Comparison of the three-level and the five-level versions of the EQ-5D

Christiansen, Anne Simone Juhl; Møller, Marie Louise Sletskov; Kronborg, Christian; Haugan, Ketil Jørgen; Køber, Lars; Højberg, Søren; Brandes, Axel; Graff, Claus; Diederichsen, Søren Zöga; Nielsen, Jonas Bille; Krieger, Derk; Holst, Anders Gaarsdal; Svendsen, Jesper Hastrup

Published in:
The European Journal of Health Economics

DOI (link to publication from Publisher):
10.1007/s10198-021-01279-z

Publication date:
2021

Document Version
Accepted author manuscript, peer reviewed version

Comparison of the three-level and the five-level versions of the EQ-5D

Authors and affiliations

Anne Simone Juhl Christiansen, MSc[a], Marie Louise Sletskov Møller, MSc[a], Christian Kronborg, PhD[a], Ketil Jørgen Haugan, MD, PhD[b], Lars Køber, MD, DMSc[c,d], Søren Højberg, MD, PhD[e], Axel Brandes, MD, DMSc[f,g], Claus Graff, PhD[h], Søren Zöga Diederichsen, MD[c], Jonas Bille Nielsen, MD, PhD[i,j,k], Derk Krieger, MD, PhD[l,m], Anders Gaarsdal Holst, MD, PhD[i], Jesper Hastrup Svendsen, MD, DMSc[c,d,i].

[a] Department of Business and Economics, University of Southern Denmark
Address: Campusvej 55, 5230 Odense, DENMARK
[b] Department of Cardiology, Zealand University Hospital Roskilde
Address: Sygehusvej 10, 4000 Roskilde DENMARK
[c] Department of Cardiology, Rigshospitalet, Copenhagen University Hospital
Address: Inge Lehmanns Vej 7, 2100 Copenhagen, DENMARK
[d] Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen
Address: Blegdamsvej 3B, 2200 Copenhagen, DENMARK
[e] Department of Cardiology, Bispebjerg Hospital, Copenhagen University Hospital
Address: Bispebjerg Bakke 23, 2400 Copenhagen DENMARK
[f] Department of Cardiology, Odense University Hospital
Address: J. B. Winsløws Vej 4, 5000 Odense, DENMARK
[g] Department of Clinical Research, Faculty of Health Sciences, University of Southern Denmark
Address: Winsløwparken 19, 5000 Odense C, DENMARK
[h] Department of Health Science and Technology, Aalborg University
Address: Fredrik Bajers Vej 7 D2, 9220 Aalborg, DENMARK
[i] Laboratory for Molecular Cardiology, Department of Cardiology, Rigshospitalet, Copenhagen University Hospital
Address: Inge Lehmanns Vej 7, 2100 Copenhagen, DENMARK
[j] Department of Epidemiology Research, Statens Serum Institut
Address: Artillerivej 5, 2300 Copenhagen, DENMARK
[k] K.G. Jebsen Center for Genetic Epidemiology, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Norway
[l] University Hospital Zurich, University of Zurich
Address: Rämistrasse 100, 8091 Zürich, SWITZERLAND
[m] Stroke Unit, Mediclinic City Hospital
Address: Building 37 - 26th St, Dubai, UAE

Corresponding author:
Christian Kronborg, PhD
Department of Business and Economics, University of Southern Denmark
Address: Campusvej 55, 5230 Odense, DENMARK
Mail: cka@sam.sdu.dk

Title page

Comparison of the three-level and the five-level versions of the EQ-5D

ORCID of authors

Marie Louise Sletskov Møller	0000-0002-3800-8118
Christian Kronborg	0000-0001-9627-1419
Ketil Jørgen Haugan	0000-0001-5584-8445
Søren Højberg	0000-0001-8385-587X
Axel Brandes	0000-0001-9145-6887
Claus Graff	0000-0001-9650-9781
Søren Zöga Diederichsen	0000-0002-9687-0857
Jonas Bille Nielsen	0000-0002-6654-2852
Derk Krieger	0000-0003-2457-7333
Anders Gaarsdal Holst	0000-0002-5343-7367
Jesper Svendsen	0000-0001-8466-8515

Title: Comparison of the three-level and the five-level versions of the EQ-5D

Abstract

EQ-5D is a generic instrument to measure health-related quality of life. In 2009, a new version, EQ-5D-5L, was introduced as an attempt to reduce ceiling effects and improve sensitivity to small changes over time.

The objective of this study was to assess the measurement properties of the EQ-5D-5L instrument compared to the EQ-5D-3L instrument in an elderly general population with a moderate to high degree of comorbidity.

A subgroup of participants in a large clinical trial completed the EQ-5D-3L and the EQ-5D-5L questionnaires. Based on the collected data, we tested for feasibility and ceiling and floor effects. Furthermore, we assessed the redistribution properties of the responses and examined the level of inconsistency, informativity., and convergent validity.

A total of 1002 persons diagnosed with hypertension, diabetes, heart failure and/or previous stroke completed both the EQ-5D-3L and the EQ-5D-5L questionnaires. The overall ceiling effect decreased from 46% with the EQ-5D-3L to 30% with the EQ-5D-5L and absolute and relative informativity were higher for EQ-5D-5L, and there was a stronger correlation between EQ-5D-5L and EQ VAS.

The EQ-5D-5L seemed to perform better than the EQ-5D-3L in terms of feasibility, ceiling effect, discriminatory power and convergent validity. The overall ceiling effect was higher than that found in patient samples in previous studies but lower than the one found in population studies.

Keywords

EQ-5D-5L, EQ-5D-3L, Health-related quality of life, Measurement properties.

Introduction

The EQ-5D questionnaire is a generic instrument to measure health-related quality of life. It contains a descriptive system that includes five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression [1]. These five dimensions are considered aspects of quality-of-life that can be affected by illness and health interventions. A key feature is the translation of the respondents' answers in the five dimensions into a single index score, the EQ-5D index score, which is obtained with a formula that attaches weights to each level in the five dimensions based on valuations by the general population [2,3]. The index score is anchored at 0 for a health state equivalent to being dead and at 1 for full health, while index scores of less than zero reflect health states regarded as worse than being dead [3]. The index score is used to calculate quality-adjusted life years in, for example, economic evaluations of health care interventions. The EQ-5D questionnaire is frequently used when assessing health in observational studies and clinical trials, and in particular in economic evaluations.

There are two versions of the EQ-5D instrument. The original three-level version known as EQ-5D-3L (or 3L) has three response levels in each dimension of the descriptive system (no problems, some problems, extreme problems/unable to/confined to bed). The new five-level version, EQ-5D-5L (or 5L), has five response levels in each dimension (no problems, slight problems, moderate problems, severe problems, extreme problems/unable to). As part of the increase from three to five response levels in the 5L version, the wording of the levels has been standardized. On the first dimension, the worst level of mobility was changed to "unable to walk about" instead of "confined to bed". In addition, the intermediate response level on the first three dimensions was changed to "moderate" instead of "some" [4].

The 3L version has been found to be reliable and valid [5-7]. However, the 5L version, introduced in 2009, was developed to reduce the occurrence of respondents reporting the best possible health (ceiling effect), and to improve the sensitivity of the EQ-5D instrument to changes in an individual's health over time [8,4].

In a recent systematic review, Buchholz et al. [9] found the measurement properties of the EQ-5D-5L to be similar or better than those of the EQ-5D-3L in terms of informativity, distributional properties and ceiling effect in different populations and settings. However, at present most studies have focused on specific patient populations as reported in the systematic review by Buchholz et al. [9] while fewer studies have focused on general samples.

The aim of this study was to compare the feasibility, ceiling and floor effects, redistribution properties, discriminatory power, and convergent validity of the EQ-5D-5L and EQ-5D-3L in an elderly general population with a moderate to high degree of comorbidity.

Methods

Study design and participants

This study is a sub-study of the LOOP study, which is a large multi-centre randomized clinical trial investigating whether continuous, long-term monitoring of heart rhythm with an implantable loop

recorder can improve atrial fibrillation detection and stroke prevention [10]. The inclusion of participants in the study has been terminated, whereas follow-up is still ongoing.

Details of the design of the LOOP study have previously been published [10]. In brief, persons aged ≥70 years and previously diagnosed with one or more of the following conditions: hypertension, diabetes mellitus, heart failure or stroke, were identified through national registries (the Danish National Patient Registry and the Danish Civil Registration System) and invited by letter. Eligible persons were randomized to either receive an implantable loop recorder or not (control group).

Before the randomization, all participants were seen in one of the four clinics that perform the trial. Prior to the visit, each participant received a letter with information about the study and the EQ-5D-5L questionnaire to fill out at home before the visit. A subset of the participants was presented with the EQ-5D-3L version during their visit. This subset was defined as the first 1000 participants who visited the clinics at either Rigshospitalet-Copenhagen University Hospital, Bispebjerg Hospital-Copenhagen University Hospital, or Zealand University Hospital in Roskilde. Once it was registered that at least 1000 participants had filled out an EQ-5D-3L questionnaire at their first visit, further collection of EQ-5D-3L data was terminated.

The sample size was decided on two premises. First, as a sub-study it was important that the validity of the main objective of the LOOP study was not jeopardised. We sought to ensure this by keeping the proportion of participants included in the sub-study low relative to number of participants in the main study. Second, the sample should be sufficiently large to be representative of the participants in the study. With a sample size of 1000 persons, the proportion of participants included in this sub-study was less than 1/6 of the participants in the main study. In previous studies that compared the two instruments sample sizes varies considerably, from less than 100 to more than 7000 participants [9].

The participants gave written informed consent. All data were stored in a centralized online case report file system. The LOOP study was approved by the Ethics Committee of the Capital Region of Denmark (H-4-2013-025) and the Danish Data Protection Agency (2007-58-0015). The trial is registered at ClinicalTrials.gov (NCT02036450).

The EQ-5D questionnaire

The EQ-5D questionnaire was developed to measure self-reported health in a simple and generic way. It contains a descriptive system that includes the five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression, and a visual analogue scale (the EQ VAS). Both the descriptive system and the EQ VAS measure the current health of the respondent. Combining the respondents' answers in each of the five dimensions of the descriptive system results in a unique health state (for example, health state 11111, which indicates no problems in any of the five dimensions of the descriptive system, or 33333 or 55555, which indicates extreme problems in all five dimensions of the EQ-5D-3L and EQ-5D-5L descriptive system, respectively). By using preference-based value sets it is possible to transform these health states into the EQ-5D index score.

On the EQ VAS the respondents rate their current health on a scale ranging from 0 to 100. Zero represents "the worst health you can imagine" and 100 represents "the best health you can imagine" [11].

We used the paper versions of the officially Danish translations of the two instruments. The study was registered with the EuroQol Research Foundation.

Analysis

Feasibility

To assess feasibility, we calculated the proportion of participants leaving the EQ-5D-3L and EQ-5D-5L questionnaires completely blank on all dimensions of the descriptive system. The proportion of participants leaving one to four of the dimensions blank was also calculated.

Ceiling and floor effects

We evaluated ceiling effects by calculating the number of participants who reported "no problems" in any dimension and on all dimensions of the descriptive systems. Floor effects were assessed by calculating the number of persons reporting "extreme problems/unable to/confined to bed" in the EQ-5D-3L descriptive system and the number of persons reporting "extreme problems/unable to" in the EQ-5D-5L descriptive system in any dimension and on all dimensions.

Absolute and relative reductions of ceiling and floor effects, when going from 3L to 5L, were also calculated. We used the following formula to compute the relative reduction:

$$\frac{ceiling/floor_{3L} - ceiling/floor_{5L}}{ceiling/floor_{3L}} * 100$$

Redistribution properties

The redistribution properties were evaluated by assessing the proportion of consistent and inconsistent 3L-5L response pairs in the descriptive system. In any dimension, a response pair was considered inconsistent if an EQ-5D-5L response of an individual was two or more levels away from the EQ-5D-3L response of the individual, as in Janssen et al. [12] (for example, if a person answered "no problems" (level 1) in EQ-5D-3L and "moderate problems" (level 3) in EQ-5D-5L).

We calculated the overall proportion of inconsistent 3L-5L response pairs and the proportion of consistent and inconsistent response pairs by each dimension. We used cross-tabulation of EQ-5D-3L and EQ-5D-5L dimension scores to assess the size of inconsistency (i.e., by how many levels an inconsistent 3L-5L response pair differed).

Discriminatory power (informativity)

The discriminatory power (informativity) of EQ-5D-3L and EQ-5D-5L was evaluated by calculating the Shannon index [13] and the Shannon evenness index for each dimension of the descriptive system. The Shannon index (H') expresses the absolute amount of informativity captured in a system, in this case the absolute amount of information obtained in a given dimension of the EQ-5D descriptive system. The Shannon evenness index (J') expresses the relative amount of information obtained.

We calculated the Shannon index with the formula:

$$H' = -\sum_{i=1}^{L} p_i \log_2 p_i$$

where L indicates the number of response levels in a dimension, and p_i is the proportion of observations at the *i*th level ($i = 1, \ldots, L$). The higher the value of H', the more information is obtained. The optimal amount of information is captured if the distribution of responses is rectangular (i.e., if responses of respondents are evenly distributed across all response levels). At this point, H' reaches its maximum value (H'_{max}), which corresponds to $\log_2 L$ (1.58 for EQ-5D-3L and 2.32 for EQ-5D-5L). H'_{max} increases if the number of response levels increases. However, the discriminatory power only increases if extra added levels are in fact used by respondents [12,14].

The Shannon evenness index (J') was calculated as:

$$J' = \frac{H'}{H'_{max}}$$

where $H'_{max} = \log_2 L$. A higher value of J' of the EQ-5D-5L compared to the J' of EQ-5D-3L questionnaire indicates better discriminatory power of EQ-5D-5L. A lower J' indicates a loss in the potential informativity [15].

Convergent validity

We tested for convergent validity between the EQ-5D descriptive system and the EQ VAS and between the EQ-5D index score and the EQ VAS for EQ-5D-3L and EQ-5D-5L, respectively, by performing Spearman's rank order correlation. The strength of the correlations was interpreted as: very weak (<0.20), weak (0.20-0.39), moderate (0.40-0.59), strong (0.60-0.79) and very strong (0.80-1) [16].

Results

A total of 459 females and 543 males completed both the 3L and 5L versions of the EQ-5D questionnaire between February 2014 and May 2016. The participants were between 70 and 90.5 years old with a mean age 75.9 (std.dev. 3.94) years.

All participants had one or more of the following conditions: Hypertension (89%), diabetes (27%), heart failure (5%) or previous stroke (19%). 647 participants (65%) suffered from one of these conditions only, 316 participants (32%) suffered from two of the conditions, and 39 participants suffered from three or more conditions.

All 1002 participants indicated an EQ VAS score on the 5L version, whereas only 979 persons indicated it on the 3L version. The mean EQ VAS scores on the two versions differed between the two instruments. On the 5L version the mean was 80.2 (std.dev. 15.51), and on the 3L version it was 78.7 (std.dev. 16.29; $p < 0.001$). In both versions the EQ VAS score ranged from 15 to 100, and with median values of 80 and 84 for the 5L and 3L versions, respectively.

Table 1 shows an analysis of the differences in the participants EQ VAS scores. In total, 615 (63%) out of 979 participants reported the same EQ VAS score for the two versions of the EQ-5D questionnaire whereas 248 (25%) reported a lower 5L EQ VAS than 3L EQ VAS, and 116 (12%) reported a higher 5L EQ VAS than the 3L EQ VAS. Most participants with a higher or lower 5L EQ VAS than the 3L EQ VAS had a difference between 1 and 5 points on the EQ VAS. The maximum difference between the two EQ VAS scores was 45

Feasibility

Out of a total of 1002 persons, seven (0.7%) left the 3L questionnaire completely blank on all dimensions of the descriptive system, whereas one person (0.1%) left the 5L questionnaire completely blank. No persons left both the 3L questionnaire and the 5L questionnaire completely blank, while 15 (1.5%) persons answered the 3L questionnaire partly and one (0.1%) person answered the 5L questionnaire partly (i.e., leaving one to four of the dimensions blank).

Ceiling and floor effects

Table 2 shows the number of persons who reported "no problems" in any dimension and on all dimensions (ceiling effect). It also shows the number of persons reporting "extreme problems/unable to" (floor effects).

Ceiling effects ranged from 57% to 96% for the 3L and 40% to 93% for the 5L. Comparing the 5L with the 3L, we found that the reduction in the ceiling effect in the individual dimensions was between 7 and 17 percentage points. The proportion of participants reporting no problems on all dimensions, "11111", decreased from 46% (95% confidence interval, CI, 42.4-48.6%) with the 3L questionnaire to 30% (95% CI 26.7-32.4%) with the 5L.

The floor effect (i.e. the proportion of persons reporting "extreme problems/unable to") was less than 1% in all dimensions except for pain/discomfort in the 3L questionnaire. There were no observations of "33333" or "55555" ("extreme problems/unable to" on all dimensions).

Redistribution properties and inconsistency

Out of the total of 1002 persons, between 381 (38%) and 919 (92%) persons reported "no problems" in one or more dimensions in both the 3L and 5L versions of the descriptive system (Table 3).

Of those answering "some/moderate problems" (level 2) with the 3L, between 38% (mobility) and 57% (self-care) answered 2, "slight problems", and between 3% (self-care) and 13% (mobility) answered 4, "severe problems", with the 5L version.

A few persons only answered 3 on any dimension of the 3L. Their answers to the 5L version were most frequently 4, "severe problems".

Overall, 160 out of 5,010 (i.e. 5 x 1002) response pairs on the five dimensions of the 3L and 5L were inconsistent. That is, the overall proportion of inconsistent responses was 3.2%. Table 3 shows the proportion of inconsistent response pairs by the dimensions of the EQ-5D descriptive system. The proportion of inconsistencies was highest for mobility and usual activities. On the mobility dimension, out of 758 persons answering "no problems" on the 3L questionnaire, 22 persons indicated "moderate" to "extreme problems" on the 5L questionnaire, and out of 193 persons answering" some problems" on 3L 21 persons answered "no problems" on the 5L questionnaire. That is a total of 43 (4.5%) inconsistent answers out of a total of 951 (=758+193) response pairs inconsistent responses on the mobility dimension. The proportion of inconsistent response pairs was also 4.3% on the usual activities dimension. The pain/discomfort dimension had 4.2% inconsistent response pairs, the anxiety/depression dimension had 2.8% and the self-care dimension had 0.5% inconsistent response pairs.

Table 3 shows that inconsistencies were more frequent for reports of level 2 (some/moderate problems) with the 3L version and levels 2-4 (slight to severe problems) with the 5L version.

Discriminatory power (informativity)

Table 4 shows the results of the discriminatory power (informativity) of the 3L and the 5L questionnaires. Both the absolute informativity (Shannon's indices, H') and the relative informativity (the evenness index, J') were higher for the 5L than the 3L on all dimensions.

Convergent validity

Table 5 shows the results of Spearman's rank order correlations for the convergent validity between EQ-5D and EQ VAS. Weak to moderate correlations were observed for both 3L and 5L. The correlations were somewhat stronger for 5L for all five dimensions of the EQ-5D descriptive system. 3L ranged from -0.215 (anxiety/depression) to -0.462 (pain/discomfort). 5L ranged from -0.282 (self-care) to -0.547 (usual activities). The correlation between the EQ-5D index scores and the EQ VAS was also slightly stronger for the 5L than the 3L (0.551 vs. 0.627) with the correlation being moderate for 3L and strong for 5L.

Discussion

In this study we investigated the measurement properties of the new 5-level version of the EQ-5D instrument compared to the original 3-level version in an elderly Danish population with a moderate to high degree of comorbidity. We found that the EQ-5D-5L had fewer missing data, fewer respondents reporting best possible health state (ceiling effects), and greater informativity of dimensions (discriminatory power) compared with the EQ-3D-3L.

The observed proportions of 3L and 5L questionnaires that were partially answered or left completely blank are within the range observed in previous studies. In a recent systematic review, Buchholz et al. [9] found that missing values on all five dimensions of the questionnaire are in the range 0-6.6% for the 3L and 0-4.0% for the 5L, and that the proportion of partially answered 3L and 5L questionnaires are in the ranges of 0-1.9% and 0-1.6%, respectively. The findings in the present study are at the lower end of these ranges.

The observed reduction in the ceiling effect in the individual dimensions of the descriptive system of EQ-5D is in accordance with findings in other studies, and the size of the absolute reduction is within the ranges observed in the systematic review by Buchholz et al. [9].

The observed ceiling effect for the profile (i.e., the proportion reporting "11111") is also comparable to the one observed in previous studies, as reported in Buchholz et al. [9]. However, the ceiling effect in the present study is higher than that found in patient samples but lower than the one in population samples. This may reflect the fact that the participants in the present study fulfilled the inclusion criteria of having at least one of four medical conditions (hypertension, diabetes, heart failure and/or previous stroke).

The results on the proportion of inconsistencies are within the range of those that have been found in previous studies. Buchholz et al. [9] report that the proportion of inconsistencies in healthy and chronic populations ranges from zero to 5.4%. In our study the overall proportion was 3.2%.

Limitations

Some limitations should be considered in the interpretation of the results of this study.

The participants received the 5L questionnaire to fill out at home before their inclusion visit and were presented with the 3L version during their inclusion visit. For this reason, the time interval between participants answering the 3L and 5L questionnaires may have varied among participants, because the partipants health may have changed in this time. Unfortunately, we do not have exact information about the time span.

Since the EQ-5D instrument asks about the respondents' health state '*today*', the observed inconsistencies may be a result of this limitation. However, since the respondents have been invited to the study based on known chronic conditions, we expect that these conditions are rather stable.

In addition, we found a statistically significant mean difference between the 3L EQ VAS, indicating a change in the overall health of the participants between the administration of the 3L and 5L versions, when judged by the respondents themselves. This may have contributed to some of the

observed differences between the 3L and the 5L versions of the EQ-5D questionnaire found in this study (e.g. explaining some of the reduced ceiling effect).

However, it should be noted that there are some differences between the 3L and 5L EQ VAS. The format of the EQ VAS, instructions and methods for marking a response on the EQ VAS, were changed in the 5L version with the aim of making the EQ VAS easier to complete and to score [17]. For this reason, the results of the 3L and 5L EQ VAS may not be completely comparable. Moreover, for most participants the differences in the EQ VAS scores were relatively small (between one and five points on the EQ VAS scale).

Strengths

Participants in this study were recruited specifically from a population with one to four common health conditions. Thus, it is not a specific patient population, but what we consider to be the common elderly population.

Another strength of the study is that participants in the study are identified from the Danish National Patient Registry, which includes standardized information about diagnoses from hospital admissions and out-patient visits; for example, diagnoses are classified according to the ICD-10, and information is continuously checked for errors and inconsistencies [18]. Thus, we consider the data on health conditions on which the participants are extracted for the study to be of high quality. Furthermore, by relying on register data on health conditions we avoid the problems that may arise from self-reported health conditions. However, the register may not cover the entire elderly population with these four conditions, since some individuals with these conditions can be treated in general practice clinics, which are not obliged to report to the Danish National Patient Registry.

Conclusions

The results from this study suggest that the EQ-5D-5L performs better that the EQ-5D-3L questionnaire in terms of feasibility, ceiling effects, discriminatory power and convergent validity. Furthermore, the overall ceiling effect was higher than that found in patient samples but lower than the one in population samples reported in previous studies.

References

1. Brooks, R.: EuroQol: the current state of play Health policy **37**(1), 53-72 (1996). doi:10.1016/0168-8510(96)00822-6
2. Poor, A.K., Rencz, F., Brodszky, V., Gulacsi, L., Beretzky, Z., Hidvegi, B., Hollo, P., Karpati, S., Pentek, M.: Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L in psoriasis patients. Quality of Life Research **26**(12), 3409-3419 (2017). doi:10.1007/s11136-017-1699-x
3. van Hout, B., Janssen, M.F., Feng, Y.S., Kohlmann, T., Busschbach, J., Golicki, D., Lloyd, A., Scalone, L., Kind, P., Pickard, A.S.: Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value in Health **15**(5), 708-715 (2012). doi:10.1016/j.jval.2012.02.008
4. Herdman, M., Gudex, C., Lloyd, A., Janssen, M.F., Kind, P., Parkin, D., Bonse, G., Badia, X.: Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Quality of Life Research **20**(10), 1727-1736 (2011). doi:10.1007/sl
5. Haywood, K.L., Garratt, A.M., Fitzpatrick, R.: Quality of life in older people: A structured review of generic self-assessed health instruments. Quality of Life Research **14**(7), 1651-1668 (2005).
6. Johnson, J.A., Pickard, A.S.: Comparison of the EQ-5D and SF-12 health surveys in a general population survey in Alberta, Canada. Medical Care **38**(1), 115-121 (2000). doi:10.1097/00005650-200001000-00013
7. Hunger, M., Sabariego, C., Stollenwerk, B., Cieza, A., Leidl, R.: Validity, reliability and responsiveness of the EQ-5D in german stroke patients undergoing rehabilitation. Quality of Life Research **21**(7), 1205-1216 (2012).
8. EuroQol Research Foundation: EQ-5D-5L | About. https://euroqol.org/eq-5d-instruments/eq-5d-5l-about/. Accessed 26. January 2020
9. Buchholz, I., Janssen, M.F., Kohlmann, T., Feng, Y.S.: A Systematic Review of Studies Comparing the Measurement Properties of the Three-Level and Five-Level Versions of the EQ-5D. PharmacoEconomics **36**(6), 645-661 (2018). doi:10.1007/s40273-018-0642-5
10. Diederichsen, S.Z., Haugan, K.J., Kober, L., Hojberg, S., Brandes, A., Kronborg, C., Graff, C., Holst, A.G., Nielsen, J.B., Krieger, D., Svendsen, J.H.: Atrial fibrillation detected by continuous electrocardiographic monitoring using implantable loop recorder to prevent stroke in individuals at risk (the LOOP study): Rationale and design of a large randomized controlled trial. American Heart Journal **187**, 122-132 (2017). doi:10.1016/j.ahj.2017.02.017
11. van Reenen, M., Janssen, B.: EQ-5D-5L User Guide, Basic information on how to use the EQ-5D-5L instrument. In. EuroQol Research Foundation, (2015). https://euroqol.org/publications/user-guides/. Accessed 9. August 2018
12. Janssen, M.F., Birnie, E., Haagsma, J.A., Bonsel, G.J.: Comparing the standard EQ-5D three-level system with a five-level version. Value Health **11**(2), 275-284 (2008). doi:10.1111/j.1524-4733.2007.00230.x
13. Shannon, C.E.: A mathematical theory of communication. Bell system technical journal **27**(3), 379-423 (1948).
14. Scalone, L., Ciampichini, R., Fagiuoli, S., Gardini, I., Fusco, F., Gaeta, L., Del Prete, A., Cesana, G., Mantovani, L.G.: Comparing the performance of the standard EQ-5D 3L with the new version EQ-5D 5L in patients with chronic hepatic diseases. Quality of Life Research **22**(7), 1707-1716 (2013). doi:10.1007/s11136-012-0318-0

15. Rencz, F., Lakatos, P.L., Gulacsi, L., Brodszky, V., Kurti, Z., Lovas, S., Banai, J., Herszenyi, L., Cserni, T., Molnar, T., Pentek, M., Palatka, K.: Validity of the EQ-5D-5L and EQ-5D-3L in patients with Crohn's disease. Quality of Life Research **28**(1), 141-152 (2019). doi:10.1007/s11136-018-2003-4
16. Campbell, M.J., Swinscow, T.D.V.: Statistics at square one 11 ed. John Wiley & Sons, (2009)
17. EuroQol Research Foundation: EQ-5D-5L User Guide: Basic information on how to use the EQ-5D-5L instrument (2019). https://euroqol.org/publications/user-guides/. Accessed 20. April 2020
18. Schmidt, M., Schmidt, S.A.J., Sandegaard, J.L., Ehrenstein, V., Pedersen, L., Sørensen, H.T.: The Danish National Patient Registry: A review of content, data quality, and research potential. Clinical Epidemiology **7**, 449-490 (2015). doi:10.2147/CLEP.S91125

Table 1: Descriptive statistics of differences in 3L VAS and 5L VAS

Differences in EQ VAS scores (0-100)	Higher 5L-VAS than 3L-VAS n=116	Lower 5L-VAS than 3L-VAS n=248
1-5	79 (68.1 %)	141 (56.9%)
6 -10	25 (21.6 %)	60 (24.2 %)
11-15	6 (5.2 %)	22 (8.9%)
>16	6 (5.2 %)	25 (10.1%)

Table 2: Respondents reporting no problems (ceiling effects) or maximum problems (floor effects) on the individual EQ-5D dimensions and on all dimensions.

	Number of observations		Ceiling effects						Floor effects					
Dimension	3L	5L	3L		5L		Absolute reduction	Relative reduction	3L		5L		Absolute reduction	Relative reduction
			N	%	N	%	%-point	%	N	%	N	%	%-point	%
Mobility	995	1001	781	79	671	67	12	15	0	0.0	3	0.3	-0.3	na
Self-Care	994	1001	958	96	930	93	3	4	1	0.1	1	0.1	0.0	0
Usual Activities	995	1001	777	78	679	68	10	13	6	0.6	4	0.4	0.2	33
Pain/ Discomfort	990	1001	561	57	402	40	17	29	12	1.2	4	0.4	0.8	67
Anxiety/ Depression	993	1001	858	86	795	79	7	8	1	0.1	0	0.0	0.1	100
Overall:	1002	1002	453	46	296	30	16	35	0	0	0	0	0	0

Table 3: Redistribution properties from 3L to 5L: cross tabulation of dimension scores (N, %)

		5L						Consistent		Inconsistent	
										Total	
3L	Dimension	1	2	3	4	5	Missing	N	%	N	%
1	Mobility	646	112	19	2	1	1	758	97%	22	3%
	Self-Care	919	37	1	0	0	1	956	99%	1	1%
	Usual Activities	646	117	11	2	0	1	763	98%	13	2%
	Pain/Discomfort	381	162	15	3	0	0	543	97%	18	3%
	Anxiety/Depression	774	70	11	2	0	1	844	98%	13	2%
2	Mobility	21	81	84	28	0	0	193	90%	21	10%
	Self-Care	4	20	10	1	0	0	31	89%	4	11%
	Usual Activities	28	108	56	19	1	0	183	86%	29	14%
	Pain/Discomfort	18	220	142	34	2	1	396	95%	20	5%
	Anxiety/Depression	15	74	38	7	0	0	119	89%	15	11%
3	Mobility	0	0	0	0	0	0	0	-	-	-
	Self-Care	0	0	0	0	1	0	1	100%	0	0%
	Usual Activities	0	1	0	3	2	0	5	83%	1	17%
	Pain/Discomfort	0	1	2	7	2	0	9	75%	3	25%
	Anxiety/Depression	0	0	0	1	0	0	1	100%	0	0%
Missing	Mobility	4	1	0	0	2	0	-	-	-	-
	Self-Care	7	0	0	1	0	0	-	-	-	-
	Usual Activities	5	0	1	0	1	0	-	-	-	-
	Pain/Discomfort	3	7	1	1	0	0	-	-	-	-
	Anxiety/Depression	6	2	1	0	0	0	-	-	-	-

Note: The size of the inconsistency is represented by the grayscale. The darker the more inconsistent.

Table 4: Discriminatory power (informativity) shown by Shannon's Index (H') and Shannon's Evenness Index (J') for 3L and 5L. Higher values indicate more informativity.

	3L		5L	
Dimension	H'	J'	H'	J'
Mobility	0.75	0.47	1.36	0.59
Self-Care	0.23	0.15	0.43	0.19
Usual Activities	0.80	0.50	1.29	0.55
Pain/Discomfort	1.07	0.67	1.71	0.74
Anxiety/Depression	0.58	0.37	0.96	0.41

Table 5: Convergent validity between EQ-5D and EQ VAS (Spearman´s correlation coefficient r_s)

	EQ VAS			
	3L		5L	
	N	r_s	N	r_s
Mobility	978	-0.3937	1001	-0.4894
Self-Care	977	-0.2374	1001	-0.2816
Usual Activities	978	-0.4494	1001	-0.5474
Pain/Discomfort	973	-0.4622	1001	-0.5198
Anxiety/Depression	976	-0.2147	1001	-0.2918

Note: Spearman's correlation coefficient is negative because dimension level increases with severity while the VAS score decrease with poor health status. The p-value is <0.0001 for all correlation coefficients.