessing health status of elderly people? *Hiroshima J Med Sci* 2008;57:85–92.

14 Ohnhaus EE, Adler R. Methodological problems in the measurement of pain: a comparison between the verbal rating scale and the visual analogue scale. *Pain* 1975;1:379–84.

15 DeLoach LJ, Higgins MS, Caplan AB, Stiff JL. The Visual analog scale in the immediate postoperative period: intrasubject variability and correlation with a numeric scale. *Anesth Analg* 1998;86:102–6.

16 Filocamo G, Davi S, Pistorio A, et al. Evaluation of 21-numbered circle and 10-centimeter horizontal line visual analog scales for physician and parent subjective ratings in juvenile idiopathic arthritis. *J Rheumatol* 2010;37:1534–41.

17 Kievit W, Welsing PM, Adang EM, et al. Comments on the use of self-reporting instruments to assess patients with rheumatoid arthritis: the longitudinal association between the DAS28 and the VAS general health. *Arthritis Rheum-Arthritis Care Res* 2006;55:745–50.

18 Jenkinson C. *Measuring Health and Medical Outcomes*. London: UCL Press Ltd., 1994.

19 Lansing RW, Moosavi SH, Banzett RB. Measurement of dyspnea: word labelled visual analog scale vs. verbal ordinal scale. *Respir Physiol Neurobiol* 2003;134:77–83.

20 Guyatt GH, Townsend M, Berman LB, Keller JL. A comparison of Likert and Visual Analog Scales for measuring changes in function. *J Chronic Dis* 1987;40:1129–33.

21 The Living Conditions, Lifestyle and Health (LLH) Project. About Living Conditions, Lifestyle and Health. Available at: http://www.llh.at/llh_project_about.html (21 October 2010, date last accessed).

22 Health in Time of Transition Project. HITT-CIS. Press and Downloads. Press Release: New study on trends in population health in neighbouring EU countries. Available at: http://www.hitt-cis.net/press-and-downloads (13 May 2010, date last accessed).

23 Cockerham WC, Hinote BP, Cockerham GB, Abbott P. Health lifestyles and political ideology in Belarus, Russia and Ukraine. *Soc Sic Med* 2006;62:1799–809.

24 Abbott P. Cultural trauma and social quality in post soviet Moldova and Belarus. *East Eur Politics Soc* 2007;21:219–58.

25 Cockerham WC, Hinote BP, Abbott P. Psychological distress, gender, and health lifestyle in Belarus, Kazakhstan, Russia, and Ukraine. *Soc Sic Med* 2006;63:2381–94.

26 Hinote BP, Cockerham WC, Abbott P. Post-communism and female tobacco consumption in the former Soviet states. *Europe-Asia Studies* 2009;61:1543–55.

27 Balabanova D, Roberts B, Richardson E, et al. Health care reforms in the former Soviet Union: beyond the transition. *Health Serv Res* 2012;47:840–64.

28 Argyrous G. *Statistics for Social & Health Research*. London: Sage Publications, 2000.

29 Petrie A, Sabin C. *Medical Statistics at a Glance*, 2nd edn. London: Blackwell Publishing, 2005.

30 Efron B, Tibshirani R. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statist Sci* 1986;1:54–77.

31 Walters SJ, Campbell MJ. The use of bootstrap methods for analysing Health-Related Quality of Life outcomes (particularly the SF-36). *Health Qual Life Outcomes* 2004;2:1–19.

32 Bowling A. Research Methods in Health. *Investigating Health and Health Services*, 2nd edn. Berkshire: Open University Press, 2004.

33 Jagsch R, Pils K. Which instrument is more suitable to assess health-related quality of life: Nottingham health profile or short-form-36? *Wien Med Wochenschr* 2006;156:149–57.

34 Knutsson I, Rydström H, Reimer J, et al. Interpretation of response categories in patient-reported rating scales: a controlled study among people with Parkinson's disease. *Health Qual Life Outcomes* 2010;8:61–9.

35 Williams VS, Morlock RJ, Feltner D. Psychometric evaluation of a visual analog scale for the assessment of anxiety. *Health Qual Life Outcomes* 2010;8:57–64.

36 Carlsson AM. Assessment of chronic pain. I. Aspects of the reliability and validity of the visual analogue scale. *Pain* 1983;16:87–101.

37 Insinga RP, Fryback DG. Understanding differences between self-ratings and population ratings for health in the EuroQol. *Qual Life Res* 2003;12:611–9.

38 Shmueli A. The relationship between the visual analog scale and the SF-36 scales in the general population: an update. *Med Decis Making* 2008;28:66–89.

39 Kelly AM. Does the clinical significant difference in visual analog scale pain scores vary with gender, age, or cause of pain? *Acad Emerg Med* 1998;5:1086–90.

40 Zisapel N, Nir T. Determination of the minimal clinically significant difference on a patient visual analog sleep quality scale. *J Sleep Res* 2003;12:291–8.