Received: 3 March 2023 | Revised: 20 July 2023 | Accepted: 14 August 2023

DOI: 10.1111/aas.14322

RESEARCH ARTICLE

acta Anaesthesiologica Scandinavica

Translation and validation of the Norwegian version of the postoperative quality of recovery score QoR-15

Linda Bergestuen[1] | Thomas Moger[1] | Kjersti Oterhals[2] | Frank Pfeffer[3] | Torunn Nestvold[4] | Stig Norderval[5] | Linn Såve Nymo[5] | Kjerstin Havnes[6] | Kristoffer Lassen[7] | Kyrre Breivik[2,8]

[1]Department of Surgery, Innlandet Hospital Trust, Lillehammer, Norway

[2]Centre on Patient-Reported Outcomes, Department of Research and Development, Haukeland University Hospital, Bergen, Norway

[3]Department of Gastrointestinal Surgery, Haukeland University Hospital, Bergen, Norway

[4]Department of Surgery, Nordland Hospital Trust, Bodø, Norway

[5]Department of Gastrointestinal surgery, University Hospital of North Norway, Tromsø and Institute of Clinical Medicine, the Arctic University of Norway, Tromsø, Norway

[6]Division of Surgery, Oncology and Women's Health, University Hospital of North Norway, Tromsø, Norway

[7]HPB surgery, Oslo university Hospital at Rikshospitalet, Oslo, Norway

[8]Regional Centre for Child and Youth Mental Health and Child Welfare, NORCE Norwegian Research Centre, Bergen, Norway

Correspondence
Thomas Moger, Department of Surgery Lillehammer, Innlandet Hospital Thrust, Pb 104, 2381 Lillehammer, Norway.
Email: thomas.moger@sykehuset-innlandet.no

Funding information
Norske Legeforening; Sykehuset Innlandet HF

Abstract

Background: As patient-centered care gains more attention, assessing the patient's perspective on their recovery has become increasingly important. In response to the need for a reliable and valid patient reported outcome measurement tool for major surgical resections in Norway. The Norwegian Registry for Gastrointestinal Surgery (NORGAST) initiated a project to translate and evaluate QoR-15's psychometric properties for patients going through general, gastrointestinal (GI), and hepato-pancreato-biliary (HPB) resectional surgery.

Methods: After a translation and adaption of the original version of QoR–15 into Norwegian, the QoR–15NO was psychometrically evaluated including a confirmatory factor analysis to test for unidimensionality, as well as tests for content validity, internal consistency, measurement error, construct validity, feasibility, and responsiveness. This process included cognitive interviews using a structured interview guide. Further, patients who underwent various types of GI/HPB surgery at five hospitals in different parts of Norway completed the QoR-15NO before surgery and on the first or second day after surgery. The impact of surgery was classified according to Surgical

Thomas Moger and Linda Bergestuen have contributed equally at the writing of this manuscript.

Thomas Moger, Frank Pfeffer, Torunn Nestvold, Stig Norderval, Linn Såve Nymo, Kjerstin Havnes, Kristoffer Lassen: "The Norwegian Registry for Gastrointestinal surgery."

Outcome Risk Tool v2 (SORT), in extra major/complex, major, intermediate, and minor.

Results: This study included 324 patients with 83% return rate with both pre- and postoperative forms. There were negative correlations between duration of surgery and postoperative QoR-15 score and the difference between post- and preoperative score (change score). Individuals who had gone through surgery with major impact had a lower postoperative mean QoR-15 score (97) than their counterparts who had experienced either medium (QoR-15: 110) or minor (QoR15: 119) impact surgery. Cronbach's alpha (0.88) and Omega Alpha Total ($\omega t = 0.90$) indicate that the scale has good to very good internal consistency. Test–retest reliability was measured by Intra-class Correlation Coefficient to ICC = 0.70. Confirmatory factor analyses supported that a one-factor model with correlated residuals had a good fit to data.

Conclusion: This study supports QoR-15NO as a valid, essentially unidimensional, feasible, and responsive instrument among patients undergoing general, GI, and HPB resectional surgery in Norway. The total QoR-15NO score provides important information that can be used in an everyday clinical setting and integrated into NORGAST.

KEYWORDS
NORGAST, psychometric evaluation, QoR-15, validation

Editorial Comment
This article presents a cross-validation of the postoperative quality of recovery scoring instrument QoR-15 in Norwegian language in population of major non-cardiac surgery and including males.

1 | INTRODUCTION

The Norwegian Registry for Gastrointestinal Surgery (NORGAST) requested a patient reported outcome measurement tool for patients undergoing major gastrointestinal (GI) and hepato-pancreato-biliary (HPB) resections evaluating the patients' status regarding essential parameters like respiration, nausea, pain, and received information. NORGAST annually retrieves data regarding perioperative complications up to 30 days postoperatively for about 6000 patient trajectories after major GI/HPB surgery.[1,2]

Outcome after surgery is typically evaluated by complication rates, length of stay in hospital (LoS), 90-day mortality rates, and survival. An increased focus on the patients' own perspective on their received healthcare has led to the development of several patient-centered measurement tools that aim to assess quality of recovery (QoR) after surgery.[3] High-quality surgical activity also entails good systems for postoperative follow–up. Patients' assessment of their own condition is important feedback to health care providers as a starting point for continuous improvement.

QoR-15 has been developed and implemented in clinical practice in at least 20 countries. Psychometric evaluations have been done in 16 countries. Myles and colleagues describe in a systematic review that QoR-15 has excellent validity and test–retest validity, in addition to high responsiveness, excellent completion, and return rates.[3] The minimal clinically important difference (MCID) was 6.0.[4] Furthermore, it shows essential unidimentional measure of QoR.[5] This supports that QoR-15 is simple-to-use outcome measure applicable across a broad range of surgical settings, and it has become the most widely reported measure of patient-assessed QoR after surgery.[3]

We aimed to translate QoR-15 into Norwegian, to do a psychometric evaluation through test for structural validity, content validity, internal consistency, measurement error, construct validity, feasibility, and responsiveness for patients going through general, GI, and HPB resectional surgery. In line with previous research,[5] we hypothesized that the instrument would be essentially unidimensional and tested this by the use of confirmatory factor analysis (CFA). As for construct validity, we hypothesized that there would be a negative correlation of age, gender, duration of surgery, and severity of surgery with the QoR-15NO score. The measurement properties have been evaluated according to the Consensus based Standards for the selection of health Measurement Instruments (COSMIN).[6,7]

A validation of QoR-15 was published in Norwegian in November 2020, after the initiation of this project.[8] However, it was validated in female patients only, after gynecological day-surgery. We aim to validate in a different and broader patient population after major

in-hospital surgery, including male participants, and perform an extensive psychometric validation.

2 | METHODS

2.1 | The instrument

The QoR-15 scale is a global measure of postoperative recovery and contains 15 questions. Each item uses an 11-point numeric rating scale, which generates a total score ranging from 0 (extremely poor QoR) to 150 (excellent QoR).[3] The questionnaire incorporates five domains of health: patient support, comfort, emotions, physical independence, and pain.[9]

2.2 | Translation and cultural adaption

After obtaining consent from the author of QoR-15[9] forward-and-back translations were performed according to the method by ISPOR Task Force for Translation and Cultural Adaption Process for Patient-Reported Outcomes (PRO).[10] Two independent bilingual translators with Norwegian as their mother tongue translated the original English version of QoR-15 into Norwegian. These two versions were compared and merged into a single forward translation by members of the NORGAST Council. The reconciliation was then blinded back translated to English by two independent translators with English as their mother tongue and compared with the original English version. Regarding question six and eight, some small adjustments were made. The first version of QoR-15 was then ready for cognitive debriefing to check understandability, interpretation, and cultural relevance for the translation.

Initially, 29 patients completed the QoR-15 NO on their first day after surgery. The time spent on completing the questionnaire was recorded. A study nurse, using a structured interview guide interviewed the patients individually to assess the level of comprehensibility and cognitive equivalence of the translation to reveal if the questions were understandable.

2.3 | Inclusion

Patients undergoing GI/HPB surgery from September 2021 until May 2022 were enrolled. Inclusion criteria were patients scheduled for elective surgery, aged 18 years or older and being fluent in Norwegian. Patients with cognitive failure, severe mental illness, or for other reasons unable to answer were excluded. Three university and two local hospitals participated, representing all Norwegian health regions; Innlandet Hospital Trust Lillehammer: 116 patients, Oslo University Hospital Rikshospitalet: 37 patients, Haukeland University Hospital: 50 patients, Nordland Hospital Trust: 30 patients, and University Hospital of North Norway: 35 patients.

2.4 | Ethics statement

The Regional Committees for Medical and Health Research Ethics concluded that no ethical approval was necessary (application number 263327). Approval by the Data Protection Officers at all five hospitals were obtained prior to start of the study. All patients provided written consent.

2.5 | Data collections

At the day of admission for surgery, eligible patients received a consent form including written information and instructions regarding the questionnaires, including a paper version of QoR–15NO. Then QoR-15NO was repeated the first day after surgery, except for one hospital where 37 patients were in intensive care unit the first day. These patients completed the postoperative questionnaire on the second day after surgery. Gender, age, ASA score (American Society of Anaesthesiologists), duration of surgery, and impact of surgery, classified by Surgical Outcome Risk Tool (SORT) v2[11] were retrieved from patient records at each site. According to the SORT classification, GI and HPB surgery are mainly classified as major and extra major/complex surgery.[11] To check reliability, test–retest was performed in 54 patients before surgery; once at the outpatient clinic and then secondly at the day of admission for surgery.

2.6 | Data analysis

Descriptive statistics, correlations, test–retest reliability, and paired *t*-test analyses were analyzed in the IBM SPSS Statistics Version 25.0 software. The remaining analyses were conducted in the R statistical environment[12] using various packages. The data were treated as continuous in all of the statistical analyses as the items of the QOR-15 had 11 response categories.[13] Missing data were handled by the use of listwise deletion in all the analyses except for the confirmatory factor analyses.

Test of unidimensionality (structural validity) was performed using CFA using the Lavaan package version 06.12.[14] The CFA models were estimated by Robust Maximum Likelihood. Missing data in the CFA analyses were handled using Full Information Maximum Likelihood and has the advantage (over listwise deletion) of using all the available data from all the individuals ($n = 268$) included in the analyses even if they have missing on some of the items.[15] Model fit was assessed using robust versions of Chi-square, the comparative fit index (CFI robust), the root mean square error of approximation (RMSEA robust), and standardized root mean squared residual (SRMR robust). CFI values greater than 0.90 together with RMSEA values of less than 0.08 and SRMR of less than 0.08 were considered acceptable,[16] whereas CFI values above 0.95 and RMSEA of below 0.06 and SRMR of less than 0.05 were preferred.[17]

Internal consistency reliability was assessed by Cronbach's alpha, Omegatotal alpha available in the in the Psych package[18] and composite reliability from the SemTools package.[19] Test–retest reliability was measured by Intra-class Correlation Coefficient (ICC 3,1) average measures, absolute agreement, two-way mixed ANOVA model. The test–retest was performed to check reliability. ICC above 0.90 was interpreted as excellent, 0.75 to 0.9 as good, 0.5 to 0.75 as moderate, and below 0.5 as poor test–retest reliability.[20]

Measurement error were assessed by the standard error of measurement (SEM) and the smallest detectable change (SDC). SEM was estimated by the following formula: Standard deviation from first measurement in the test–retest * (√1-intra-class coefficient (ICC)). SDC was calculated by multiplying SEM with 1.96.[5]

Hypothesis testing for construct validity was assessed by the use of Pearson's correlation coefficients and one way ANOVA analysis (using Least significant difference test for post hoc tests) when analyzing associations and group mean differences, respectively.

Regarding responsiveness, changes from baseline were assessed by paired *t*-tests. The amount of change was assessed by Cohen's *d* (change score divided by Standard deviation pretest score) and standardized response mean (change score divided by Standard deviation change scores) of which 0.2, 0.5, and 0.8 standardized units are regarded as small, medium, and large effect sizes, respectively.[3,21] Floor and ceiling effect was defined as 15% or more receiving the highest or lowest score, respectively.[22]

2.7 | Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

3 | RESULTS

3.1 | Feasibility

A total of 324 patients were included. Of these 56 patients did not return the postoperative questionnaire and were not eligible for analyses. The completion and return rate were 83%. Of the 268 patients considered eligible, 57 returned incomplete forms, with one or more of the items not answered (see Table 3).

Time spent completing QoR-15NO were measured for the 29 patients who underwent the cognitive interviews, ranging from 4 to 6 min, mean time 4.5 min.

3.2 | Patient characteristics

Patient characteristics are presented in Table 1. The mean duration of surgery was 154 min (SD 88), and 85% of the patients underwent major or Xmajor/complex surgery.

TABLE 1 Patient characteristics ($N = 268$).

Male	137 (51%)
Age, (years) Mean (SD)	64 (16)
ASA score	
I	14 (5%)
II	172 (64%)
III	81 (30%)
IV	1 (0,5%)
Duration of surgery (minutes) Mean (SD)	154 (88)
Severity of surgery, according to SORT	
Minor	5 (2%)
Intermediate	34 (13%)
Major	76 (28%)
Extra major/complex	153 (57%)

Note: Data are presented as frequencies (percentages) unless otherwise stated; ASA, American Society of Anesthesiologists[23]; SORT, Surgical Outcome Risk Tool.[11]

3.3 | Content validity

The results from the structured interviews were assessed by an expert group comprised of a nurse, surgeon and PROM expert, evaluating the questionnaires content validity. Repeated patient feedback regarding unclear words and phrases would lead to revisions, but no such feedback occurred. The QoR-15 NO was then ready to use for further validation.

3.4 | Structural validity—Dimensionality assessment

The one-factor model (M1) had an unsatisfactory fit to the data (Table 2). When including the 10 correlated error terms, implied by the bifactor model in Kleif et al.,[5] the fit of the model (M2) became acceptable (CFI >0.90, RMSEA and SRM <0.08). Modification indices suggested that this model could be improved by including correlated error terms between q13 (nausea) and q14 (anxious) and between q13 and q15 (depressed). Three pairs of correlated error terms could be removed due to statistical non-significance (q8 & q5, $r = 0.06$, $p = 0.35$; q8 & q11, $r = 0.03$, $p = 0.74$; q8 & q12, $r = -0.10$, $p = 0.13$. These changes (M3) led to an improved fit. CFI (>0.95) and RMSEA (<0.05) were now both beneath, and SRMS slightly above (<0.05), the preferred cut off values suggested by Hu and Bentler.[17]

The final revised model (M3) is shown in Figure 1. While significant, the factor loadings tied to q3 (feeling rested), q9 (feeling comfortable and in control), and q10 (having a feeling of general well-being) were very large (>0.75), the items assessing severe pain (q12), nausea (q13), feeling anxious (q14), and feeling depressed (q15) had low loadings on the QoR factor (0.24–0.33). All the correlated residuals were statistically significant ($p < .05$) and most were of medium size ($r = 0.30$–0.50). The correlation between the residuals of feeling anxious and depressed was, however, large ($r = 0.75$).

TABLE 2 Fit of models.

	Chi-squared robust	CFI robust	RMSEA robust (CI)	SRMS robust
M1: One-factor model	471.102, df = 90, $p < .001$	0.727	0.145 (0.132, 0.158)	0.103
M2: One-factor model with 10 correlated residuals[a]	175.874, df = 80, $p < .001$	0.936	0.075 (0.060, 0.090)	0.073
M3: One factor model with 9 correlated residuals	123.346, df = 81, $p < .01$	0.972	0.049 (0.031, 0.066)	0.057

Abbreviations: CFI, comparative fit index; RMSEA, root mean square error of approximation; SRMS, standardized root mean squared residual.
[a]Based upon Kleif et al.[5]

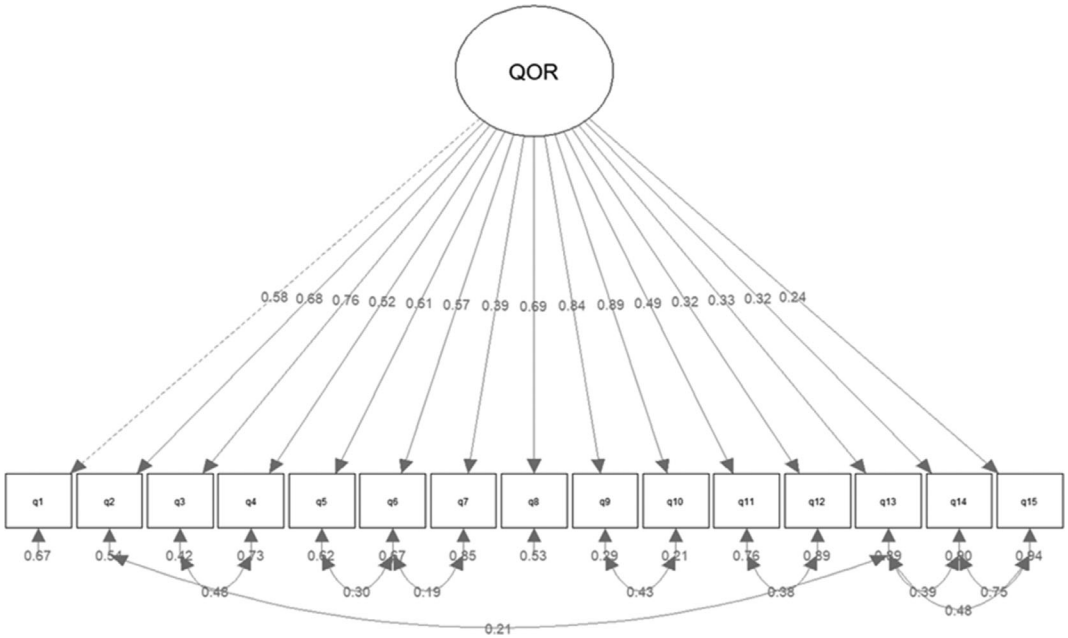

FIGURE 1 Final revised model M3. One factor model with nine correlated residuals. QoR, quality of recovery.

3.5 | Reliability

Cronbach's alpha (0.88) and Omega alphaTotal ($\omega t = 0.90$) indicate that the scale has good to very good internal consistency. Composite reliability based on the final correlated residual model (model 3) also suggest that the reliability is good (0.80). For the test–retests the time from consultation to surgery varied from 1 to 113 days, median time was 23 days. Test–retest reliability was assessed as moderate (ICC = 0.70, CI = 0.47–0.83).

3.6 | Measurement error

SEM and the upper 95% confidence limit of the SDC were estimated to be 10.80 and 29.84, respectively.

3.7 | Hypothesis testing for construct validity

There were significantly negative correlations between duration of surgery and postoperative QoR-15 score ($r = -0.30$, $p < .001$) and change score (difference between post- and preoperative score) ($r = -0.24$, $p < 0.001$). Individuals who had gone through extra major/complex surgery had a statistically significantly lower postoperative mean QoR-15 score (96.91, SD = 25.91) than their counterparts who had experienced either major (109.93, SD = 25.32, $p < .01$) or minor/intermediate (119.17, SD = 23.07, $p < .001$) surgery. The mean difference between major and minor/intermediate was not statistically significant. Female gender was negatively correlated with pre-test ($r = -0.14$, $p < .05$) and post-test ($r = -0.21$, $p < 0.01$) but not with change score. Age was correlated with posttest score ($r = 0.14$, $p < 0.05$) but not with change score.

3.8 | Responsiveness to change

A substantial decrease of 21.5 (mean) in total QoR-15 score from baseline to after surgery was observed (Cohen's $d = -1.02$, $p < .001$; Table 3). The effect sizes expressed as Cohen's d for the individual items ranges from 0.06 to 1.67. Except for two items

TABLE 3 Change score and effect sizes for single items and the total scale score of the QoR-15NO questionnaire from before to after surgery in 268 patients (range 211 to 250 on specific items/total score).

QoR-15 NO item	Pre-mean	Post-mean	Change score	Cohen's *d*	SRM	sig
1. Able to breathe easy ($n = 249$)	9.21	8.27	−0.95	−0.62	−0.40	***
2. Been able to enjoy food ($n = 250$)	8.39	6.14	−2.25	−0.92	−0.66	***
3. Feeling rested ($n = 249$)	7.57	5.82	−1.76	−0.75	−0.57	***
4. Have had a good sleep ($n = 250$)	7.63	5.83	−1.80	−0.74	−0.55	***
5. Able to look after personal toilet and hygiene unaided ($n = 250$)	9.70	7.80	−1.90	−1.55	−0.61	***
6. Able to communicate with family or friends ($n = 250$)	9.82	9.13	−0.69	−0.88	−0.35	***
7. Getting support from hospital doctors and nurses ($n = 234$)	8.45	9.51	1.06	0.33	0.33	***
8. Able to return to work or usual home activities ($n = 237$)	8.32	3.58	−4.73	−1.67	−1.14	***
9. Feeling comfortable and in control ($n = 247$)	8.45	6.48	−1.97	−0.87	−0.64	***
10. Having a feeling of general well-being ($n = 247$)	7.79	5.76	−2.03	−0.80	−0.64	***
11. Moderate pain ($n = 240$)	7.81	5.28	−2.53	−0.88	−0.78	***
12. Severe pain ($n = 241$)	8.63	6.68	−1.94	−0.76	−0.58	***
13. Nausea or vomiting ($n = 240$)	8.85	7.51	−1.34	−0.52	−0.40	***
14. Feeling worried or anxious ($n = 243$)	7.45	7.64	0.19	0.06	0.06	
15. Feeling sad or depressed ($n = 241$)	8.36	8.19	−0.17	−0.07	−0.06	
Total QOR-15 scale score ($n = 211$)	126.64	105.18	−21.45	−1.02	−0.53	***

Abbreviations: SRM, standardized response mean; QoR, quality of recovery.
***$p < .001$.

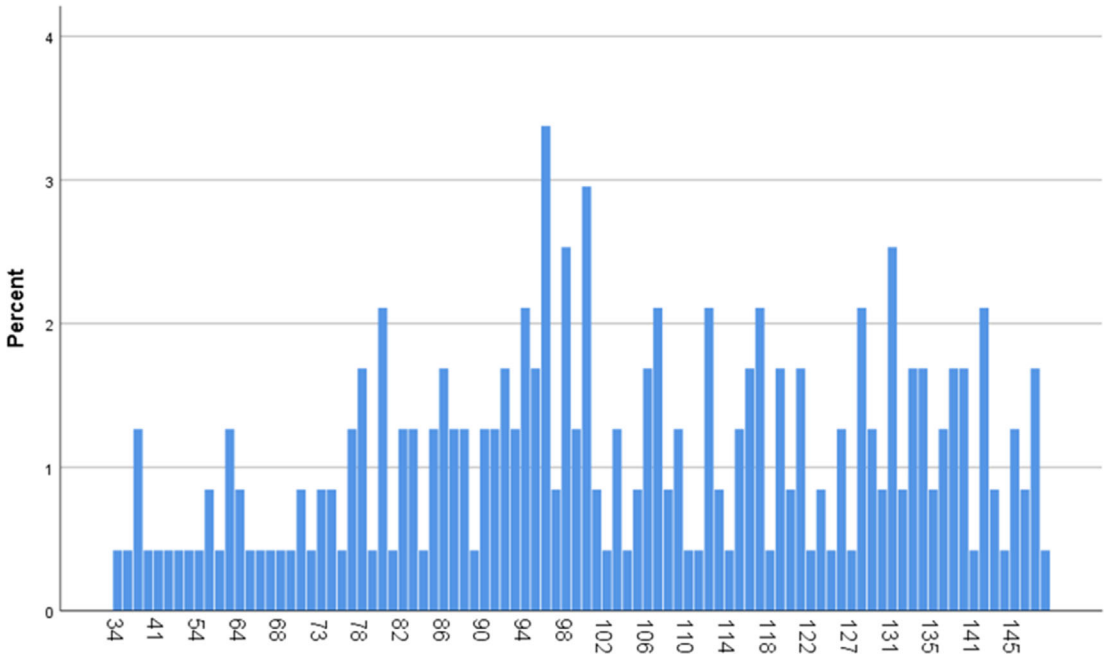

FIGURE 2 Distribution of QoR-15NO postoperative score ($N = 268$).

(q14 [anxious] and q15 [depressed]), the items significantly changed from pre- to post-assessment ($p < .001$). Among these, the mean score decreased except for q7 (support from hospital doctors and nurses) that increased over time.

As minimum total score is 0 and maximum is 150, there was no indication of either a ceiling or a floor effect as these scores were clearly lower than the cut off of 15%. The first and second day after surgery total scores ranged from 34 to 149 (Figure 2).

4 | DISCUSSION

This study has translated and validated the QoR-15 PROMS tool in a Norwegian clinical context. The Norwegian version of the QoR-15 has preserved the validity, moderate to very high reliability, the high degree of responsiveness and the clinical feasibility of the original English version. To support the validity of the QoR-15 NO, all the main hypotheses for construct validity were proven.

Regarding construct validity, the postoperative QoR-15NO total score and change score were strongly associated with the duration of surgery. As known from several meta studies,[3,5] this study reproduced a good capacity of QoR-15NO to discriminate between patients undergoing xmajor/complex versus major and minor/intermediate surgery. The mean difference between major and minor/intermediate-did not reach statistically significance, probably caused by the small number of patients in minor/intermediate group. The findings are all concurrent with the original paper and later metanalyses regarding psychometric evaluation.

Women had significantly lower pre- and postoperative QoR-15 score than men but had the same change score. This is concurrent with earlier studies.[9,24] The present study found that the high age correlated with higher postoperative QoR-15 score, but not with change score. Older people have more comorbidity and physical complaints, and a decrease in QoR-15 score could be expected. An earlier study reported no negative correlation with age[5] and points to the possibility that older people generally tend to underreport their health status and overscore their satisfaction with health care. The present study has similar findings which suggest that the instrument is less responsive in older patients.

Cronbach's alpha and Omega alpha total indicated very good internal consistency, with Cronbach's alpha being higher than previous studies.[3] Test–retest overall reliability was only moderate. Additionally, the values for the SEM and SDC were larger compared to previous studies.[5] This is concerning as the SDC is larger than the minimal clinical important difference (MCID) value of 6,[4] indicating that it is a considerable chance that a MCID could be caused by measurement error. We find it plausible, however, that the moderate ICC and large SEM and SDC values in the present study may be due to the large time gap (median 21 days) between the assessments. Other studies have typically done the test/retest postoperatively with a relatively short time between the two responses. In this study's set-up test–retest patients were included in the outpatient clinic where they gave their first score. The retest was done immediately before surgery. Furthermore, the moderate reliability could also be explained by the fact that patients were in a very different physical environment and in a different psychological state when doing the retest, and this might impact their answers. When baseline QoR-15 should be performed is rarely addressed in earlier studies. The instrument has earlier proven ability to detect clinically important changes in quality of recovery and it is likely that several of the questions could be answered differently at different time points preoperatively. Further research should explore these issues. The Norwegian version of the QoR-15 has still preserved a moderate to high degree of reliability depending on whether test–retest or internal consistency were used to assess it.

The present study found a large change in QoR-15NO from baseline to after surgery, which indicates a meaningful change in total score as defined by Myles.[4] Except for two items (q14 [anxious] and q15 [depressed]), the items significantly changed from pre- to post-assessment. The QoR-15NO total score were evenly distributed and further indicates that the well documented ability of this responsive instrument to measure and detect changes in quality of recovery has been maintained.

This study supports earlier research[5] which concludes that the instrument is essentially unidimensional as 80 percentage of the variance was explained by a strong general factor. It was in fact nearly identical to what Kleif and colleague found when using a bifactor model to analyze the synthesized QoR-15 data from three studies (81 percentage). This underlines the applicability of the total score of QoR-15. However, the scale also seems to measure some smaller domains reflected by the correlated residuals found in the dimensional analyses in both the present study and in Kleif et al.[5] It is probably wise in both clinical practice and research to focus on the specific items/domains in addition to the total score. An example of this is the two items assessing anxiety (q14) and depression (q15). These two items had a large correlation between their residuals after accounting for the variance explained by the general factor, had low loadings on this general factor (<0.35) and were the two items of which did not have statistically significant change score. Other studies show that the change score of these two items were either not significant or was in the opposite direction of the change score of the scale ([9,24–26]). Future research should explore when it is most beneficial to use the total score versus domains or even single items of this scale. The results show a high degree of feasibility and 83% who fulfilled the inclusion criteria agreed to complete QoR-15NO.

The clinical feasibility and high proportion of patients returning postoperative schemes show that it was easily implemented with a slightly active follow-up by hospital staff. The population underwent more complex surgery than earlier reported studies. Due to the Covid-19 pandemic and the related reduced capacity of surgery in Norwegian hospitals, enrolment of patients to the study took longer time than expected. The present study does not have full overview of dropouts in the inclusion period. The plan was to include a predetermined number of consecutive patients at each hospital. The study nurses had variable time schedules in the different hospitals and indirectly this caused a non-biased inclusion of patients. This type of inclusion instead of a real consecutive cohort is a methodological weakness, but the coherency between the present results and previous studies indicates no such bias issue. In this study the time gap between performing test and retest prior to surgery was too long.

In conclusion, the Norwegian version of QoR-15 appears valid with good reliability and a high degree of responsiveness. The instrument appears clinically feasible in a population of general, GI, and HPB patients with a high impact of surgery. Future studies should focus on timing of baseline data and gather several postoperative scores to monitor the patients to full recovery. QoR-15NO will now

be implemented as an integrated part of clinical practice and incorporated into the NORGAST.

AUTHOR CONTRIBUTION

KO, LB and TM conceived and designed the study. LB, TM, FP, SN, KH, TN, LN and KL all contributed to collection of data in their hospitals. KB performed the statistical analysis. LB, TM, KO and KB drafted the first version of manuscript. All authors discussed the results and worked on the final version and revised manuscript.

ACKNOWLEDGMENTS

We would like to thank Hartwig Kørner, Tom Mala, and Morten Tandberg Eriksen for their contribution to this manuscript. Geir Vegard Berg contributed with the study protocol and ethic approval. The first and second author has received a research funding from the Norwegian Medical Association's foundation for quality improvement and patient safety, and Innlandet Hospital Trust, Department of Research and Innovation. The authors have no conflict of interest.

ORCID

Thomas Moger https://orcid.org/0000-0003-1798-9639

REFERENCES

1. Lassen K, Nymo LS, Kørner H, et al. The New National Registry for Gastrointestinal Surgery in Norway: NoRGast. 2018.
2. Norsk register for gastrokirurgi (NoRGast). Nasjonalt servicemiljø for medisinske kvalitetsregistre. https://www.kvalitetsregistre.no/register/mage-og-tarm/norsk-register-gastrokirurgi-norgast
3. Myles PS, Shulman MA, Reilly J, Kasza J, Romero L. Measurement of quality of recovery after surgery using the 15-itemquality of recovery scale: a systematic review and meta-analysis. *Br J Anaesth*. 2022;128(6):1029-1039. doi:10.1016/j.bja.2022.03.009
4. Myles PS, Myles DB. An updated minimal clinically important difference for the QoR-15 scale. *Anesthesiology*. 2021;135(5):934-935. doi:10.1097/ALN.0000000000003977
5. Kleif J, Waage J, Christensen KB, Gögenur I. Systematic review of the QoR-15 score, a patient- reported outcome measure measuring quality of recovery after surgery and anaesthesia. *Br J Anaesth*. 2018;120(1):28-36. doi:10.1016/j.bja.2017.11.013
6. Gagnier JJ, Lai J, Mokkink LB, Terwee CB. COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures. *Qual Life Res*. 2021;30(8):2197-2218. doi:10.1007/s11136-021-02822-4
7. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;63(7):737-745. doi:10.1097/ALN.0000000000003977
8. Leonardsen A, Svensen K. Pasienterfaringer etter dagkirurgi—oversettelse og validering av spørreskjemaet Quality of Recovery-15 (QoR-15nor). *Sykepleien Forskning*. 2020;15(82819):e-82819. doi:10.4220/Sykepleienf.2020.82819
9. Stark PA, Myles PS, Burke JA. Development and psychometric evaluation of a postoperative quality of recovery score: the QoR-15. *Anesthesiology*. 2013;118(6):1332-1340. doi:10.1097/ALN.0b013e318289b84b
10. Wild D, Grove A, Martin M, et al. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR task force for translation and cultural adaptation. *Value Health*. 2005;8(2):94-104. doi:10.1111/j.1524-4733.2005.04054.x
11. Wong DJN, Harris S, Sahni A, et al. Developing and validating subjective and objective risk-assessment measures for predicting mortality after major surgery: an international prospective cohort study. *PLoS Med*. 2020;17(10):e1003253. doi:10.1371/journal.pmed.1003253
12. R Development Core Team. A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2020. Retrieved from https://www.R-project.org/
13. Rhemtulla M, Brosseau-Liard PÉ, Savalei V. When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychol Methods*. 2012;17(3):354-373. doi:10.1037/a0029315
14. Rosseel Y. Lavaan: an R package for structural equation modeling. *J Stat Softw*. 2012;48(2):1-36. http://www.jstatsoft.org/v48/i02/
15. Schafer JL, Graham JW. Missing data: our view of the state of the art. *Psychol Methods*. 2002;7:147-177. doi:10.1037/1082-989X.7.2.147
16. Brown TA. *Confirmatory Factor Analysis for Applied Research*. 2nd ed. Guilford Press; 2015.
17. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Modeling*. 1999;6(1):1-55.
18. Revelle W. (2014). psych: Procedures for Personality and Psychological Research. North- western University, Evanston. R package version 1.4.1
19. Jorgensen TD, Pornprasertmanit S, Schoemann AM, Rosseel Y. semTools: Useful Tools for Structural Equation Modeling. R Package Version 0. 2019 https://CRAN.R-project.org/package=semTools
20. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155-163.
21. Cohen J. *Statistical Power Analysis*. 2nd ed. Erlbaum; 1988.
22. Terwee CB, Bot SD, de Vet HC, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(1):34-42. doi:10.1016/j.jclinepi.2006.03.012.
23. ASA Physical Status Classification System. 2014 https://www.asahq.org/standards-and-guidelines/asa-physical-status-classification-system
24. Kleif J, Edwards HM, Sort R, Vilandt J, Gögenur I. Translation and validation of the Danish version of the postoperative quality of recovery score QoR-15. *Acta Anaesthesiol Scand*. 2015;59(7):912-920. doi:10.1111/aas.12525
25. Demumieux F, Ludes PO, Diemunsch P, et al. Validation of the translated quality of Recovery-15 questionnaire in a French-speaking population. *Br J Anaesth*. 2020;124(6):761-767.
26. Lyckner S, Böregård IL, Zetterlund EL, Chew MS. Validation of the Swedish version of quality of recovery score – 15: a multicentre, cohort study. *Acta Anaesthesiol Scand*. 2018;62(7):893-902. doi:10.1111/aas.13086

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: Bergestuen L, Moger T, Oterhals K, et al. Translation and validation of the Norwegian version of the postoperative quality of recovery score QoR-15. *Acta Anaesthesiol Scand*. 2023;1-8. doi:10.1111/aas.14322