


# Heterogeneity in Quality of Life of Long-Term Colon Cancer Survivors: A Latent Class Analysis of the Population-Based PROFILES Registry

FELIX J. CLOUTH <sup>a,d</sup> ARTURO MONCADA-TORRES <sup>d</sup> GIJS GELEIJNSE <sup>d</sup> FLOORTJE MOLS <sup>b,d</sup> FELICE N. VAN ERNING,<sup>d</sup>  
IGNACE H.J.T. DE HINGH <sup>e,f</sup> STEFFEN C. PAUWS,<sup>c</sup> LONNEKE V. VAN DE POLL-FRANSE <sup>b,d,g</sup> JEROEN K. VERMUNT <sup>a</sup>

Departments of <sup>a</sup>Methodology and Statistics, <sup>b</sup>Medical and Clinical Psychology, and <sup>c</sup>Communication and Cognition, Tilburg University, Tilburg, The Netherlands; <sup>d</sup>Netherlands Comprehensive Cancer Organisation (IKNL), Utrecht, The Netherlands; <sup>e</sup>Department of Surgery, Catharina Hospital, Eindhoven, The Netherlands; <sup>f</sup>Department of Epidemiology, GROW School for Oncology and Developmental Biology, Maastricht University, The Netherlands; <sup>g</sup>Division of Psychosocial Research and Epidemiology, The Netherlands Cancer Institute, Amsterdam, The Netherlands

Disclosures of potential conflicts of interest may be found at the end of this article.

**Key Words.** Clustering • Latent class analysis • Long-term survivors • Netherlands Cancer Registry • Patient-reported outcomes measures • PROFILES registry

## ABSTRACT

**Background.** Long-term colon cancer survivors present heterogeneous health-related quality of life (HRQOL) outcomes. We determined unobserved subgroups (classes) of survivors with similar HRQOL patterns and investigated their stability over time and the association of clinical covariates with these classes.

**Materials and Methods.** Data from the population-based PROFILES registry were used. Included were survivors with nonmetastatic (TNM stage I–III) colon cancer ( $n = 1,489$ ). HRQOL was assessed with the Dutch translation of the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire C30 version 3.0. Based on survivors' HRQOL, latent class analysis (LCA) was used to identify unobserved classes of survivors. Moreover, latent transition analysis (LTA) was used to investigate changes in class membership over time. Furthermore, the effect of

covariates on class membership was assessed using multinomial logistic regression.

**Results.** LCA identified five classes at baseline: class 1, excellent HRQOL ( $n = 555$ , 37.3%); class 2, good HRQOL with prevalence of insomnia ( $n = 464$ , 31.2%); class 3, moderate HRQOL with prevalence of fatigue ( $n = 213$ , 14.3%); class 4, good HRQOL with physical limitations ( $n = 134$ , 9.0%); and class 5, poor HRQOL ( $n = 123$ , 8.3%). All classes were stable with high self-transition probabilities. Longer time since the diagnosis, no comorbid conditions, and male sex were associated with class 1, whereas older age was associated with class 4. Clinical covariates were not associated with class membership.

**Conclusion.** The identified classes are characterized by distinct patterns of HRQOL and can support patient-centered care. LCA and LTA are powerful tools for investigating HRQOL in cancer survivors. *The Oncologist* 2021;26:e492–e499

**Implications for Practice:** Long-term colon cancer survivors show great heterogeneity in their health-related quality of life. This study identified five distinct clusters of survivors with similar patterns of health-related quality of life and showed that these clusters remain stable over time. It was also shown that these clusters do not significantly differ in tumor characteristics or received treatment. Cluster membership of long-term survivors can be identified by sociodemographic characteristics but is not predetermined by diagnosis and treatment.

Correspondence: Felix J. Clouth, M.Sc., Tilburg University, Department Methodology and Statistics, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. Telephone: +31134663687; e-mail: f.j.clouth@tilburguniversity.edu. Received April 23, 2020; accepted for publication December 9, 2020; published Online First on January 11, 2021. <http://dx.doi.org/10.1002/onco.13655>

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

## INTRODUCTION

Many cancer survivors struggle with long-term symptoms or diminished physical or psychosocial functioning that negatively affects their health-related quality of life (HRQOL). For instance, colon cancer survivors can experience nerve pain and fatigue that affect their psychological health and social roles during everyday life, primarily because of persistent side effects of the treatment [1, 2]. Cancer survivorship research has made substantial progress, specifically on the immediate and late effects of cancer and its treatment [1, 3, 4]. Long-term follow-up and registry studies on health status and HRQOL based on an individual's self-reported well-being (e.g., the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire [EORTC QLQ] C30) [5] have collected a vast amount of data to understand the outcomes and risks of new and existing treatments [6–8]. This is of crucial importance because the number of patients with a cancer diagnosis who will survive for 10 years or more is growing because of an aging population and ongoing advances in screening, diagnosis, and treatment [9, 10]. Understanding HRQOL is pivotal to health care professionals and patients in deciding on treatments at the time of diagnosis.

HRQOL in patients with and survivors of colon cancer is well studied. However, the majority of studies have only investigated specific aspects of HRQOL (such as chemotherapy-induced neuropathy [2, 11] and fatigue [12]) or modeled the subdomains of common HRQOL instruments separately by testing differences in means or using regression models [13–17]. Other studies have combined these subdomains into a single HRQOL score [18]. However, using such methods is problematic as cancer survivors present vast heterogeneity in their HRQOL [19, 20]. For example, some survivors might maintain optimal physical functioning after treatment but suffer from declined emotional well-being, whereas other survivors might experience fatigue or dyspnea. This heterogeneity is not reflected in an overall HRQOL score and cannot be adequately modeled when investigating subdomains separately, as information about the relation of these scales will be lost [21]. Pioneering work on this topic has been done by Kenzik et al. [19] and Pinheiro et al. [20] using latent class analysis (LCA) [22, 23] to investigate heterogeneity in HRQOL of lung and breast cancer survivors. This clustering approach allows modeling HRQOL by identifying unobservable classes of survivors with similar HRQOL patterns [23]. Because many aspects of HRQOL are long lasting or develop years after the treatment [7, 11], investigating heterogeneity in HRQOL for long-term survivors is particularly important. However, to what extent heterogeneity influences our understanding of HRQOL in long-term colon cancer survivors remains unknown. For instance, symptoms caused by treatment might be mitigated after many years, whereas limitations in functioning might still be prevalent or even increase with aging.

The aim of this study was to investigate heterogeneity and stability in HRQOL for a cohort of long-term colon cancer survivors. We used three-step LCA [24] (an extension of traditional LCA) to obtain unbiased estimates of the effects

of covariates on the classification into the latent HRQOL classes. Furthermore, we used latent transition analysis (LTA) [23, 25] to investigate if survivors transitioned between classes at follow-up.

## MATERIALS AND METHODS

### Data

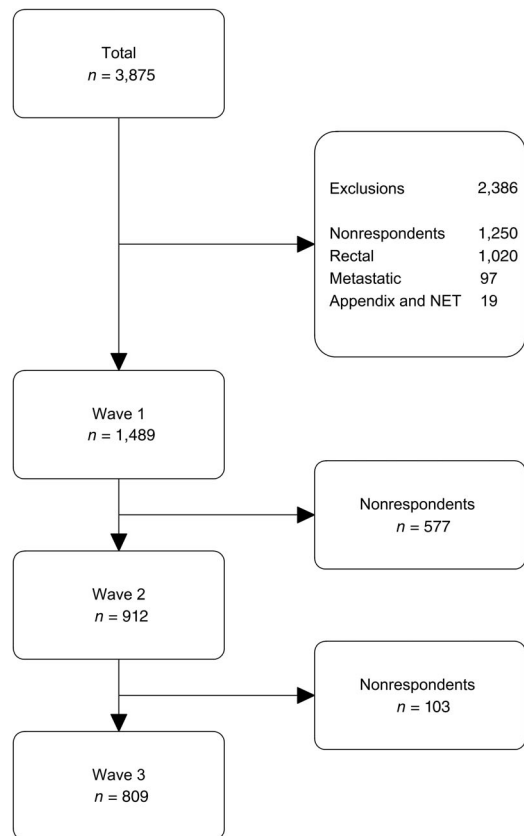
Data collection was performed within the PROFILES (Patient-Reported Outcomes Following Initial Treatment and Long-Term Evaluation of Survivorship) Registry [6]. This large dynamic population-based registry was set up to investigate the physical and psychosocial impact of cancer and its treatment and is continuously updated. Currently, it contains patient-reported outcomes data of over 25,000 cancer survivors with various tumors. PROFILES data are directly linked to the population-based Netherlands Cancer Registry (NCR) [26]. The NCR routinely collects information of all patients in The Netherlands, such as date of diagnosis, tumor characteristics, clinical stage, and treatment.

For this study, data of 1,489 survivors were used from the 2010 PROFILES colorectal cancer survey. Data collection was described in more detail previously [2]. In summary, the first wave of data collection started in December 2010. Additionally, there were three follow-up waves in yearly intervals, of which two were used for this analysis. Initially, 3,875 survivors diagnosed with colorectal cancer between 2000 and 2009 (as recorded in the NCR) from the southern region of The Netherlands were invited for participation. Survivors who already participated in an earlier study and who had an unverifiable address, cognitive impairment, or died before the start of the study were excluded [2]. Furthermore, nonrespondents ( $n = 1,250$ ), survivors diagnosed with rectal cancer ( $n = 1,020$ ), with a neuroendocrine tumor or a tumor in the appendix ( $n = 19$ ), and those not treated with curative intent (i.e., survivors with metastatic cancer [ $n = 97$ ]) were excluded. This resulted in a final sample of 1,489 colon cancer survivors at the first measurement occasion (Fig. 1).

### Measures

#### *Sociodemographic and Clinical Characteristics*

Survivors' age, sex, and clinical information about the diagnosis were obtained from the NCR. A survivor's vital status was obtained from the civil municipality registers. Comorbidities at time of the study were assessed with the adapted Self-Administered Comorbidity Questionnaire [27]. These include heart condition, stroke, high blood pressure, chronic obstructive pulmonary disease, diabetes, peptic ulcer disease, kidney disease, liver disease, anemia, thyroid disease, depression, arthritis, backache, and rheumatism. This information was updated at each follow-up wave. Socioeconomic status was determined by an indicator developed by Statistics Netherlands, which is based on aggregated individual fiscal data on the economic value of the home and household income. This indicator was estimated based on a sample of, on average, 17 households for each postal code [28]. Socioeconomic status for the



**Figure 1.** Inclusion criteria of the data used in the present study.

Abbreviation: NET, neuroendocrine tumor.

survivors in this study was determined by linking this information to the NCR based on the postal code at the time of diagnosis.

### HRQOL

HRQOL was assessed with the Dutch translation of the EORTC QLQ-C30 (version 3.0) [29]. This questionnaire contains five functional scales on physical, role, cognitive, emotional, and social functioning; a global health status/quality of life scale; three symptoms scales on fatigue, nausea and vomiting, and pain; and six single items assessing dyspnea, insomnia, loss of appetite, constipation, diarrhea, and financial impact. Most of the functioning and symptom scales consist of two items that range from 1 (not at all) to 4 (very much), except for the global quality of life scale, which ranges from 1 (very poor) to 7 (excellent) [30]. For the estimation of the LCA and LTA models, all scales were treated as ordinal measures. For comparison with previous literature, results of the LCA were then mapped on a scale with a range of 0 to 100.

### Statistical Analysis

Sociodemographic and clinical sample characteristics were quantified using descriptive statistics. LCA was used to identify clusters of survivors with similar response patterns on the EORTC QLQ-C30 scales and single items. LCA is a model-based unsupervised clustering technique that classifies individuals into unobserved classes based on probability. That is, for each survivor, a posterior class membership

probability is estimated based on the individual response pattern at the first time point and the class for which this probability is highest is assigned. Furthermore, a cross-sectional multinomial logistic (MNL) regression model was

**Table 1.** Demographic and clinical characteristics at baseline

Characteristic	n (%)
Total	1,489
Age	
Mean $\pm$ SD, years	70.5 $\pm$ 9.2
Missing	0 (0)
Sex	
Male	778 (52)
Female	709 (48)
Missing	2 (0)
Socioeconomic status	
Low	321 (21)
Medium	589 (39)
High	518 (35)
Missing	61 (5)
Time since diagnosis	
Mean $\pm$ SD, years	5.1 $\pm$ 2.9
Missing	0 (0)
Number of comorbidities	
None	329 (22)
1	394 (27)
> 1	678 (46)
Missing	76 (5)
TNM stage	
I	365 (24)
II	664 (45)
III	460 (31)
Missing	0 (0)
Differentiation grade	
Poorly differentiated	210 (14)
Moderately differentiated	986 (66)
Well differentiated	160 (11)
Missing	133 (9)
Topography	
Distal	812 (55)
Proximal	677 (45)
Missing	0 (0)
Histology	
Adenocarcinoma	1,287 (87)
Mucinous	198 (13)
Signet ring cell	4 (0)
Missing	0 (0)
Treatment	
Surgery only	1,032 (69)
Surgery and chemotherapy	457 (31)
Missing	0 (0)

**Table 2.** Goodness-of-fit indices for one-to-seven-class models at baseline

No. of classes	No. of parameters	LL	BIC	AIC	AIC3	CAIC
1	45	-16,727.8	33,784.4	33,545.6	33,590.6	33,829.4
2	61	-14,483.5	29,412.6	29,088.9	29,149.9	29,473.6
3	77	-14,097.5	28,757.6	28,349.0	28,426.0	28,834.6
4	93	-13,980.8	28,641.0	28,147.5	28,240.5	28,734.0
<b>5</b>	<b>109</b>	<b>-13,903.3</b>	<b>28,602.9</b>	<b>28,024.5</b>	<b>28,133.5</b>	<b>28,711.9</b>
6	125	-13,858.5	28,630.2	27,966.9	28,091.9	28,755.2
7	141	-13,820.2	28,670.5	27,922.4	28,063.4	28,811.5

The selected model is in bold.

Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion; CAIC, consistent Akaike information criterion; LL, log-likelihood.

used to assess the effect of covariates such as socio-demographic and clinical patient characteristics on the class membership. For this step, the three-step correction procedure proposed by Vermunt [24] was used to account for uncertainty in the class assignment and to prevent biased MNL regression estimates. The MNL regression model was restricted to the first time point to prevent a loss in statistical power because of dropouts at follow-up. To investigate dynamic changes in class membership at follow-up, transition probabilities were estimated using LTA. Because of the beforementioned dropout of participants at follow-up (Fig. 1), the measurement model of the LTA was estimated on data from the first measurement occasion only. The classification rule obtained from this model was then applied to the observed response patterns at follow-up where classifications were considered missing values in subsequent analyses for survivors who dropped out. Based on these classifications, the transition probabilities over the three measurement occasions were estimated. LTA is particularly suited for this situation as missingness at random can be

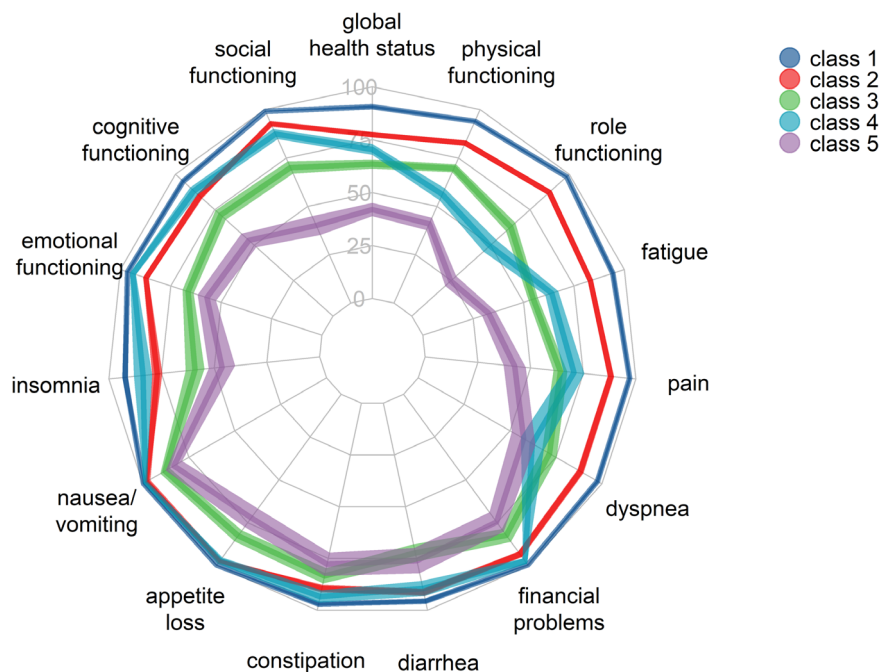
assumed. A detailed description of three-step LCA and LTA is provided as supplemental online Appendix 1.

To determine the statistically optimal number of classes, the number of classes was increased stepwise, and model fit was evaluated using the Bayesian information criterion, Akaike information criterion (AIC), AIC3, consistent AIC, and bivariate residuals. The model with the number of classes that minimized these information criteria was selected using multiple starting values to avoid local minima. Analyses were performed in R (version 3.6.0) [31] and Latent GOLD 5.1 [32]. Model syntax for R and Latent GOLD are available on GitHub [33].

## RESULTS

### Sociodemographic and Clinical Characteristics

In total, 1,489 survivors completed the questionnaire in 2010, 902 (61%) survivors in 2011, and 794 (53%) survivors in 2012 (Fig. 1). Descriptive statistics of this cohort are presented in Table 1.



**Figure 2.** Means and confidence intervals of all five health-related quality of life classes for the 15 European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire C30 dimensions. The symptom and single item scales were reversed.

**HRQOL**

For the first measurement occasion, a five-class solution was identified (Table 2).

The classes were characterized as follows.

Class 1: Excellent HRQOL (*n* = 555, 37.3%); very high probabilities for the highest scores on all 15 EORTC QLQ-C30 dimensions. This class exceeds average HRQOL scores of the Dutch normative population matched for age and gender [34].

Class 2: Good HRQOL with prevalence of insomnia (*n* = 471, 31.6%); very high probabilities for high scores on the EORTC QLQ-C30 symptom scales and slightly lower scores on the functioning scales. Survivors of this class are more likely to experience insomnia than survivors with comparable overall HRQOL.

Class 3: Moderate HRQOL with limitations in the functioning scales and fatigue (*n* = 209, 14.0%); high probabilities for moderate to good scores on most symptom scales and high probabilities for moderate scores on the functioning scales with prevalence of fatigue.

Class 4: Good HRQOL with physical limitations (*n* = 133, 8.9%); very high probabilities for the highest scores on emotional, cognitive, social functioning, and most symptom items but significantly lower scores on physical functioning, role functioning, fatigue, pain, and dyspnea.

Class 5: Poor HRQOL with severe limitations (*n* = 121, 8.1%); high probabilities for moderate scores on most symptom scales and low scores on the functioning scales. High probabilities for severe limitations in role functioning, social functioning, and fatigue.

Figure 2 shows the class-specific means and their confidence intervals for each dimension of the EORTC QLQ-C30 measure. Table 3 presents the exact values for these means with reference values for an age-matched norm-population [34] and indicates clinically relevant differences [35].

Male survivors were more likely than female survivors to be classified into class 1. Furthermore, the probability of being classified into class 1 increased with time passed between receiving the diagnosis and completing the questionnaire, whereas older survivors were more likely to be classified in class 4. A strong effect can be observed for the prevalence of comorbid conditions: with more than one comorbidity, the probability of poorer HRQOL outcome classes increases substantially. Clinical tumor characteristics and the indicator for receiving chemotherapy did not yield significant effects on class membership (Table 4).

Results from the LTA indicate that the probabilities of staying in one class over the three measurement occasions are high with 86.9% for class 1, greater than 75% for classes 2 and 5, and greater than 65% for classes 3 and 4 (Fig. 3). Further survivors were more likely to transition into a better HRQOL outcome class than into a worse HRQOL outcome class.

**DISCUSSION**

We identified five distinct classes of HRQOL for long-term colon cancer survivors. The first class was homogeneously characterized by an excellent HRQOL that even exceeded average HRQOL in the normal population (matched for age and sex). The narrow confidence intervals around the means (Fig. 2) show that almost all survivors in this class

**Table 3.** Means and SDs of all five health-related quality of life classes for the 15 European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire C30 dimensions. Presented are mean values for the age-matched Dutch population [34] and values for clinically meaningful differences [35]

Dimension	Mean ± SD					Norm-population [34], mean	Clinically meaningful differences [35]			
	Class 1	Class 2	Class 3	Class 4	Class 5		Trivial	Small	Medium	Large
Global health status/quality of life	90.6 ± 8.5	77.4 ± 12.4	63.5 ± 16.2	70.4 ± 14.2	42.0 ± 17.4	79	0-4	4-10	10-15	>15
Physical functioning	94.0 ± 7.5	82.7 ± 13.8	69.7 ± 15.2	56.2 ± 17.1	40.9 ± 18.0	89	0-5	5-14	14-22	>22
Role functioning	98.3 ± 5.8	87.3 ± 15.5	62.8 ± 22.1	49.6 ± 23.0	24.7 ± 19.5	88	0-6	6-19	19-29	>29
Emotional functioning	96.5 ± 7.2	87.3 ± 14.2	66.4 ± 20.2	93.7 ± 8.3	57.1 ± 25.9	90	0-3	3-7	7-10	>10
Cognitive functioning	95.1 ± 8.3	84.6 ± 17.2	71.3 ± 24.3	88.2 ± 13.6	53.4 ± 26.5	92	0-3	3-9	9-14	>14
Social functioning	99.2 ± 3.7	92.7 ± 12.4	70.1 ± 23.2	87.4 ± 14.9	40.1 ± 27.2	95	0-5	5-11	11-15	>15
Fatigue	5.7 ± 8.6	16.9 ± 13.2	45.2 ± 16.6	36.0 ± 17.8	67.1 ± 19.3	5.0	0-5	5-13	13-19	>19
Nausea/vomiting	0.3 ± 2.3	2.5 ± 8.1	12.1 ± 17.9	1.4 ± 5.1	17.0 ± 25.1	1.8	0-3	3-8	8-15	>15
Pain	3.0 ± 7.9	11.9 ± 16.1	35.8 ± 24.1	29.4 ± 27.3	57.5 ± 29.7	18	0-6	6-13	13-19	>19
Dyspnea	2.4 ± 8.8	11.7 ± 18.2	28.5 ± 29.1	37.9 ± 34.4	42.7 ± 35.9	7.4	0-4	4-9	9-15	>15
Insomnia	7.8 ± 16.3	23.0 ± 26.6	40.9 ± 33.1	16.2 ± 23.8	53.7 ± 35.1	5.0	0-4	4-13	13-24	>24
Appetite loss	0.2 ± 2.4	2.5 ± 9.6	17.3 ± 23.0	2.8 ± 10.3	27.8 ± 29.9	3.0	0-5	5-14	14-23	>23
Constipation	2.9 ± 9.9	10.4 ± 19.9	16.6 ± 27.1	6.6 ± 16.6	22.2 ± 31.6	7.7	0-5	5-13	13-19	>19
Diarrhea	4.4 ± 12.6	8.6 ± 19.2	27.7 ± 30.8	10.5 ± 20.3	23.9 ± 34.4	4.2	0-3	3-7	>7	—
Financial problems	0.4 ± 3.4	6.6 ± 17.3	17.5 ± 27.4	2.5 ± 8.9	25.4 ± 31.9	2.1	0-3	3-10	>10	—

**Table 4.** Odds ratios and 95% confidence intervals of factors associated with latent classes of health-related quality of life

Covariates	Class 2, OR (95% CI)	Class 3, OR (95% CI)	Class 4, OR (95% CI)	Class 5, OR (95% CI)
Intercept (ref = class 1)	0.33 (0.06–1.66)	0.38 (0.05–2.73)	0.00 (0.00–0.02) <sup>a</sup>	0.05 (0.00–0.61) <sup>a</sup>
Sex (ref = male): Female	1.63 (1.14–2.32) <sup>a</sup>	2.01 (1.31–3.10) <sup>a</sup>	1.48 (0.82–2.69)	1.86 (1.13–3.07) <sup>a</sup>
Age, years	1.00 (0.98–1.02)	0.98 (0.96–1.01)	1.07 (1.02–1.14) <sup>a</sup>	1.00 (0.96–1.03)
Socioeconomic status (ref = low)				
Medium	1.31 (0.80–2.13)	0.82 (0.48–1.39)	0.60 (0.29–1.26)	1.18 (0.62–2.25)
High	0.99 (0.60–1.64)	0.55 (0.31–0.98) <sup>a</sup>	0.95 (0.47–1.92)	0.96 (0.49–1.89)
No. comorbid conditions (ref = none)				
One	1.76 (1.09–2.85) <sup>a</sup>	1.45 (0.73–2.87)	3.75 (0.82–17.16)	3.09 (1.07–8.89) <sup>a</sup>
More than one	4.34 (2.67–7.05) <sup>a</sup>	7.66 (4.20–13.98) <sup>a</sup>	17.66 (4.27–73.11) <sup>a</sup>	20.67 (7.72–55.38) <sup>a</sup>
Time since diagnosis, years	0.92 (0.87–0.98) <sup>a</sup>	1.00 (0.92–1.07)	0.94 (0.85–1.04)	0.84 (0.77–0.93) <sup>a</sup>
TNM stage (ref = stage I)				
Stage II	1.08 (0.69–1.67)	1.20 (0.66–2.19)	0.98 (0.49–1.95)	0.82 (0.42–1.57)
Stage III	1.48 (0.76–2.91)	2.07 (0.84–5.13)	1.27 (0.44–3.64)	2.11 (0.86–5.19)
Differentiation grade (ref = moderately differentiated)				
Poorly differentiated	0.92 (0.51–1.64)	0.77 (0.39–1.52)	1.67 (0.51–5.41)	1.04 (0.45–2.37)
Well differentiated	0.81 (0.39–1.69)	0.91 (0.40–2.09)	1.59 (0.37–6.78)	1.41 (0.52–3.80)
Topography (ref = distal): Proximal	1.23 (0.84–1.79)	1.26 (0.80–1.98)	0.98 (0.56–1.73)	1.68 (0.98–2.86)
Histology (ref = adenocarcinoma): Mucinous	1.20 (0.69–2.08)	0.84 (0.42–1.68)	1.41 (0.60–3.28)	1.87 (0.96–3.62)
Treatment (ref = surgery only): Surgery and adjuvant chemotherapy	0.82 (0.45–1.50)	0.76 (0.33–1.72)	0.56 (0.20–1.52)	0.55 (0.25–1.23)

Odds ratios reported in this table are based on a multinomial logistic regression and therefore adjusted for all other covariates in the model. Covariates with missing cases (see Table 1) were imputed using mean imputation.

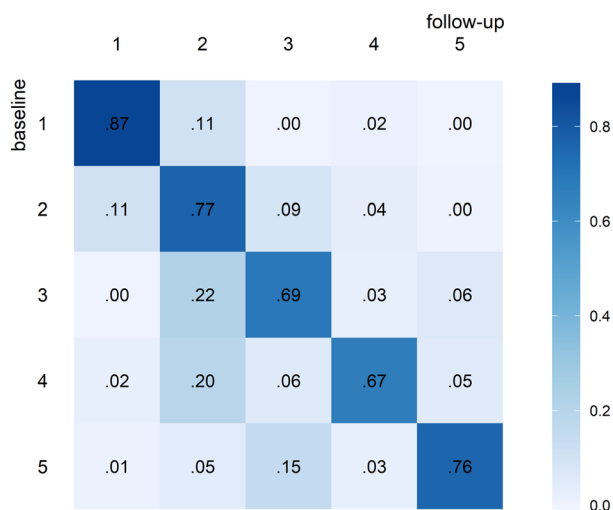
<sup>a</sup>Statistically significant effects.

Abbreviations: CI, confidence interval; OR, odds ratio.

answered with the highest score on all dimensions of the EORTC QLQ-C30. This class was also the most prevalent one (37%), which suggests an exceptional prognosis for HRQOL for a large group of patients who survive their disease. Survivors in this class likely had better than average HRQOL already before their diagnosis and a good prognosis for

survival, which could explain the high prevalence in this cohort. Additionally, response shift is a well-documented phenomenon in cancer survivorship research. Facing a disease such as cancer might shift a survivor's view on his or her own HRQOL. High HRQOL as found in this study might reflect to some extent changes in one's internal standards, changes in importance attributed to specific domains of HRQOL, or a general redefinition of HRQOL [36].

The composition of class 4 is highly interesting. Although survivors in this class maintain an excellent emotional and cognitive functioning, they experience severe limitations in their physical and role functioning, symptoms of fatigue, insomnia, pain, and dyspnea. Results from the MNL regression show that older age is a significant predictor for this class. The composition of class 4 also highlights the advantages of using LCA for analyzing HRQOL outcomes. A survivor group with such peculiar values of HRQOL would have been extremely difficult to identify by analyzing the dimensions of the EORTC QLQ-C30 separately. Moreover, LCA allows us to identify dimensions of HRQOL that are important for clinical practice. Generally, it can be observed that scores on the functioning scales are lower than on the symptom scales, indicating that the long-term effects of colon cancer relate to limitations in functioning rather than prevalence of symptoms. However, although symptoms of nausea, constipation, diarrhea, and appetite loss seem to be less relevant for long-term survivors, they



**Figure 3.** Transition probabilities between classes from baseline to follow-up. For example, in this case the probability of transitioning from class 3 at baseline to class 2 at follow-up is 0.22.

experience higher levels of fatigue, insomnia, pain, and dyspnea, as well as more limitations in role functioning for almost all classes. This is especially relevant because many of these symptoms (e.g., pain) can be successfully managed and should be screened for in subsequent care paths after initial treatment [37].

Results of the MNL regression and the LTA show that HRQOL after colon cancer increases with time. Probabilities of class membership decreased for classes with poorer HRQOL with increasing time between the diagnosis and filling out the questionnaire. Furthermore, although the probabilities of staying in the same class over the two follow-up years were high and especially high for class 1 with excellent HRQOL, probabilities for transitioning to better HRQOL classes were considerably higher than for transitioning to poorer HRQOL classes. The results from the MNL regression indicate survivorship bias in the data. However, LTA revealed that even for these long-term survivors HRQOL increases in some cases. Moreover, the MNL regression model did not identify clinical predictors for class membership: the type of tumor and receiving adjuvant chemotherapy had no effect on HRQOL as assessed with the EORTC QLQ-C30 in long-term colon cancer survivors. These results support the findings of Kenzik and colleagues [19], which indicate that long-term effects in HRQOL might not be of major concern when choosing treatment. Although it is important to note that the absence of long-term effects is conditional on long-term survival, it is valuable information at the time of decision making that colon cancer and its treatment do not affect long-term HRQOL. Furthermore, comorbid conditions are an important factor for determining HRQOL, and attention to them needs to have a prominent role in the follow-up care path of colon cancer survivors.

We believe that this study greatly contributes to the understanding of HRQOL in long-term colon cancer survivors. However, there are a couple of shortcomings that are worth discussing. Patients in our cohort survived on average 5 years before they filled in the questionnaire for the first time. Survivorship bias limits the generalizability of our results. Additionally, all effects of potential predictors of class membership are conditional on survival. For example, we do not know if there is a treatment effect on distinct HRQOL classes close to the treatment itself. Especially for patients with a poorer prognosis, this would be relevant information and therefore limits our findings. This type of analysis requires a relatively large number of observations. Although even for less prevalent classes we had adequate cell counts for all categories of the covariates, it is possible that weak effects on the class membership could not be detected.

Because the presented results are based on an observational study, we propose the use of LCA in randomized controlled trials for consecutive studies to further evaluate the effect of treatment on HRQOL in cancer survivors.

## REFERENCES

1. Wu HS, Harden JK. Symptom burden and quality of life in survivorship: A review of the literature. *Cancer Nurs* 2015;38:29–54.
2. Mols F, Beijers T, Lemmens V et al. Chemotherapy-induced neuropathy and its association with quality of life among 2-to 11-year colorectal cancer survivors: Results from the population-based PROFILES registry. *J Clin Oncol* 2013;31:2461–2470.
3. Bottomley A, Reijneveld JC, Koller M et al. Current state of quality of life and patient-reported outcomes research. *Eur J Cancer* 2019;121:55–63.
4. Jansen L, Herrmann A, Stegmaier C et al. Health-related quality of life during the 10 years after diagnosis of colorectal cancer: A population-based study. *J Clin Oncol* 2011;29:3263–3269.
5. Sprangers MAG, Cull A, Bjordal K et al. The European Organization for Research and

Additionally, the usability of causal inference techniques in the context of LCA applied to data such as presented in this study could be investigated.

## CONCLUSION

In this study, we showed that LCA is a powerful method to investigate HRQOL in cancer survivors taking into account the multidimensionality of the construct and the heterogeneity of the data. Our model allowed us to discriminate five classes of long-term colon cancer survivors, which compositions yielded a clinically meaningful explanation of HRQOL in this cohort. Furthermore, we used three-step LCA to correctly estimate the effect of covariates on class membership and showed the stability of our class solution over time using LTA. We found that tumor characteristics and receiving adjuvant chemotherapy had no effect on class membership in this population of long-term survivors. Although this poses a limitation for using this information in decision making, it also suggests that most survivors return to good or excellent HRQOL after surviving colon cancer.

## ACKNOWLEDGMENTS

This article draws on data of the PROFILES Registry. We would like to thank the registration team of the Netherlands Comprehensive Cancer Organization (IKNL) for the collection of the data for the Netherlands Cancer Registry. We would like to acknowledge The Netherlands Organisation for Scientific Research (NWO) for grant 628.001.030, “Helping cancer patients to choose the best treatment: Data-driven shared decision making on cancer treatment for individual patients.”

## AUTHOR CONTRIBUTIONS

**Conception/design:** Felix J. Clouth, Arturo Moncada-Torres, Gijs Geleijnse, Steffen C. Pauws, Lonneke V. van de Poll-Franse

**Provision of study material or patients:** Floortje Mols, Lonneke V. van de Poll-Franse

**Collection and/or assembly of data:** Floortje Mols, Lonneke V. van de Poll-Franse

**Data analysis and interpretation:** Felix J. Clouth, Floortje Mols, Felice N. van Erning, Ignace H.J.T. de Hingh, Lonneke V. van de Poll-Franse, Jeroen K. Vermunt

**Manuscript writing:** Felix J. Clouth, Arturo Moncada-Torres, Steffen C. Pauws, Jeroen K. Vermunt

**Final approval of manuscript:** Felix J. Clouth, Arturo Moncada-Torres, Gijs Geleijnse, Floortje Mols, Felice N. van Erning, Ignace H.J.T. de Hingh, Steffen C. Pauws, Lonneke V. van de Poll-Franse, Jeroen K. Vermunt

## DISCLOSURES

**Ignace H.J.T. de Hingh:** Roche, QP&S/RanD Biotech (RF—institution). The other authors indicated no financial relationships. (C/A) Consulting/advisory relationship; (RF) Research funding; (E) Employment; (ET) Expert testimony; (H) Honoraria received; (OI) Ownership interests; (IP) Intellectual property rights/inventor/patent holder; (SAB) Scientific advisory board

Treatment of Cancer approach to quality of life assessment: Guidelines for developing questionnaire modules. *Qual Life Res* 1993;2:287–295.

6. van de Poll-Franse L V, Horevoorts N, Eenbergen M Van et al. The Patient Reported Outcomes Following Initial Treatment and Long Term Evaluation of Survivorship Registry: Scope, rationale and design of an infrastructure for the study of physical and psychosocial outcomes in cancer survivorship cohorts. *Eur J Cancer* 2011; 47:2188–2194.
7. Arndt V, Koch-Gallenkamp L, Jansen L et al. Quality of life in long-term and very long-term cancer survivors versus population controls in Germany. *Acta Oncol* 2017;56:190–197.
8. Chambers SK, Meng X, Youl P et al. A five-year prospective study of quality of life after colorectal cancer. *Qual Life Res* 2012;9:1551–1564.
9. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA Cancer J Clin* 2019;69:7–34.
10. Miller KD, Nogueira L, Mariotto AB et al. Cancer treatment and survivorship statistics, 2019. *CA Cancer J Clin* 2019;69:363–385.
11. Bonhof CS, van de Poll-Franse L V, Vissers PAJ et al. Anxiety and depression mediate the association between chemotherapy-induced peripheral neuropathy and fatigue: Results from the population-based PROFILES registry. *Psychooncology* 2019;:1–8.
12. Thong MSY, Mols F, van de Poll-Franse LV et al. Identifying the subtypes of cancer-related fatigue: Results from the population-based PROFILES registry. *J Cancer Surviv* 2018;12:38–46.
13. Arndt V, Merx H, Stegmaier C et al. Quality of life in patients with colorectal cancer 1 year after diagnosis compared with the general population: A population-based study. *J Clin Oncol* 2004;22:4829–4836.
14. Thong MSY, Mols F, Lemmens VEPP et al. Impact of chemotherapy on health status and symptom burden of colon cancer survivors: A population-based study. *Eur J Cancer* 2011;47: 1798–1807.
15. Adams S V, Ceballos R, Newcomb PA. Quality of life and mortality of long-term colorectal cancer survivors in the Seattle Colorectal Cancer Family Registry. *PLoS One* 2016;11:e0156534.
16. Teker F, Demirag G, Erdem D et al. Quality of life in colorectal cancer patients during chemotherapy in the era of monoclonal antibody therapies. *J BUON* 2015;20:443–451.
17. Burton-Chase AM, Parker WM, Donato KM et al. Health-related quality of life in colorectal cancer survivors: Are there differences between sporadic and hereditary patients? *J Patient Rep Outcomes* 2017;2:21.
18. Ratjen I, Schafmayer C, Enderle J et al. Health-related quality of life in long-term survivors of colorectal cancer and its association with all-cause mortality: A German cohort study. *BMC Cancer* 2018;18:1156.
19. Kenzik KM, Martin MY, Fouad MN et al. Health-related quality of life in lung cancer survivors: Latent class and latent transition analysis. *Cancer* 2015;121:1520–1528.
20. Pinheiro LC, Tan X, Olshan AF et al. Examining health-related quality of life patterns in women with breast cancer. *Qual Life Res* 2017; 26:1733–1743.
21. Snyder CF, Garret-Mayer E, Blackford AL et al. Concordance of cancer patients' function, symptoms, and supportive care needs. *Qual Life Res* 2009;18:991–998.
22. McCutcheon AL. *Latent Class Analysis*. Beverly Hills, CA: Sage Publications, 1987.
23. Collins LM, Lanza ST. *Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences*. Hoboken, NJ: Wiley, 2010.
24. Vermunt JK. Latent class modeling with covariates: Two improved three-step approaches. *Polit Anal* 2010;18:450–469.
25. Bartolucci F, Montanari GE, Pandolfi S. Three-step estimation of latent Markov models with covariates. *Comput Stat Data Anal* 2015;83: 287–301.
26. Schouten LJ, Höppener P, van den Brandt P et al. Completeness of cancer registration in Limburg, The Netherlands. *Int J Epidemiol* 1993;22: 27. Sangha O, Stucki G, Liang MH et al. The Self-Administered Comorbidity Questionnaire: A new method to assess comorbidity for clinical and health services research. *Arthritis Rheum* 2003;49:156–163.
28. van Duijn C, Keij I. Sociaal-economische status indicator op postcode niveau. *Maandstat van Bevolk* 2002;50:32–35.
29. Niezgod HE, Pater JL. A validation study of the domains of the core EORTC Quality of Life Questionnaire. *Qual Life Res* 1993;2:319–325.
30. Fayers P, Aaronson NK, Bjordal K et al. The EORTC QLQ-C30 Scoring Manual. 3rd ed. Brussels, Belgium: European Organisation for Research and Treatment of Cancer, 2001.
31. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing, 2019. Available at <https://www.r-project.org/>.
32. Vermunt JK, Magidson J. *Upgrade Manual for Latent GOLD 5.1*. Belmont, MA: Statistical Innovations, 2016.
33. Clouth FJ. Manuscript LCA colon cancer. GitHub. 2020. doi:<https://doi.org/10.5281/zenodo.3698006>. [https://github.com/IKNL/manuscriptLCA\\_coloncancer](https://github.com/IKNL/manuscriptLCA_coloncancer)
34. van de Poll-Franse L V, Mols F, Gundy CM et al. Normative data for the EORTC QLQ-C30 and EORTC-sexuality items in the general Dutch population. *Eur J Cancer* 2011;47:667–675.
35. Cocks K, King MT, Velikova G et al. Evidence-based guidelines for determination of sample size and interpretation of the European Organisation for the Research and Treatment of Cancer Quality of Life Questionnaire Core 30. *J Clin Oncol* 2011;29:89–96.
36. Hagedoorn M, Sneeuw KCA, Aaronson NK. Changes in physical functioning and quality of life in patients with cancer: Response shift and relative evaluation of one's condition. *J Clin Epidemiol* 2002;55:176–183.
37. Mulvey MR, Boland EG, Bouhassira D et al. Neuropathic pain in cancer: Systematic review, performance of screening tools and analysis of symptom profiles. *Br J Anaesth* 2017;119:765–774.



See <http://www.TheOncologist.com> for supplemental material available online.