



Universiteit
Leiden
The Netherlands

Sensitivity to change in well-being and health-related quality of life in adults with eating disorder symptoms: ICECAP-A versus EQ-5D-5L

Kaam, F.E. van; Rohrbach, P.J.; Marle, M.E.V.; Furth, E.F. van; Dingemans, A.E.

Citation

Kaam, F. E. van, Rohrbach, P. J., Marle, M. E. V., Furth, E. F. van, & Dingemans, A. E. (2024). Sensitivity to change in well-being and health-related quality of life in adults with eating disorder symptoms: ICECAP-A versus EQ-5D-5L. *International Journal Of Eating Disorders*, 57(3), 593-601. doi:10.1002/eat.24142

Version: Publisher's Version

License: [Licensed under Article 25fa Copyright Act/Law \(Amendment Taverne\)](#)

Downloaded from: <https://hdl.handle.net/1887/4214419>

Note: To cite this publication please use the final published version (if applicable).

ORIGINAL ARTICLE

Sensitivity to change in well-being and health-related quality of life in adults with eating disorder symptoms: ICECAP-A versus EQ-5D-5L

Fleur E. van Kaam MSc¹  | Pieter J. Rohrbach PhD²  |
M. Elske van den Akker-Van Marle PhD³  | Eric F. van Furth PhD^{1,4}  |
Alexandra E. Dingemans PhD^{1,4} 

¹GGZ Rivierduinen Eating Disorders Ursula, Leiden, The Netherlands

²Department of Clinical Psychology, Open University, Heerlen, The Netherlands

³Department of Biomedical Data Sciences, Section of Medical Decision Making, Leiden University Medical Center, Leiden, The Netherlands

⁴Department of Psychiatry, Leiden University Medical Center, Leiden, The Netherlands

Correspondence

Fleur E. van Kaam, GGZ Rivierduinen Eating Disorders Ursula, Leiden, The Netherlands.
Email: f.vankaam@rivierduinen.nl

Funding information

ZonMw, Grant/Award Number: 636310001

Action Editor: Ruth Striegel Weissman

Abstract

Objective: Economic evaluations of treatments help to inform decisions on allocating health care resources. These evaluations involve comparing costs and effectiveness in terms of quality of life. To calculate quality-adjusted life years, generic health related quality of life is often used, but is criticized for not being sensitive to change in mental health populations. Another approach, using experienced well-being measured through capabilities with the ICECAP-A, has been proposed as an alternative. The aim of this study was to investigate whether changes in individuals with eating disorder (ED) symptoms can be better captured using health related quality of life (EQ-5D-5L) or well-being (ICECAP-A).

Method: Measurements at two time points with an interval of 1 year were used from a sample of 233 participants with self-reported ED symptoms. An analysis of variance was used to test whether the EQ-5D-5L and ICECAP-A differed in sensitivity to change over time. In order to compare the two questionnaires in terms of clinically significant outcome, the reliable change index and clinical cut-off score were calculated.

Results: The two questionnaires did not differ in sensitivity to change. More individuals had recovered but also more had deteriorated according to the EQ-5D-5L compared to the ICECAP.

Discussion: The present study revealed no differences in sensitivity to change in health-related quality of life or well-being in individuals with ED symptoms in the context of mild clinical change. Results corroborated the pervasiveness of low quality of life in this population, even after alleviation of ED symptoms.

Public significance statement: Measuring treatment benefits in terms of improvements in quality of life is an integral part of economic evaluations in health care. It was expected that these treatment benefits would be better captured as changes in well-being (measured with the ICECAP-A) than as changes in health-related quality of life (measured with the EQ-5D-5L) for individuals with ED symptoms. Based on the results of this study, no preference for one of the two approaches was found.

KEYWORDS

capabilities, eating disorders, EQ-5D-5L, health-related quality of life, ICECAP-A, quality of life, responsiveness, well-being

1 | INTRODUCTION

Eating disorders (EDs) are serious psychiatric disorders characterized by disturbances in attitudes and behaviors toward eating, body weight, and body shape (American Psychiatric Association, 2013). Patients with EDs have high rates of mortality, chronicity, relapse, and comorbidity (Keel & Brown, 2010; Smink et al., 2013). A longer duration of ED symptoms is associated with poor outcome of treatment, lower chances of recovery and lower levels of quality of life (QoL) (Keel & Brown, 2010). Individuals who recovered from an ED still report a lower QoL compared to individuals from the general population (Hoeken & Hoek, 2020; Keel & Brown, 2010). In health care, health-related quality of life (HRQoL) is an assessment of how the individual's well-being may be affected over time by a disorder which includes all physical, mental, and social aspects of the individual's life (Mitchell et al., 2017). Generic instruments such as the EQ-5D-5L and the SF-36 are used to measure HRQoL (van Krugten, Busschbach, et al., 2021).

Besides a health related burden, there is also an economic burden on the individual patient with an ED as well as on society as a whole. Costs consist of direct healthcare costs and of non-healthcare costs such as loss of earnings and financial burden (Aardoom et al., 2016). In the United States, the yearly healthcare costs of patients with an ED are on average 48% higher than those of individuals in the general population (Samnaliev et al., 2015). Cost-effectiveness is therefore, together with effectiveness, an important consideration in the choice of treatment. Health economic evaluations can contribute in decision making on how to optimally allocate (scarce) health care resources (Byford et al., 2003).

Measuring treatment benefits in terms of improvements in QoL is an integral part of economic evaluations in health care. Cost-effectiveness analysis is a form of economic evaluation where relative costs and effects of different courses of action are compared. In health care, the effectiveness of treatments for both physical and mental health conditions can be captured with disorder specific questionnaires, but also with generic HRQoL instruments, such as the EQ-5D-5L (Devlin & Krabbe, 2013). These instruments use utility value (a measure of overall life satisfaction) to specify treatment benefits. This specific form of cost-effectiveness analysis, with utility as effectiveness measure, is referred to as cost-utility analysis. Utility values, which can be derived from raw scores on instruments like the EQ-5D-5L (Versteegh et al., 2016), can be transformed into quality-adjusted life years (QALYs). One QALY indicates one extra life year in perfect health. QALYs are often used in economic evaluation to assess the value of medical interventions, since they are generic and thus offer the possibility to compare across different health situations (Mitchell et al., 2017).

However, Mitchell et al. (2017) suggested that the EQ-5D-5L is not sensitive enough to capture subtle changes in mental well-being over time, which is important for evaluating treatment outcome in patients with mental health conditions, like EDs. The EQ-5D-5L

mainly focusses on physical aspects of QoL and only covers determinants important to QoL in mental health problems on one dimension (i.e., anxiety/depression). In psychiatric patients, who are likely to suffer less from physical distress compared to patients with somatic complaints, the use of the EQ-5D-5L may result in a score that does not reflect the true quality of life a patient experiences and subtle changes may not be reflected in the score (Al-Janabi et al., 2011; Keeley et al., 2013; Mitchell et al., 2017; Rohrbach et al., 2019). An alternative approach for assessing outcomes of treatment is by focusing on a patients' capabilities, which provide a broad approach to well-being and are defined by what a person can "do" and can "be" in their life and their ability to experience freedom and opportunities (Keeley et al., 2013; Mitchell et al., 2017). By assessing capabilities it is possible to evaluate the individual benefit from treatment based on the ability to function well on areas that are valuable to a person. For example, patients with EDs improving their self-efficacy in an early stage of treatment are more likely to show less severe eating ED symptoms at the end of treatment (Keshen et al., 2017). This indicates that a sense of control and achievement (i.e., self-efficacy) may be an important aspect of quality of life that may not be captured by generic instruments such as the EQ-5D-5L and SF-36.

The ICECAP-A (Al-Janabi et al., 2011) measures QoL in terms of capabilities, which might be more relevant in evaluating treatment outcome in mental health care compared to (physical) health (Al-Janabi et al., 2011; Al-Janabi et al., 2012; Keeley et al., 2013; Mitchell et al., 2017). The EQ-5D-5L measures a related but distinct concept than the ICECAP-A. The first measures HRQoL, while the latter measures a concept that might be described as well-being. The way these concepts are used in the present study, is that well-being (captured with the ICECAP-A) and HRQoL (captured with the EQ-5D-5L) are related in the sense that they are both subject of the concept of QoL, but have a different emphasis. Well-being puts more emphasis on the experienced freedom to be or do the things that one wishes to be or do, while HRQoL underscores (dis)ability and functioning.

It has been suggested that the ICECAP-A can be used in economic evaluations in a similar way as QALYs (Rohrbach, Dingemans, Groothuis-Oudshoorn, et al., 2022). Mitchell et al. (2017) have investigated the validity of the ICECAP-A in assessing treatment outcome in a sample of adults with depression. They found that changes in depression are more likely to be captured and valued using the ICECAP-A than with the EQ-5D-5L, and suggested that the ICECAP-A is therefore more suitable for economic evaluation of benefits of mental health care than the EQ-5D-5L. However, the heightened sensitivity to change of the ICECAP-A compared to the EQ-5D-5L has yet to be determined in an ED population.

The aim of this study was to investigate whether changes in ED symptoms can better be captured by changes in well-being (measured with the ICECAP-A) or changes in HRQoL (measured with the EQ-5D-

5L) and whether changes in ED symptoms and in perceived self-efficacy are more likely to be captured by the ICECAP-A than the EQ-5D-5L.

2 | METHOD

2.1 | Design and procedure

The present study was part of a larger randomized control trial (RCT) study aiming to evaluate the effectiveness and cost-effectiveness of Featback, a fully automated online self-help intervention, online expert-patient (i.e., someone who has successfully recovered from an ED) support, and a combination of Featback and online expert-patient support (Rohrbach et al., 2019; Rohrbach, Dingemans, van Furth, et al., 2022). Data of the ICECAP-A, EQ-5D-5L, Eating Disorder Examination Questionnaire (EDE-Q) and General Self-Efficacy Scale (GSES) collected at T1 (post-intervention) and T5 (12 months after T1) were used for this study. T1 was used instead of T0, because baseline measures of the ICECAP-A and EQ-5D-5L were not collected.

After screening, eligible participants were asked to give informed consent and fill in online baseline measures (T0). After the baseline assessment, participants were randomized to one of the four conditions by an independent researcher. After 8 weeks of intervention or waiting list, participants completed a post-intervention assessment (T1) online. Additional online post-intervention assessments were repeated after 3 (T2), 6 (T3), 9 (T4), and 12 (T5) months after T1. Ethical approval was obtained by an independent medical ethics committee (METC-LDD; NL64553.058.18).

2.2 | Participants

Participants were mainly recruited via the Dutch e-community “Proud2Bme” (<http://www.proud2bme.nl>) (Proud2Bme, 2009), the Featback website (<http://www.featback.nl>), and the network of the patient organization WEET (<http://www.WEET.info>). Table 1 shows the demographical and clinical characteristics of the participants included in the study sample.

The sample consisted of 233 participants, ranging from 18 to 63 years of age. All participants had access to the internet and had self-reported ED symptoms defined as scoring 52 or higher on the Weight Concerns Scale (WCS) (Killen et al., 1994), or reporting one or more ED symptoms assessed by the Short Evaluation of Eating Disorders (SEED) (Bauer et al., 2005). Participants younger than 16 years or not reporting any ED symptoms were excluded. Data were collected between October 2018 and December 2021.

2.3 | Measures

2.3.1 | ICECAP-A

The Dutch version of the ICECAP-A (Al-Janabi et al., 2011; Hoof et al., 2016; Rohrbach et al., 2021; Rohrbach, Dingemans,

TABLE 1 Characteristics of participants at T1 (post-intervention).

T1 (N = 233)	
Gender, n (%)	
Female	226 (97.0)
Male	5 (2.1)
Other	2 (0.9)
Age M, SD (in years), Nationality, n (%)	27.3, ±9.2 (range 15–63)
Dutch	209 (89.7)
Belgian	22 (9.4)
Other	2 (0.9)
Education, n (%)	
University, doctoral, post-doctoral	58 (25.0)
Higher vocational education	62 (26.7)
Post-secondary vocational education	46 (19.8)
Other	67 (28.4)
Socioeconomic status, n (%)	
At school/studying	90 (38.6)
Working	86 (36.9)
Not working	10 (4.3)
Unemployed	7 (3.0)
Incapacitated	40 (17.2)
Self-reported diagnosis status, n (%)	
Officially diagnosed with ED	155 (66.5)
No diagnosis, but assumed to have an ED	52 (22.3)
Eating problems, but likely no ED	26 (11.2)
Eating disorder treatment history, n (%)	
ED treatment history	140 (60.1)
No ED treatment history	93 (39.9)
Mean scores (SD)	
ICECAP-A	0.67 (0.19)
EQ-5D-5L	0.64 (0.24)
EDE-Q	4.80 (1.24)
GSES	25.82 (5.80)

Abbreviations: ED, Eating Disorder; EDE-Q, Eating Disorder Examination Questionnaire; GSES, General Self-Efficacy Scale.

Groothuis-Oudshoorn, et al., 2022) measures five capabilities which are important to an individual's QoL: “stability,” “attachment,” “autonomy,” “achievement,” and “enjoyment.” The five items (one item per capability) attempt to capture individuals capability to live a life that they values. Each capability item has four levels of response ranging from (1) not being able to experience a capability at all to (4) being able to fully experience a capability (Rohrbach et al., 2021). A tariff for the Dutch general population was used to transform raw scores on the ICECAP-A to capability index scores (Rohrbach, Dingemans, Groothuis-Oudshoorn, et al., 2022). Capability index scores range from 1 (full capability) to 0 (no capability). The ICECAP-A has good test-retest reliability (ICC = 0.79) and construct validity in a large Dutch

population (Rohrbach et al., 2021), and has acceptable levels of content validity (Keeley et al., 2013).

2.3.2 | EQ-5D-5L

The Dutch version of the EQ-5D-5L measures health determinants of QoL on five dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression (EuroQol Group, 2021; Herdman et al., 2011; Versteegh et al., 2016). Each dimension has five response levels ranging from “no problems” to “extreme problems/unable to” (Devlin & Krabbe, 2013), resulting in a health state that can be transformed to a utility score, ranging from 1 (perfect health) to −0.446 (worse than death) and anchored at 0 (death). QALYs can then be calculated by multiplying utility scores by the time (in years) a participant lived with that utility (Mitchell et al., 2017; Rohrbach, Dingemans, Groothuis-Oudshoorn, et al., 2022). One QALY indicates one extra life year in perfect health. QALYs are used in economic evaluation to assess the value of medical interventions since they offer the possibility to compare across different health situations (Mitchell et al., 2017). The current study used a tariff from the Dutch general population for this transformation to utility scores (Versteegh et al., 2016). The EQ-5D-5L has a good to excellent reliability and acceptable levels of content validity (Feng et al., 2020; Keeley et al., 2013).

2.3.3 | Eating Disorder Examination Questionnaire

The Dutch translation of the Eating Disorder Examination Questionnaire (EDE-Q 6.0) (Fairburn, 2008) was used to determine the range and severity of specific attitudes, cognitions and behaviors associated with ED symptomatology over the past 28 days. The EDE-Q generates two types of data: Frequency data on key behavioral features of EDs (6 items) (i.e., objective binge episodes, subjective binge episodes, self-induced vomiting, laxative misuse, diuretic misuse, excessive exercising), and a global ED psychopathology score (average of 22 items). Responses on these 2 items are recorded on a 7-point scale ranging from 0 to 6. Higher scores indicate higher ED psychopathology (Aardoom et al., 2012). Internal consistency, test-retest reliability, and discriminative validity of the EDE-Q have been found to be acceptable to high (Berg et al., 2011).

2.3.4 | General Self-Efficacy Scale

The GSES (Schwarzer & Jerusalem, 1995) is a 10-item psychometric scale designed to measure a general sense of perceived self-efficacy, with the aim to predict coping with daily hassles and adaptation after experiencing various stressful life events. Responses are recorded on a four-point scale ranging from 1 to 4. The individual items are summed to produce the final composite score with a range of 10–40. Psychometric properties of the GSES are adequate (Schwarzer & Jerusalem, 1995).

2.4 | Statistical analyses

Data of completers at T5 were used for statistical analyses. In order to investigate the association between the ICECAP-A and the EQ-5D-5L the correlation between the scores on both instruments were calculated at T1 and at T5. Additionally, the correlation between both the difference scores of T1 and T5, and the residual gain scores of the ICECAP-A and the EQ-5D-5L were calculated. The residual gain score takes into account differences in scores between participants at T1 and were obtained by performing a regression analysis, with T1 as independent and T5 as dependent variables.

For comparing responsiveness of two instruments, an interval of two measurements is required. Kind of treatment and the effect of treatment is not relevant for these analyses. Two comparable sets of statistics were used to determine whether the ICECAP-A and the EQ-5D-5L differed in responsiveness: effect size (Cohen's *d*) and standardized response mean (SRM) (Beurs et al., 2012; Carlier et al., 2015; Dingemans & van Furth, 2017). The effect size was calculated by dividing the difference between pre-test (T1) and post-test (T5) by the standard deviation (SD) of the pre-test (T1). The SRM was calculated by dividing the difference between the pre-test (T1) and the post-test (T5) by the SD of the difference score. Positive effect sizes and SRM values reflect the degree of (standardized) improvement relative to baseline scores or the score differences, respectively.

An analysis of variance for repeated measures in a 2 (questionnaire) × 2 (time) design was conducted to determine whether the ICECAP-A differed significantly from the EQ-5D-5L at T1 and at T5, and whether there was a significant difference between both instruments in sensitivity to change over time. A significant interaction between time and questionnaire indicates a difference in responsivity between the ICECAP-A and the EQ-5D-5L. In this analysis T1 was the independent variable and T5 the dependent variable.

In order to compare the ICECAP-A and the EQ-5D-5L in terms of clinically significant outcome, the reliable change index (RCI) and clinical cut-off score (CO) were calculated (Dingemans & van Furth, 2017; Jacobson & Truax, 1991). The RCI offers the possibility to determine whether people changed sufficiently that the change is unlikely to be due to simple measurement unreliability. A score below the CO implies that it resembles a score typical of the normal population more than a score typical of a dysfunctional population (Dingemans & van Furth, 2017). The RCI and CO were both calculated as formulated by Jacobson and Truax (1991). The formula used to calculate the RCI was: $RCI = 1.96 \times \sqrt{2(SEM)^2}$, with $SEM = SD_{T1} \times \sqrt{(1 - Cronbach's \alpha)}$. The CO was calculated with the following formula:

$$CO = (SD_{\text{typical population}} \times \text{Mean}_{\text{clinical sample T1}}) + (SD_{\text{clinical sample T1}} \times \text{Mean}_{\text{typical population}}) / (SD_{\text{typical population}} + SD_{\text{clinical sample T1}}).$$

Psychometrics from the typical (Dutch) population used for calculating the RCI and the CO of the ICECAP-A and the EQ-5D-5L were drawn from the study performed in the general population by

Rohrbach, Dingemans, Spinhoven, et al. (2022). The RCI and CO were also calculated for the EDE-Q for which psychometrics from the typical population were drawn from the study performed by Aardoom et al. (2012). The participants were categorized in four groups based on the RCI and the CO: recovered, improved, stable, and deteriorated. A Chi square-test was used to determine how the results, regarding frequencies of improved and recovered patients, differed between the ICECAP-A and the EQ-5D-5L. The same calculation was done for the EDE-Q.

To assess differences in responsiveness between the ICECAP-A and the EDE-Q on the one hand and between the EQ-5D-5L and the EDE-Q on the other, a multivariate analysis of variance for repeated measurements was performed with the scores of the three instruments on T1 and T5 (de Beurs et al., 2021). A significant interaction effect between time and the three instruments would indicate a difference between the instruments in change over time. A “simple” contrast was tested for the instrument effect, which compared the total score of the EDE-Q with the ICECAP-A and with the EQ-5D-5L. T1 and T5 were compared with the repeated contrast for time. The partial η^2 , as provided by SPSS, represents the between-instruments effect, the time-effect, and the interaction effect. An η^2 of 0.01 indicates a small effect, 0.06 a medium effect, and 0.14 a large effect (de Beurs et al., 2018). The same analyses were used to investigate whether the ICECAP-A was more sensitive than the EQ-5D-5L to capture changes in perceived self-efficacy measured with the GSES (Schwarzer & Jerusalem, 1995).

All analyses were performed using IBM SPSS Statistics 22.0 for Windows, with a significance level of $p < .05$.

3 | RESULTS

3.1 | Baseline characteristics of participants

The baseline characteristics of the participants in this study are shown in the Table 1. The percentage of participant's scores on the different scales of ICECAP-A and EQ-5D-5L at T1 and T5 can be found in the Supplemental Materials S1.

3.2 | Comparison of the ICECAP-A and the EQ-5D-5L

There was a strong correlation between the ICECAP-A and the EQ-5D-5L at T1 ($r = .64$) and T5 ($r = .67$). This may point to overlap in the concepts the two questionnaires measure. The difference scores and the residual gain scores (T1–T5), showed a moderate association on both indices ($r = .39$ and $r = .45$, respectively).

Contrary to our expectations, no significant interaction was found between time and questionnaire when comparing the ICECAP-A and the EQ-5D-5L with a repeated measures ANOVA in a 2×2 factorial design (Table 2) ($p = .25$), indicating that the change scores on the two instruments did not differ significantly from each other. The result suggested that the instruments captured QoL changes in the investigated sample similarly. When comparing responsiveness with the SRM however, analyses showed a trivial effect on the EQ-5D-5L and a small effect on the ICECAP-A, suggesting that although responsiveness is still small, the ICECAP-A is slightly more responsive. Responsiveness with the ES (Cohen's d) showed a trivial effect on both the ICECAP-A and the EQ-5D-5L (Table 2).

For comparing the ICECAP-A and the EQ-5D-5L in terms of clinically significant outcome, the RCI and the CO of both questionnaires were calculated (Table 3). Analysis revealed that the percentage of participants who had recovered and the percentage of participants who had deteriorated was higher on the EQ-5D-5L, while the percentage of participants who remained stable between T1 and T5 was higher for the ICECAP-A. The percentage of participants in the improved category was similar between the two instruments.

3.3 | Change in the ICECAP-A and the EQ-5D-5L in relation to change on the EDE-Q

There was a low negative correlation between the ICECAP-A and the EDE-Q at T1 as well as at T5 ($r = -.38$ and $r = -.48$, respectively), indicating that lower ED symptoms were associated with higher scores on the ICECAP-A. The same was true for correlations between the EQ-5D-5L and the EDE-Q at T1 and T5 ($r = -.43$ and $r = -.40$, respectively). Multivariate analysis of variance for repeated measurements

TABLE 2 Comparison of the responsiveness of the ICECAP-A, EQ-5D-5L, EDE-Q, and the GSES.

	T1		T5		F-test Time (T1 vs. T5) $\times$ Instrument (ICECAP-A vs. EQ-5D-5L)	ES (95% CI)	SRM
	M	SD	M	SD			
ICECAP-A (Capability score; range 0–1)	0.66	0.19	0.70	0.20	$F(1, 232) = 1.35, p = .25$	0.21 (0.08–0.34)	0.21
EQ-5D-5L (Utility score; range 0–1)	0.65	0.24	0.67	0.28		0.08 (–0.05–0.21)	0.09
EDE-Q (total score; range 0–6)	4.80	1.24	4.27	1.48		–0.43 (–0.56–0.29)	–0.44
GSES (self-efficacy score; range 10–40)	25.82	5.83	26.71	6.13		0.15 (0.02–0.28)	0.19

Note: Time $\times$ Instrument: interaction effect between time and instrument.

Abbreviations: CI, confidence interval; EDE-Q, Eating Disorder Examination Questionnaire; ES, effect size; GSES, General Self-Efficacy Scale; M, mean score; SD, standard deviation; SRM, standardized response mean.

TABLE 3 Comparing clinical significance and reliable change index between the ICECAP-A, the EQ-5D-5L, and the EDE-Q.

	CO-score	RCI	Clinically significant change				
			Recovered	Improved	Not changed	Deteriorated	
ICECAP-A	0.80	0.27	12 (5.2%)	7 (3.0%)	208 (89.3%)	6 (2.6%)	$\chi^2(9) = 38.93$
EQ-5D-5L	0.76	0.32	15 (6.4%)	7 (3.0%)	195 (83.7%)	16 (6.9%)	$p < .001$
EDE-Q	2.55	0.74	26 (11.2%)	55 (23.6%)	128 (54.9%)	24 (10.3%)	

Abbreviations: CO, cut-off score; EDE-Q, Eating Disorder Examination Questionnaire; RCI, reliable change index.

suggested a significant time (2 levels: T1 versus T5) $\times$ instrument (3 levels: EDE-Q versus ICECAP-A versus EQ-5D-5L) interaction effect, ($F = 44.93$, $p < .001$, $\eta^2 = 0.16$). Post hoc comparisons indicated that the EDE-Q varied significantly more over time compared to both the ICECAP-A, $F(1, 23) = 44.06$, $p < .001$, $\eta^2 = 0.16$, and the EQ-5D-5L, $F(1, 23) = 47.10$, $p < .001$, $\eta^2 = 0.17$. Specifically, deduced from the means at T1 and T5 of the three instruments, there was a significant decrease in ED symptoms while there was only a small increase in EQ-5D-5L and ICECAP-A scores. In other words, changes in ED symptoms are not strongly associated with changes in ICECAP-A and the EQ-5D-5L scores.

3.4 | ICECAP-A and EQ-5D-5L in relation to self-efficacy

There was a moderate correlation between the ICECAP-A and the GSES at T1, $r = .51$, as well as at T5, $r = .61$. The correlations between the EQ-5D-5L and the GSES at T1, $r = .36$, and T5, $r = .44$, could be considered weak. These results suggest that, compared to the EQ-5D-5L, the ICECAP-A is more associated with self-efficacy as measured by the GSES reflecting a difference in underlying concept.

When responsiveness of the GSES was compared with that of the ICECAP-A and the EQ-5D-5L, multivariate analysis of variance for repeated measures suggested significant time (2 levels: T1 versus T5) $\times$ instrument (3 levels: GSES versus ICECAP-A versus EQ-5D-5L) interaction effect, $F(2, 23) = 8.54$, $p < .001$, $\eta^2 = 0.04$. Post hoc comparisons indicated that the GSES increased more over time compared to both the ICECAP-A, $F(1, 23) = 8.36$, $p = .004$, $\eta^2 = 0.04$ and the EQ-5D-5L, $F(1, 23) = 8.74$, $p = .003$, $\eta^2 = 0.04$. These results suggest that changes in perceived self-efficacy are not strongly associated with changes in ICECAP-A and the EQ-5D-5L scores.

4 | DISCUSSION

The aim of the present study was to investigate whether changes over time in individuals with ED symptoms could be better captured by measuring well-being (with the ICECAP-A) or HRQoL (with the EQ-5D-5L). In contrast with previous studies that found the well-being approach (ICECAP-A) to be more sensitive to change compared to the HRQoL approach (EQ-5D-5L) in assessing outcome of patients with schizophrenia (Helter et al., 2020) and depressive disorder

(Mitchell et al., 2017), the present study revealed no statistically significant differences between the ICECAP-A and the EQ-5D-5L in terms of responsiveness in a population with ED symptoms.

The present study showed that while ED symptoms were significantly decreased at 12 months follow-up, QoL was not increased. This is in line with several studies which showed that QoL in patients with EDs was impaired for much longer after recovery from ED symptoms and remained low compared to controls with no history of ED (Hoeken & Hoek, 2020; Pohjolainen et al., 2016; Tomba et al., 2019); Pohjolainen et al. (2016) found that EDs have a long-lasting impact on both physical and mental QoL. Although there was improvement over an 8-year follow-up after treatment, QoL was still significantly impaired compared to the general population. The time frame in which this study took place may have been too short to see a meaningful improvement in QoL and to reveal any differences in responsiveness between the ICECAP-A and the EQ-5D-5L. Doll et al. (2005) showed that individuals with an ED history had poorer scores on emotional dimensions of the SF36, but did not differ significantly on the physical dimensions compared to individuals with no ED history. Additionally, improvement of physical impairments in EDs may be quicker compared to improvement of symptoms in the mental domain (De La Rie et al., 2005; Panea-Pizarro et al., 2021). Consequently, for adequately capturing changes in QoL of individuals with an ED, assessing improvements in both the physical and mental domain appears important, especially in the first stages of recovery. Both HRQoL, which may be especially relevant for somatic complaints that are involved in EDs, and well-being, which more broadly captures the physical and mental domains, may thus be applicable for this group. Based on the results of this study, there does not appear to be a preference between the EQ-5D-5L and the ICECAP-A in the ability to capture these improvements. That the heightened sensitivity to change of the ICECAP in other mental disorders was not replicated in this population with ED symptoms, may be because compared to other disorders EDs have a significant impact on physical functioning (Jenkins et al., 2011; Winkler et al., 2021), which are adequately captured by the HRQoL approach.

Unsurprisingly, the EDE-Q showed significantly more variation than both the ICECAP-A and the EQ-5D-5L. This is in line with evidence suggesting that QoL remains low, even after treatment has successfully decreased ED symptoms (Tomba et al., 2019). The pervasiveness of low QoL in individuals with ED symptoms makes it challenging to determine whether, if changes occur, these are better conceptualized as changes in well-being or HRQoL and are better captured by the ICECAP-A or EQ-5D-5L.

Furthermore, results showed that the GSES and the ICECAP-A are more associated than the GSES and the EQ-5D-5L. The association between the GSES and the ICECAP-A supports the idea that self-efficacy, which is the individual's belief about their own capabilities and control over their life, is part of the underlying construct the ICECAP-A intends to measure (Al-Janabi et al., 2011). However, no significant difference appeared between the ICECAP-A and the EQ-5D-5L in responsiveness to changes in self-efficacy.

4.1 | Strengths and limitations

A strong aspect of the present study is the fact that the sample consisted of a large number of participants with few dropouts. This study was limited by the fact that no pre-intervention data (T0) were available. The availability of pre-intervention data could have resulted in more variation in HRQoL and well-being, and possibly different findings on the sensitivity to change of the ICECAP-A and the EQ-5D-5L. Another limitation was that no distinction between ED types was made. This is important to address in future studies, as type of ED and BMI appears to be associated with HRQoL (Hart et al., 2020; Panea-Pizarro et al., 2021). Possibly, distinguishing between ED types and BMI would have led to different conclusions in this study.

4.2 | Implications and future directions

Although previous studies by Helter et al. (2020) and Mitchell et al. (2017) showed promising results in patients with depression and schizophrenia, future research is needed to further investigate the usefulness of the use of the ICECAP-A in evaluating treatment outcome in mental health care. Other instruments such as the Mental Health Quality of Life questionnaire (MHQoL), specifically designed to measure quality of life for economic evaluations in individuals with mental health problems (van Krugten, Feskens, et al., 2021), should also be considered in future research.

Furthermore, it would be valuable to investigate the responsiveness of the ICECAP-A and the EQ-5D-5L in an ED population across a longer time period, since recovery in QoL appears to exceed the period of 12 months the present study covered and often remains lower in patients with a lifetime history of ED compared to healthy controls.

5 | CONCLUSION

The present study showed no preference for using well-being (measured with the ICECAP-A) or HRQoL (measured with the EQ-5D-5L) in capturing changes in symptomatology and self-efficacy in a sample with individuals with ED symptoms in the context of mild clinical change. This raises the question if the ICECAP-A is a suitable instrument for economic evaluation of benefits from mental healthcare interventions, as suggested by Mitchell et al. (2017), at least in the

case of ED's. Based on the results of this study, there seems to be no preference for using the ICECAP-A or EQ-5D-5L for economic evaluations in an ED population.

AUTHOR CONTRIBUTIONS

Fleur E. van Kaam: Conceptualization; formal analysis; investigation; methodology; writing – original draft; writing – review and editing. **Pieter J. Rohrbach:** Data curation; investigation; project administration; supervision; writing – review and editing. **M. Elske van den Akker-Van Marle:** Conceptualization; writing – review and editing. **Eric F. van Furth:** Conceptualization; funding acquisition; writing – review and editing. **Alexandra E. Dingemans:** Conceptualization; funding acquisition; methodology; project administration; supervision; writing – original draft; writing – review and editing.

ACKNOWLEDGMENTS

The authors would like to thank E. de Beurs for advising in the statistical analysis.

FUNDING INFORMATION

ZonMw (Study number 636310001) is the subsidizing party for the study from which data for the current study were used.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no competing interests.

DATA AVAILABILITY STATEMENT

The datasets generated and analyzed for the current study are available from the corresponding author on reasonable request, and are included in the DANS repository (<https://doi.org/10.17026/dans-z6z-ctye>).

ETHICS STATEMENT

Ethical approval for this study was obtained from the Leiden University Medical Centre Ethical Committee (CME LUMC Leiden, reference number NL64553.058.18). All participants gave online informed consent for participating in the study.

ORCID

Fleur E. van Kaam  <https://orcid.org/0009-0002-2850-6591>

Pieter J. Rohrbach  <https://orcid.org/0000-0003-0795-466X>

M. Elske van den Akker-Van Marle  <https://orcid.org/0000-0002-5269-509X>

Eric F. van Furth  <https://orcid.org/0000-0002-1131-0886>

Alexandra E. Dingemans  <https://orcid.org/0000-0002-5499-7139>

REFERENCES

- Aardoom, J. J., Dingemans, A. E., Slof Op't Landt, M. C., & van Furth, E. F. (2012). Norms and discriminative validity of the eating disorder examination questionnaire (EDE-Q). *Eating Behaviors*, 13(4), 305–309. <https://doi.org/10.1016/j.eatbeh.2012.09.002>
- Aardoom, J. J., Dingemans, A. E., van Ginkel, J. R., Spinhoven, P., van Furth, E. F., & van den Akker-Van Marle, M. E. (2016). Cost-utility of an internet-based intervention with or without therapist support in

- comparison with a waiting list for individuals with eating disorder symptoms: A randomized controlled trial. *International Journal of Eating Disorders*, 49(12), 1068–1076. <https://doi.org/10.1002/eat.22587>
- Al-Janabi, H., Flynn, N., & Coast, J. (2011). Development of a self-report measure of capability wellbeing for adults: The ICECAP-A. *Quality of Life Research*, 21(1), 167–176. <https://doi.org/10.1007/s11136-011-9927-2>
- Al-Janabi, H., Peters, T. J., Brazier, J., Bryan, S., Flynn, T. N., Clemens, S., Moody, A., & Coast, J. (2012). An investigation of the construct validity of the ICECAP-A capability measure. *Quality of Life Research*, 22(7), 1831–1840. <https://doi.org/10.1007/s11136-012-0293-5>
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5)* (5th ed.). American Psychiatric Publishing.
- Bauer, S., Winn, S., Schmidt, U., & Kordy, H. (2005). Construction, scoring and validation of the short evaluation of eating disorders (SEED). *European Eating Disorders Review*, 13(3), 191–200. <https://doi.org/10.1002/erv.637>
- Berg, K., Peterson, C., Frazier, P., & Crow, S. (2011). Psychometric evaluation of the eating disorder examination and eating disorder examination-questionnaire: A systematic review of the literature. *International Journal of Eating Disorders*, 45(3), 428–438. <https://doi.org/10.1002/eat.20931>
- Beurs, E. D., Barendregt, M., Flens, G., van Dijk, E., Huijbrechts, I., & Meerding, W. J. (2012). Vooruitgang in de behandeling meten – Een vergelijking van vragenlijsten voor zelfrapportage (pp. 1–11). <https://mgv.boomtijdschriften.nl/Artikelen/gv-0-0-1.htm>
- Byford, S., McCrone, P., & Barrett, B. (2003). Developments in the quantity and quality of economic evaluations in mental health. *Current Opinion in Psychiatry*, 16(6), 703–707. <https://doi.org/10.1097/00001504-200311000-00017>
- Carlier, I., Kovács, V., van Noorden, M., van der Feltz-Cornelis, C., Mooij, N., Schulte-Van Maaren, Y., van Hemert, A., Zitman, F., & Giltay, E. (2015). Evaluating the responsiveness to therapeutic change with routine outcome monitoring: A comparison of the symptom Questionnaire-48 (SQ-48) with the brief symptom inventory (BSI) and the outcome Questionnaire-45 (OQ-45). *Clinical Psychology & Psychotherapy*, 24(1), 61–71. <https://doi.org/10.1002/cpp.1978>
- de Beurs, E., Carlier, I., & van Hemert, A. (2021). Psychopathology and health-related quality of life as patient-reported treatment outcomes: Evaluation of concordance between the brief symptom inventory (BSI) and the short Form-36 (SF-36) in psychiatric outpatients. *Quality of Life Research*, 31(5), 1461–1471. <https://doi.org/10.1007/s11136-021-03019-5>
- de Beurs, E., Vissers, E., Schoevers, R., Carlier, I. V. E., van Hemert, A. M., & Meesters, Y. (2018). Comparative responsiveness of generic versus disorder-specific instruments for depression: An assessment in three longitudinal datasets. *Depression and Anxiety*, 36(1), 93–102. <https://doi.org/10.1002/da.22809>
- de la Rie, S., Noordenbos, G., & van Furth, E. F. (2005). Quality of life and eating disorders. *Quality of Life Research*, 14(6), 1511–1521. <https://doi.org/10.1007/s11136-005-0585-0>
- Devlin, N. J., & Krabbe, P. F. M. (2013). The development of new research methods for the valuation of EQ-5D-5L. *The European Journal of Health Economics*, 14(S1), 1–3. <https://doi.org/10.1007/s10198-013-0502-3>
- Dingemans, A., & van Furth, E. (2017). Het meten van verandering tijdens behandeling voor eetstoornissen: een vergelijking van algemene en specifieke vragenlijst. *Tijdschrift voor Psychiatrie*, 59(5), 278–285.
- Doll, H., Petersen, S., & Stewart-Brown, S. (2005). Eating disorders and emotional and physical well-being: Associations between student self-reports of eating disorders and quality of life as measured by the SF-36. *Quality of Life Research*, 14(3), 705–717. <https://doi.org/10.1007/s11136-004-0792-0>
- EQ-5D-5L. (2021). Euroqol. <https://euroqol.org/eq-5d-instruments/eq-5d-5l-about/>
- Fairburn, C. (2008). *Cognitive behavior therapy and eating disorders*. Guilford Press.
- Feng, Y., Kohlmann, T., Janssen, M., & Buchholz, I. (2020). Psychometric properties of the EQ-5D-5L: A systematic review of the literature. *Quality of Life Research*, 30(3), 647–673. <https://doi.org/10.1007/s11136-020-02688-y>
- Hart, L. M., Gordon, A. R., Sarda, V., Calzo, J. P., Sonnevile, K. R., Samnaliev, M., & Austin, S. B. (2020). The association of disordered eating with health-related quality of life in U.S. young adults and effect modification by gender. *Quality of Life Research*, 29(5), 1203–1215. <https://doi.org/10.1007/s11136-019-02396-2>
- Helter, T. M., Coast, J., Łaszewska, A., Stamm, T., & Simon, J. (2020). Do alternative capability instruments capture differing aspects of mental health and quality of life in comparison to the EQ-5D for schizophrenic patients with depression? *Research Square*, 29(6), 1433–1464. <https://doi.org/10.1007/s11136-019-02393-5>
- Herdman, M., Gudex, C., Lloyd, A., Janssen, M., Kind, P., Parkin, D., Bonsel, G., & Badia, X. (2011). Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of Life Research*, 20(10), 1727–1736. <https://doi.org/10.1007/s11136-011-9903-x>
- Hoeken, D., & Hoek, H. W. (2020). Review of the burden of eating disorders: Mortality, disability, costs, quality of life, and family burden. *Current Opinion in Psychiatry*, 33(6), 521–527. <https://doi.org/10.1097/ycp.0000000000000641>
- Hoof, M. V., Jeuring, S., Jonkers, D., Masclee, A., Pierik, M., Stokroos, R., & Joore, M. (2016). De Nederlandse vertaling en inblikvalidatie van de ICECAP-A: meten van kwaliteit van leven volgens de capability-benadering. *Tijdschrift Voor Gezondheidswetenschappen*, 94(8), 313–320. <https://doi.org/10.1007/s12508-016-0109-z>
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59(1), 12–19. <https://doi.org/10.1037/0022-006x.59.1.12>
- Jenkins, P. E., Hoste, R. R., Meyer, C., & Blissett, J. M. (2011). Eating disorders and quality of life: A review of the literature. *Clinical Psychology Review*, 31(1), 113–121. <https://doi.org/10.1016/j.cpr.2010.08.003>
- Keel, P. K., & Brown, T. A. (2010). Update on course and outcome in eating disorders. *International Journal of Eating Disorders*, 43, 195–204. <https://doi.org/10.1002/eat.20810>
- Keeley, T., Al-Janabi, H., Lorgelly, P., & Coast, J. (2013). A qualitative assessment of the content validity of the ICECAP-A and EQ-5D-5L and their appropriateness for use in Health Research. *PLoS One*, 8(12), e85287. <https://doi.org/10.1371/journal.pone.0085287>
- Keshen, A., Helson, T., Town, J., & Warren, K. (2017). Self-efficacy as a predictor of treatment outcome in an outpatient eating disorder program. *Eating Disorders*, 25(5), 406–419. <https://doi.org/10.1080/10640266.2017.1324073>
- Killen, J., Taylor, C., Hayward, C., Wilson, D., Haydel, K., Hammer, L., Simmonds, B., Robinson, T., Litt, I., & Varady, A. (1994). Pursuit of thinness and onset of eating disorder symptoms in a community sample of adolescent girls: A three-year prospective analysis. *The International Journal of Eating Disorders*, 16(3), 227–238.
- Mitchell, P. M., Al-Janabi, H., Byford, S., Kuyken, W., Richardson, J., Iezzi, A., & Coast, J. (2017). Assessing the validity of the ICECAP-A capability measure for adults with depression. *BMC Psychiatry*, 17(1), 1–13. <https://doi.org/10.1186/s12888-017-1211-8>
- Panea-Pizarro, I., Moran, J. M., Lavado-García, J., Beato-Fernández, L., Domínguez-Martin, A. T., Huerta-González, S., Novo, A., Pedrera-Zamorano, J. D., & López-Espuela, F. (2021). Health-related quality of life in Spanish women with eating disorders. *Nutrients*, 13(2), 403403. <https://doi.org/10.3390/nu13020403>
- Pohjolainen, V., Koponen, S., Räsänen, P., Roine, R. O., Sintonen, H., & Karlsson, H. (2016). Long-term health-related quality of life in eating

- disorders. *Quality of Life Research*, 25(9), 2341–2346. <https://doi.org/10.1007/s11136-016-1250-5>
- Proud2Bme. (2009). Online hulp voor iedereen met een eetstoornis. <https://www.proud2bme.nl/>
- Rohrbach, P. J., Dingemans, A. E., Essers, B. A., van Furth, E. F., Spinhoven, P., Groothuis-Oudshoorn, C. G. M., van Til, J. A., & van den Akker-Van Marle, M. E. (2021). The ICECAP-A instrument for capabilities: Assessment of construct validity and test-retest reliability in a general Dutch population. *Quality of Life Research*, 31(3), 687–696. <https://doi.org/10.1007/s11136-021-02980-5>
- Rohrbach, P. J., Dingemans, A. E., Groothuis-Oudshoorn, C. G., van Til, J. A., Essers, B. A., van Furth, E. F., & van den Akker-Van Marle, M. E. (2022). The ICEpop capability measure for adults instrument for capabilities: Development of a tariff for the Dutch general population. *Value in Health*, 25(1), 125–132. <https://doi.org/10.1016/j.jval.2021.07.011>
- Rohrbach, P. J., Dingemans, A. E., Spinhoven, P., van den Akker-Van Marle, E., van Ginkel, J. R., Fokkema, M., Moessner, M., Bauer, S., & van Furth, E. F. (2019). A randomized controlled trial of an internet-based intervention for eating disorders and the added value of expert-patient support: Study protocol. *Trials*, 20(1), 509. <https://doi.org/10.1186/s13063-019-3574-2>
- Rohrbach, P. J., Dingemans, A. E., Spinhoven, P., van Ginkel, J. R., Fokkema, M., Wilderjans, T. F., Bauer, S., & van Furth, E. F. (2022). Effectiveness of an online self-help program, expert-patient support, and their combination for eating disorders: Results from a randomized controlled trial. *The International Journal of Eating Disorders*, 55, 1361–1373. <https://doi.org/10.1002/eat.23785>
- Rohrbach, P. J., Dingemans, A. E., van Furth, E. F., Spinhoven, P., van Ginkel, J. R., Bauer, S., & van den Akker-Van Marle, M. E. (2022). Cost-effectiveness of three internet-based interventions for eating disorders: A randomized controlled trial. *International Journal of Eating Disorders*, 55(8), 1143–1155. <https://doi.org/10.1002/eat.23763>
- Samnaliev, M., Noh, H. L., Sonnevile, K. R., & Austin, S. B. (2015). The economic burden of eating disorders and related mental health comorbidities: An exploratory analysis using the U.S. medical expenditures panel survey. *Preventive Medicine Reports*, 2, 32–34. <https://doi.org/10.1016/j.pmedr.2014.12.002>
- Schwarzer, R., & Jerusalem, M. (1995). Generalized self-efficacy scale. In J. Weinman, S. Wright, & M. Johnston (Eds.), *Measures in health psychology: A user's portfolio. Causal and control beliefs*. NFER-NELSON.
- Smink, F. R. E., van Hoeken, D., & Hoek, H. W. (2013). Epidemiology, course, and outcome of eating disorders. *Current Opinion in Psychiatry*, 26(6), 543–548. <https://doi.org/10.1097/ycp.0b013e328365a24f>
- Tomba, E., Tecuta, L., Crocetti, E., Squarcio, F., & Tomei, G. (2019). Residual eating disorder symptoms and clinical features in remitted and recovered eating disorder patients: A systematic review with meta-analysis. *International Journal of Eating Disorders*, 52(7), 759–776. <https://doi.org/10.1002/eat.23095>
- van Krugten, F. C. W., Busschbach, J. J. V., Versteegh, M. M., Hakkaart-Van Roijen, L., & Brouwer, W. B. F. (2021). The mental health quality of life questionnaire (MHQoL): Development and first psychometric evaluation of a new measure to assess quality of life in people with mental health problems. *Quality of Life Research*, 31(2), 633–643. <https://doi.org/10.1007/s11136-021-02935-w>
- van Krugten, F. C. W., Feskens, K., Busschbach, J. J. V., Hakkaart-Van Roijen, L., & Brouwer, W. B. F. (2021). Instruments to assess quality of life in people with mental health problems: A systematic review and dimension analysis of generic, domain- and disease-specific instruments. *Health and Quality of Life Outcomes*, 19(1), 249. <https://doi.org/10.1186/s12955-021-01883-w>
- Versteegh, M., Vermeulen, M., Evers, M., de Wit, G. A., Prenger, R., & Stolk, A. (2016). Dutch tariff for the five-level version of EQ-5D. *Value in Health*, 19(4), 343–352. <https://doi.org/10.1016/j.jval.2016.01.003>
- Winkler, L. A., Gudex, C., Lichtenstein, M. B., Røder, M. E., Adair, C. E., Sjögren, M., & Støving, R. K. (2021). Explanatory factors for disease-specific health-related quality of life in women with anorexia nervosa. *Journal of Clinical Medicine*, 10(8), 15921592. <https://doi.org/10.3390/jcm10081592>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: van Kaam, F. E., Rohrbach, P. J., van den Akker-Van Marle, M. E., van Furth, E. F., & Dingemans, A. E. (2024). Sensitivity to change in well-being and health-related quality of life in adults with eating disorder symptoms: ICECAP-A versus EQ-5D-5L. *International Journal of Eating Disorders*, 57(3), 593–601. <https://doi.org/10.1002/eat.24142>