



Do patients suffering from chronic diseases retrospectively overestimate how healthy they were before they fell ill?



Andreas Hinz^{a,*}, Markus Zenger^{b,c}, Katja Leuteritz^a,
Anja Mehnert-Theuerkauf^a, Katja Petrowski^d

^a Department of Medical Psychology and Medical Sociology, University of Leipzig, Germany

^b Department of Applied Human Studies, University of Applied Sciences Magdeburg-Stendal, Germany

^c Integrated Research and Treatment Center Adiposity Diseases, University of Leipzig, Germany

^d Medical Psychology and Medical Sociology, Department of Psychosomatic Medicine and Psychotherapy, University of Mainz, Germany

Received 13 January 2021; accepted 8 March 2021

KEYWORDS

Self-rated health;
Response shift;
Retrospective
assessment;
Descriptive survey
study

Abstract *Background/Objective:* To examine the impact of a disease on a person's subjective health state, patients are often asked to assess their current health state and to retrospectively assess how healthy they were before they fell ill. The objective of this study was to test whether patients generally overestimated the quality of their pre-disease health. *Method:* Six samples of patients with chronic diseases (cancer patients, cardiovascular patients, and patients diagnosed with sarcoidosis, N between 197 and 1,197) were analyzed. The patients assessed their current health states and their health states at the time before diagnosis. The retrospective scores were compared with matched data from general population studies. *Results:* In three of the six studies, the retrospective health ratings of the patients were significantly higher than the general population norms (effect sizes between 0.24 and 0.46), two studies yielded non-significant effects, and in one study there was an opposite trend. The general overestimation of pre-disease health was more pronounced in older patients as compared with younger ones, and it was more pronounced when global health/quality of life was to be assessed. *Conclusions:* Retrospective assessments of pre-disease health states are not appropriate for assessing disease-related changes in a person's health state.

© 2021 Asociación Española de Psicología Conductual. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

* Corresponding author. Department of Medical Psychology and Medical Sociology, University of Leipzig, Philipp-Rosenthal-Str. 55, 04103 Leipzig, Germany.

E-mail address: andreas.hinz@medizin.uni-leipzig.de (A. Hinz).

PALABRAS CLAVE

Autoinforme de salud;
Cambios en la
respuesta;
Evaluación
retrospectiva;
Estudio descriptivo
mediante encuestas

Pacientes que sufren enfermedades crónicas: ¿sobrestiman qué tan saludables eran antes de enfermar?

Resumen *Antecedentes/Objetivo:* Cuando se quiere examinar el impacto de una enfermedad en la percepción de salud de una persona se pide a la misma evaluar su estado actual y retrospectivamente evaluar también como era su salud antes de enfermar. El objetivo de este estudio era comprobar si los pacientes generalmente sobrestiman la calidad de su salud pre-enfermedad. *Método:* Seis muestras de pacientes con enfermedades crónicas (pacientes con cáncer, enfermedad cardiovascular y sarcoidosis, n entre 197 y 1.197) fueron analizadas. Los pacientes evaluaron su estado de salud actual y su estado de salud antes del diagnóstico. Los puntajes retrospectivos fueron comparados con datos pareados de estudios con la población general. *Resultados:* En tres de los seis estudios los puntajes retrospectivos de salud de los pacientes fueron significativamente más altos que la norma poblacional (tamaños del efecto entre 0,24 y 0,26), dos estudios mostraron efectos no significativos y uno de los estudios mostró la tendencia opuesta. La sobrestimación de la salud fue más pronunciada en pacientes mayores comparados con pacientes jóvenes. *Conclusiones:* Las evaluaciones retrospectivas de estados de salud pre-enfermedad no son apropiadas para evaluar cambios relacionados con la enfermedad en el estado de salud de una persona.

© 2021 Asociación Española de Psicología Conductual. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Quality of Life (QoL) has become a relevant outcome criterion in clinical research and practice. When the effects of a disease or a treatment on QoL are examined, changes in QoL are relevant. However, pre-disease QoL data are generally not available. Moreover, simple difference scores (post-disease vs. pre-disease or post-treatment vs. pre-treatment) are valid only under the assumption that the frames of reference for these assessments did not change. However, multiple studies have shown that this is not the case, a phenomenon referred to as ‘‘response shift’’ (Lindberg et al., 2017; Sprangers & Schwartz, 1999; ten Ham et al., 2020). Several techniques have been developed to detect such response shift effects, the thentest method (Schwartz et al., 2006), structural equation modeling (Oort, 2005), and the approach using anchoring vignettes (Hinz et al., 2020).

When patients retrospectively assess how good their health was before they fell ill, they tend to overestimate their previous (disease-free) health state, due to such response shift effects. One of the first examinations of this effect was performed by Brickman et al. (1978) who examined lottery winners, accident victims, and controls. As was to be expected, the accident victims assessed their current happiness as being lower than the controls did, but when they retrospectively assessed how happy they had been pre-accident, the victims’ ratings were higher than the controls’ estimates. This was confirmed by a Dutch study that found that patients who had experienced an injury assessed their previous health state (pre-injury) as having been significantly better than Dutch normative values (Graaf et al., 2019). A systematic review, based on 31 examinations, proved that nearly all groups of patients generally tended to overestimate their pre-injury health states (Scholten et al., 2017).

While many studies have examined these effects among people who have experienced a sudden event such as an injury, those focused on patients suffering from chronic dis-

eases are rare. Moreover, it is unknown whether there are age and gender effects on overestimations of previous health states. Since the persons’ subjective assessments of their own health may diverge with the assessments of their physicians with increasing age (Schnittker, 2005), it is possible that the overestimation of previous health state increases with increasing age.

Global assessments of health and QoL are not identical with the average of assessments in the specific domains that constitute the concept of QoL. Several studies that compared oncological patients’ QoL with that of the general population found that the patients reported clear detriments in multiple functioning domains and heightened levels in multiple symptom variables, even though they nevertheless rated their global health and QoL as being relatively good (Arndt et al., 2005; Hinz et al., 2017). The above-mentioned studies indicating overestimations occurring in retrospective health assessments generally adopted such general health assessments. The question of whether such overestimations can also be observed when specific domains of QoL have to be judged has not been investigated so far. Our hypothesis is that overestimations of previous health are more pronounced when health is to be assessed on a global level as compared with specific aspects of QoL.

One further problem in the studies with retrospective assessments is that the sample sizes of these studies are often relatively low. The Brickman et al. (1978) study included 29 accident victims, and in the systematic review (Scholten et al., 2017), only 29% of the studies had sample sizes of 300 or above. For reliable estimates of retrospective judgment effects, studies with sufficiently large sample sizes are necessary. Therefore, in this study we concentrate on larger sample sizes.

The aims of this study were: (a) to test whether patients suffering from chronic diseases retrospectively overestimate how healthy they were pre-disease, (b) to investigate age

Table 1 Overview on the samples.

Sample	Patients	<i>n</i>	Age (years) <i>M</i> ± <i>SD</i>	Gender (% female)	Time since diagnosis	Reference
BR-CA	Breast cancer patients	323	66.20 ± 9.60	100%	7.4 yrs.	(Friedrich, Zenger, & Hinz, 2019)
MIXED-CA	Mixed cancer patients	428	60.60 ± 10.50	25.70%	167 d.	(Zenger et al., 2010; Zenger et al., 2011)
GYN-CA	Gynecologic cancer patients	354	61.20 ± 14.20	100%	127 d.	(Kecke et al., 2017)
URO-CA	Urologic cancer patients	197	63.10 ± 10	0%	38 d.	(Preiß et al., 2019)
CARDIO	Cardiologic rehabilitation	356	55.60 ± 8.10	16.30%	2.4 yrs.	(Friedrich, Karoff, & Hinz, 2019)
SARCOID	Sarcoidosis patients	1,197	54.30 ± 11.70	65.40%	12.8 yrs.	(Bosse-Henck et al., 2017)

Note. *M*: Mean; *SD*: Standard deviation; yrs.: years; d.: days.

and gender differences in this effect of retrospective overestimation, and (c) to examine whether the global health assessments are more strongly affected by such overestimations than the specific components of health and QoL.

Method

Patient samples

This study comprised six samples which had been examined with different objectives. In all of these samples, the patients also gave retrospective health judgments. Table 1 summarizes the most relevant features of the samples. Further descriptions of the samples have been published in the articles referenced in Table 1. All of the six studies were approved by the Ethics Committee of the University of Leipzig.

BR-CA: Breast cancer patients. This sample comprised of 308 women who took part in a routine radiologic after-treatment (breast cancer) examination (Friedrich, Zenger, & Hinz, 2019). Immediately after the radiologic examination, the participants were asked to fill in several questionnaires.

MIXED-CA: Mixed cancer patients. This sample was composed of 317 male patients with urologic cancer (Zenger et al., 2010) and 103 female patients with gynecological cancer (Zenger et al., 2011) treated in a German university hospital. In this analysis, we only use the data from the first measurement, obtained while the patients were hospitalized.

GYN-CA: Gynecologic cancer patients. Patients with gynecological or breast cancer ($N=354$) were consecutively recruited for this study in the gynecological clinics of three German hospitals. In most cases, the test was performed one or two days before hospital discharge. Further details are given elsewhere (Kecke et al., 2017).

URO-CA: Urologic cancer patients. This sample is composed of 197 male patients suffering from urologic cancer, in most cases prostate cancer. The patients were treated in a German university hospital. As in the GYN-CA sample, most patients received the questionnaires one or two days before discharge (Preiß et al., 2019).

CARDIO: Patients with cardiovascular diseases treated in a rehabilitation clinic. A group of cardiologic patients treated in a German rehabilitation clinic took part in this

study. The central inclusion criterion was survival of an acute cardiovascular event. A total of 348 patients consented to participate and filled in the questionnaire. Further details are described elsewhere (Friedrich, Karoff, & Hinz, 2019).

SARCOID: Patients diagnosed with sarcoidosis. Sarcoidosis is multi-system inflammatory disease, characterized by epitheloid granulomas (Starshinova et al., 2020). The members of the German Sarcoidosis Society were invited to take part in the survey, and 1,197 of them consented. Further details of the sample have been published elsewhere (Bosse-Henck et al., 2017).

Instruments

Health assessment. The questionnaire Fragebogen zur Lebenszufriedenheit - Module FLZ-M (Questions of Life Satisfaction) was developed to assess QoL and life satisfaction in eight dimensions, one of them being health. Participants from all of the samples responded to the health question (How satisfied are you with your health?) on a five-point rating scale from *dissatisfied* to *very satisfied*, coded with scores from 1 to 5 (Henrich & Herschbach, 2000). Comparison scores of the general population were taken from the normative study (Daig et al., 2009) which comprises 5,030 representatively selected persons from the German general population.

EORTC QLQ-C30. The quality of life questionnaire EORTC QLQ-C30 (Aronson et al., 1993) was also included in one of the six studies (BR-CA). It consists of 30 items and includes five functioning scales (physical, role, emotional, social, and cognitive), three symptom scales (Fatigue, Pain, and Nausea/vomiting), a global health status/QoL scale, and six single items. All scores range from 0 to 100. Higher functioning scores represent better functioning/QoL, whereas higher symptom scores represent more severe symptoms. A QoL sum score that aggregates across all of the functioning and symptom scales except for financial difficulties can be calculated (Giesinger et al., 2016). Normative values are available for several countries (Hinz et al., 2014; Nolte et al., 2019).

Statistical analysis

We calculated expected mean scores for each sample and each scale based on published normative studies for the FLZ-

Table 2 Mean health scores (range: 1-5) for patients and the general population.

Sample	Current (Patients)		Retrospective (Patients)		General population		<i>d</i> (Current, General population)		<i>d</i> (Retrospective, General population)	
	<i>M</i>	(<i>SD</i>)	<i>M</i>	(<i>SD</i>)	<i>M</i>	(<i>SD</i>)				
BR-CA	3.36	(1.07)	4.06	(0.91)	3.64	(0.92)	-0.28	***	0.46	***
MIXED-CA	3.07	(1.28)	3.89	(1.02)	3.66	(0.91)	-0.54	***	0.24	***
GYN-CA	2.88	(1.14)	3.68	(1.05)	3.62	(0.93)	-0.71	***	0.06	<i>ns</i>
URO-CA	3.04	(1.22)	3.84	(0.87)	3.63	(0.91)	-0.55	***	0.24	***
CARDIO	2.79	(1.13)	3.52	(0.96)	3.80	(0.90)	-1.00	***	-0.30	***
SARCOID	2.44	(1.03)	3.95	(0.98)	3.92	(0.88)	-1.55	***	0.03	<i>ns</i>

Note. *M*: mean; *SD*: standard deviation; *d*: effect size; ***: $p < .001$; *ns*: not significant.

M (Daig et al., 2009) and for the EORTC QLQ-C30 (Hinz et al., 2014) (data from Table 2 of that paper). These normative studies present mean scores stratified by age groups and gender, based on large representatively selected samples of the German general population. First, we calculated the pre-disease age of the participants as being the difference between their age when they completed the questionnaire and the time since diagnosis, since the retrospective health assessments referred to the pre-diagnosis time. For each of the samples, we then calculated the percentages of the (pre-diagnosis) age and gender groups according to the age stratification given in the normative studies. Then we calculated the weighted mean of the health-related variables in the six samples, based on the distribution of the age and gender groups. For the EORTC QLQ-C30, we used previously published data from the general population (Friedrich, Zenger, & Hinz, 2019). Effect sizes *d* were calculated by relating the mean score difference to the pooled standard deviations. Differences between the patients' mean scores (current values and retrospective values) and the means of the general population were tested with *t*-tests. All statistical calculations were performed with SPSS version 24.

Results

Assessments of the general health state

Table 2 presents the mean current health state and the retrospectively assessed health state (as measured with the FLZ-M) of the patient groups, together with the expected mean scores derived from the general population samples. These expected mean scores refer to age distributions according to the ages of the patients at diagnosis. The slight differences among the general population mean scores assigned to the six samples in Table 2 are due to differing age and gender distributions in these samples. The relatively high general population mean score in the line CARDIO is due to the high proportion of males in the CARDIO sample, and the high score in the SARCOID line reflects that the SARCOID patients are relatively young.

Table 2 shows that all mean scores measuring patients' current health were lower than the general population mean scores, with effect sizes between -1.55 and -0.28. More importantly in the context of this study, the patients' retrospectively assessed health states were higher than the expected mean scores in five of the six samples; the

strongest effect ($d=0.46$) was found for breast cancer survivors. Only sample 5 (cardiologic patients) showed an opposite trend.

The effect of age and gender on retrospective health assessments

Table 3 shows the comparisons between the patients' assessments and the general population's mean scores, broken down by age groups and gender. For the first five samples, we used the age cutoff 60 years (age at diagnosis). In the SARCOID sample, the participants were younger than in the other samples; therefore we used this group's median age at diagnosis (41 years) as the cutoff.

In five of the six samples, the health overestimation as reflected in the retrospective judgments was higher in the older age groups than in the younger age groups. The only exception was the SARCOID sample with low effect sizes for both age groups. All older subgroups overestimated their previous health, in two cases with effect sizes greater than 0.60. In the younger groups, however, these effects were weaker, and in three samples, the effects were negative, indicating an underestimation of the retrospectively assessed health in the younger age groups.

Gender comparisons were possible for those three samples which included males and females. In two of these (MIXED-CA and SARCOID), the males overestimated their previous health more strongly than the females did. In the third sample (CARDIO), we observed a general underestimation, and this was more pronounced in the females' group than in the males' group.

Assessments of QoL with the EORTC QLQ-C30

For sample 1 (BR-CA), retrospective assessments of QoL with the questionnaire EORTC QLQ-C30 were available. All functioning scales and all symptom scales indicated worse current QoL scores in the patient group compared with the general population (Table 4). The results for the retrospective assessments were inconsistent. In both domains, functioning and symptoms, there were subscales with higher scores for the patients and other subscales with higher scores for the general population. Concerning the sum score, averaging across the functioning and symptom scores (Giesinger et al., 2016), the patients' retrospectively

Table 3 Retrospective health assessments, broken down by age group and gender.

Sample	<i>n</i>	Current (Patients)		Retrospective (Patients)		General population		<i>d</i> (Current, General population)	<i>d</i> (Retrospective, General population)
		<i>M</i>	(<i>SD</i>)	<i>M</i>	(<i>SD</i>)	<i>M</i>	(<i>SD</i>)		
Age comparisons									
BR-CA									
≤ 60 y.	171	3.31	(1.11)	4.13	(0.83)	3.90	(0.89)	-0.59	0.27
> 60 y.	152	3.41	(1.04)	3.97	(0.98)	3.36	(0.96)	0.05	0.63
MIXED-CA									
≤ 60 y.	187	2.79	(1.33)	3.86	(1.07)	3.88	(0.89)	-0.98	0.00
> 60 y.	241	3.28	(1.20)	3.93	(0.98)	3.48	(0.93)	-0.19	0.47
GYN-CA									
≤ 60 y.	167	2.86	(1.18)	3.76	(1.06)	3.90	(0.89)	-1.00	-0.14
> 60 y.	187	2.90	(1.10)	3.61	(1.03)	3.36	(0.96)	-0.45	0.25
URO-CA									
≤ 60 y.	71	2.55	(1.23)	3.72	(0.97)	3.88	(0.89)	-1.25	-0.17
> 60 y.	126	3.32	(1.13)	3.90	(0.79)	3.36	(0.96)	-0.04	0.62
CARDIO									
≤ 60 y.	283	2.75	(1.15)	3.49	(0.96)	3.86	(0.89)	-1.09	-0.40
> 60 y.	73	2.93	(1.07)	3.64	(0.96)	3.48	(0.93)	-0.55	0.17
SARCOID									
≤ 41 y.	610	2.52	(1.05)	4.07	(0.95)	4.01	(0.86)	-1.56	0.07
> 41 y.	587	2.35	(1.01)	3.84	(0.99)	3.82	(0.90)	-1.54	0.02
Gender comparisons									
MIXED-CA									
Males	318	3.27	(1.20)	3.90	(1.00)	3.64	(0.91)	-0.35	0.27
Females	110	2.48	(1.32)	3.87	(1.08)	3.71	(0.91)	-1.10	0.16
CARDIO									
Males	298	2.81	(1.12)	3.55	(0.94)	3.80	(0.90)	-0.98	-0.27
Females	58	2.67	(1.19)	3.39	(1.05)	3.79	(0.90)	-1.07	-0.41
SARCOID									
Males	414	2.51	(1.08)	4.12	(0.86)	3.92	(0.88)	-1.44	0.23
Females	783	2.40	(1.01)	3.87	(1.03)	3.92	(0.88)	-1.61	-0.05

Note. *M*: mean; *SD*: standard deviation; *d*: effect size.

assessed mean score was nearly identical with that of the general population, and the effect size ($d = -0.04$) was negligible. However, the 2-item general health/QoL scale yielded a significant positive effect size ($d = 0.37$), indicating better retrospectively assessed health of the patients than the general population norms.

Discussion

The first aim of this study was to test whether patients suffering from a chronic disease retrospectively overestimate their pre-disease health states. We hypothesized that the retrospective health assessments were higher than the norm values (Scholten et al., 2017). With one exception (sample of the cardiology patients), this effect was confirmed in all samples. There are two possible explanations for this overestimation of pre-disease quality of health. One explanation is response shift, a change in the internal frame of reference (Sprangers & Schwartz, 1999). When people experience health problems they may become more tolerant of several detriments, and, as measured with this newly adapted internal scale, they may perceive their previous health problems as negligible. A second possible explanation

is a memory effect, combined with the implicit theory of change (Blome & Augustin, 2015). This means that the patients did not precisely remember their pre-disease health state, but that they have an implicit theory about the effect of a disease (e.g., cancer) on the health state of an individual, and, based on that theory, inferred that they must have been much healthier before they fell ill.

In the group of cardiologic patients, we observed a reverse effect featuring poorer retrospective health assessments. A possible reason for this peculiarity is that these patients might have experienced more risk factors or more objective health problems leading up to their cardiac event. Furthermore, patients attending a rehabilitation program are not representative of all cardiology patients since patients in relatively good health (which is associated with a good health previously as well) participate less frequently in rehabilitation programs. The comparison of the mean health scores in Table 2 also shows that the cardiology patients rated their current health as being poor. We cannot clarify the specific reasons for the low health scores of the cardiology patients here, however, one should be aware that retrospectively assessed health states are not always higher than the self-reported health states of the general population, and that the results obtained in studies with injury

Table 4 QoL mean scores (EORTC QLQ-C30) of the BR-CA sample.

	Current patients		Retrospective patients		General population		<i>d</i> (Current, General population)		<i>d</i> (Retrospective, General population)	
	<i>M</i>	(<i>SD</i>)	<i>M</i>	(<i>SD</i>)	<i>M</i>	(<i>SD</i>)				
Functioning scales										
Physical	74.4	(20.6)	89.7	(17.9)	90.2	(16.7)	-0.85	***	-0.03	<i>ns</i>
Role	68.9	(29.1)	88.1	(22.8)	88.6	(22.2)	-0.77	***	-0.02	<i>ns</i>
Emotional	64.2	(26.3)	76.6	(23.3)	81.6	(19.2)	-0.76	***	-0.24	***
Cognitive	80.4	(23.6)	93.2	(14.1)	89.4	(15.7)	-0.46	***	0.26	***
Social	75.4	(27.8)	87.8	(20.2)	90.0	(18.1)	-0.64	***	-0.11	*
Symptoms										
Fatigue	38.3	(27.0)	18.7	(22.9)	22.1	(21.7)	0.67	***	-0.15	**
Nausea/Vomiting	5.7	(14.3)	3.0	(11.4)	3.3	(9.1)	0.21	**	-0.03	<i>ns</i>
Pain	32.9	(29.7)	18.0	(24.5)	22.1	(25.0)	0.39	***	-0.17	**
Dyspnea	26.9	(32.1)	12.6	(23.9)	11.9	(21.8)	0.56	***	0.03	<i>ns</i>
Insomnia	43.3	(34.5)	23.4	(29.5)	21.6	(24.6)	0.73	***	0.07	<i>ns</i>
Appetite loss	14.3	(27.1)	6.6	(19.4)	4.9	(13.1)	0.47	***	0.10	<i>ns</i>
Constipation	13.5	(25.0)	12.3	(22.4)	7.6	(8.2)	0.36	***	0.31	***
Diarrhea	11.0	(21.9)	7.2	(16.9)	4.9	(13.8)	0.34	***	0.15	**
Financial difficult.	23.1	(32.6)	11.6	(23.9)	6.9	(16.4)	0.66	***	0.23	***
Global scores										
Global health/QoL	63.0	(19.3)	78.7	(18.2)	71.5	(21.2)	-0.42	***	0.37	***
Sum score	75.2	(17.4)	87.2	(13.7)	87.8	(14.4)	-0.79	***	-0.04	<i>ns</i>

Note. * $p < .05$; ** $p < .01$; ***: $p < .001$; *ns*: not significant.

patients cannot be generalized to all patients with chronic diseases. While the effect of retrospective overestimations of the health state is not at all new, our study adds information on patients with chronic diseases.

Current and previous health were assessed with single questions that might be considered weak measures with unsecure reliability. However, many studies reported in the literature use such single-item variables with five answer alternatives for measuring the subjectively experienced health state. While such short instruments are not appropriate for deriving conclusions on the individual level, we believe that conclusions on the group level are nevertheless justified when the sample sizes are large enough.

In one of the samples (BR-CA), we used a complex questionnaire for assessing health-related QoL with current and retrospective measurements. The mean scores of the current functioning scales were much lower and the means of the symptom scales were much higher than the means of the general population; the most relevant differences (effect sizes above 0.60) were found for physical functioning, role functioning, emotional functioning, social functioning, fatigue, insomnia, and financial difficulties. Such comparisons between patients and the general population concerning their QoL have been frequently reported in the literature. The focus of this study, however, was on retrospective assessments. The comparison between the retrospective judgments and the population norms does not support the hypothesis of a general overestimation of all health aspects. The results for the 15 scales were inconsistent, with overestimations and underestimations. Most effects sizes were small. Despite the similarity of the patients' retrospective assessments and the population norms in most functioning and symptom scales, the

difference in the two-item general health/QoL scale was remarkable with a statistically significant ($p < .001$) effect size of $d = 0.37$. The patients' retrospectively assessed global health/QoL was 78.7, while the corresponding value of the general population was only 71.5. In contrast to that, the effect size of the sum score, which aggregated across the functioning and symptoms scales, was negligible ($d = -0.04$). This seemingly contradictory result is in line with findings from other studies that found that global aspects of QoL undergo other response shift effects than more specific aspects of QoL do (Hinz et al., 2017). Health is a complex construct with several implicit components such as physical health, mental health, and functional limitations (Idler & Cartwright, 2018), with an individual-specific combination of positive and negative aspects of health (Schönfeld et al., 2017). Our results support the idea that complex constructs such as general health are more affected by response shift and memory effects than more specific variables.

Age effects of retrospective overestimations were found in five of the six samples. With the exception of the SARCOID sample, older patients showed stronger overestimations of their previous health states than younger patients. A large study with more than 150,000 participants showed that older people tend to overestimate their subjective health state when compared to objective assessments made by physicians (Idler & Cartwright, 2018). Moreover, there is a declining alignment of self-rated health and objective health and/or physical limitations for older persons, while depression symptoms are more strongly correlated with self-rated health in the oldest group compared to any other group (Schnitker, 2005). Our observation of larger retrospective health overestimations in the older age group is in line with those findings that showed physical limitations

being disregarded in older people's self-assessments of their health. Examinations that include objective measurements of health, subjective measurements of health, and retrospective health assessments could clarify the degree to which response shift effects and memory effects explain such overestimations of previous health states. Concerning gender, the three samples did not show consistent effects and do not allow for drawing generalizable conclusions about gender differences in response shift or memory effects.

Some limitations of this study should be mentioned. The normative scores which served as comparison values were taken from specific normative studies, and it is possible that there was some bias in these studies. However, normative studies tend to select positive samples (Enzenbach et al., 2019), which means that the overestimations of the previous health states reported in this paper would become even stronger if the samples of the normative studies were fully representative. In the examinations analyzed here, there might also have been a selection effect, resulting in an underrepresentation of patients in overall poor health, while there are arguments for a negative bias in the sample of cardiology rehabilitation patients. The patients were asked to assess the quality of their health at the time before the diagnosis. It is possible that the patients interpreted this instruction as referring to disease-free moments in their life before the diagnosis, leading to an overestimation of health. We found overestimations of previous health states in five of the six examinations, but we could not clarify which of the possible explanations (response shift, memory effect, implicit theory of health change) contributed most to these effects. As already mentioned above, the reliability of the single-item measurement of subjectively experienced health is insecure, therefore, the results should be only interpreted on the group level. While the most relevant finding, the overestimation of retrospectively assessed health, has already been reported in the literature, our study adds information on patients with specific chronic diseases and on age and gender differences of this effect. Finally, we examined heterogeneous samples of patients with different diagnoses, treatments, and settings. This limits the comparability of the results from these samples, however, it gives an impression of the generalizability and robustness of the findings.

Taken together, the overestimation of previous health states indicates that, in the case of chronic diseases, retrospective health assessments are inappropriate tools for assessing the impact of a disease on a person's subjectively experienced health.

Acknowledgement

We acknowledge support from the German Research Foundation (DFG; Grant Number: HI 1108/8-1) and Universität Leipzig within the program of Open Access Publishing.

References

- Aaronson, N. K., Ahmedzai, S., Bergman, B., Bullinger, M., Cull, A., Duez, N. J., Filiberti, A., Flechtner, H., Fleishman, S. B., Dehaes, J. C. J. M., Kaasa, S., Klee, M., Osoba, D., Razavi, D., Rofe, P. B., Schraub, S., Sneeuw, K., Sullivan, M., & Takeda, F. (1993). The European-Organization-For-Research-And-Treatment-Of-Cancer QLQ-C30 - A quality-of-life instrument for use in international clinical trials in oncology. *Journal of the National Cancer Institute*, 85, 365–376. <https://doi.org/10.1093/jnci/85.5.365>
- Arndt, V., Merx, H., Stegmaier, C., Ziegler, H., & Brenner, H. (2005). Persistence of restrictions in quality of life from the first to the third year after diagnosis in women with breast cancer. *Journal of Clinical Oncology*, 23, 4945–4953. <https://doi.org/10.1200/JCO.2005.03.475>
- Blome, C., & Augustin, M. (2015). Measuring change in quality of life: Bias in prospective and retrospective evaluation. *Value in Health*, 18, 110–115. <https://doi.org/10.1016/j.jval.2014.10.007>
- Bosse-Henck, A., Koch, R., Wirtz, H., & Hinz, A. (2017). Fatigue and excessive daytime sleepiness in sarcoidosis: Prevalence, predictors, and relationships between the two symptoms. *Respiration*, 94, 186–197. <https://doi.org/10.1159/000477352>
- Brickman, P., Coates, D., & Janoff-Bulman, R. (1978). Lottery winners and accident victims - Is happiness relative? *Journal of Personality and Social Psychology*, 36, 917–927.
- Daig, I., Herschbach, P., Lehmann, A., Knoll, N., & Decker, O. (2009). Gender and age differences in domain-specific life satisfaction and the impact of depressive and anxiety symptoms: a general population survey from Germany. *Quality of Life Research*, 18, 669–678.
- Enzenbach, C., Wicklein, B., Wirkner, K., & Loeffler, M. (2019). Evaluating selection bias in a population-based cohort study with low baseline participation: The LIFE-Adult-Study. *BMC Medical Research Methodology*, 19, Article 135 <https://doi.org/10.1186/s12874-019-0779-8>
- Friedrich, M., Karoff, J., & Hinz, A. (2019). Response shift effects in patients' assessments of their quality of life after cardiac rehabilitation. *Quality of Life Research*, 28, 2609–2620. <https://doi.org/10.1007/s11136-019-02195-9>
- Friedrich, M., Zenger, M., & Hinz, A. (2019). Response shift effects of quality of life assessments in breast cancer survivors. *European Journal of Cancer Care*, 28, e12979. <https://doi.org/10.1111/ecc.12979>
- Giesinger, J. M., Kieffer, J. M., Fayers, P. M., Groenvold, M., Petersen, M. A., Scott, N. W., Sprangers, M. A., Velikova, G., & Aaronson, N. K. (2016). Replication and validation of higher order models demonstrated that a summary score for the EORTC QLQ-C30 is robust. *Journal of Clinical Epidemiology*, 69, 79–88. <https://doi.org/10.1016/j.jclinepi.2015.08.007>
- Graaf, M. W. de, Reininga, I. H. F., Wendt, K. W., Heine-man, E., & El Mounni, M. (2019). Pre-injury health status of injured patients: A prospective comparison with the Dutch population. *Quality of Life Research*, 28, 649–662. <https://doi.org/10.1007/s11136-018-2035-9>
- Henrich, G., & Herschbach, P. (2000). Questions on Life Satisfaction (FLZ^(M)) - A short questionnaire for assessing subjective quality of life. *European Journal of Psychological Assessment*, 16, 150–159.
- Hinz, A., Karoff, J., Kittel, J., Brähler, E., Zenger, M., Schmalbach, B., & Kocalevent, R.-D. (2020). Associations between self-rated health and the assessments of anchoring vignettes in cardiovascular patients. *International Journal of Clinical and Health Psychology*, 20, 100–107. <https://doi.org/10.1016/j.ijchp.2020.04.001>
- Hinz, A., Mehnert, A., Dégi, C., Reissmann, D. R., Schotte, D., & Schulte, T. (2017). The relationship between global and specific components of quality of life, assessed with the EORTC QLQ-C30 in a sample of 2019 cancer patients. *European Journal of Cancer Care*, 26, Article e12416 <https://doi.org/10.1111/ecc.12416>
- Hinz, A., Singer, S., & Brähler, E. (2014). European reference values for the quality of life questionnaire EORTC QLQ-C30: Results of a German investigation and a summarizing analysis of six Euro-

- pean general population normative studies. *Acta Oncologica*, 53, 958–965. <https://doi.org/10.3109/0284186X.2013.879998>
- Idler, E., & Cartwright, K. (2018). What do we rate when we rate our health? Decomposing age-related contributions to self-rated health. *Journal of Health and Social Behavior*, 59, 74–93. <https://doi.org/10.1177/0022146517750137>
- Kecke, S., Ernst, J., Eienkel, J., Singer, S., & Hinz, A. (2017). Psychometric properties of the fatigue questionnaire EORTC QLQ-FA12 in a sample of female cancer patients. *Journal of Pain and Symptom Management*, 54, 922–928. <https://doi.org/10.1016/j.jpainsymman.2017.08.007>
- Lindberg, P., Netter, P., Koller, M., Steinger, B., & Klinkhammer-Schalke, M. (2017). Breast cancer survivors' recollection of their quality of life: Identifying determinants of recall bias in a longitudinal population-based trial. *PLoS One*, 12, Article e0171519 <https://doi.org/10.1371/journal.pone.0171519>
- Nolte, S., Liegl, G., Petersen, M. A., Aaronson, N. K., Costantini, A., Fayers, P. M., Groenvold, M., Holzner, B., Johnson, C. D., Kemmler, G., Tomaszewski, K. A., Waldmann, A., Young, T. E., & Rose, M. (2019). General population normative data for the EORTC QLQ-C30 health-related quality of life questionnaire based on 15,386 persons across 13 European countries, Canada and the United States. *European Journal of Cancer*, 107, 153–163. <https://doi.org/10.1016/j.ejca.2018.11.024>
- Oort, F. J. (2005). Using structural equation modeling to detect response shifts and true change. *Quality of Life Research*, 14, 587–598. <https://doi.org/10.1007/s11136-004-0830-y>
- Preiß, M., Friedrich, M., Stolzenburg, J.-U., Zenger, M., & Hinz, A. (2019). Response shift effects in the assessment of urologic cancer patients' quality of life. *European Journal of Cancer Care*, 28, Article e13027 <https://doi.org/10.1111/ecc.13027>
- Schnittker, J. (2005). When mental health becomes health: Age and the shifting meaning of self-evaluations of general health. *Milbank Quarterly*, 83, 397–423. <https://doi.org/10.1111/j.1468-0009.2005.00407.x>
- Scholten, A. C., Haagsma, J. A., Steyerberg, E. W., van Beeck, E. F., & Polinder, S. (2017). Assessment of pre-injury health-related quality of life: A systematic review. *Population Health Metrics*, 15, Article 10 <https://doi.org/10.1186/s12963-017-0127-3>
- Schönfeld, P., Brailovskaia, J., & Margraf, J. (2017). Positive and negative mental health across the lifespan: A cross-cultural comparison. *International Journal of Clinical and Health Psychology*, 17, 197–206. <https://doi.org/10.1016/j.ijchp.2017.06.003>
- Schwartz, C. E., Bode, R., Repucci, N., Becker, J., Sprangers, M. A. G., & Fayers, P. M. (2006). The clinical significance of adaptation to changing health: a meta-analysis of response shift. *Quality of Life Research*, 15, 1533–1550. <https://doi.org/10.1007/s11136-006-0025-9>
- Sprangers, M. A. G., & Schwartz, C. E. (1999). Integrating response shift into health-related quality of life research: a theoretical model. *Social Science in Medicine*, 48, 1507–1515. [https://doi.org/10.1016/S0277-9536\(99\)00045-3](https://doi.org/10.1016/S0277-9536(99)00045-3)
- Starshinova, A. A., Malkova, A. M., Basantsova, N. Y., Zinchenko, Y. S., Kudryavtsev, I. V., Ershov, G. A., ... & Yablonskiy, P. K. (2020). Sarcoidosis as an autoimmune disease. *Frontiers in Immunology*, 10, Article 2933 <https://doi.org/10.3389/fimmu.2019.02933>
- ten Ham, R. M. T., Broering, J. M., Cooperberg, M. R., Carroll, P., & Wilson, L. S. (2020). Understanding the major factors affecting response shift effects on health-related quality of life: What the then-test measures in a longitudinal prostate cancer registry. *Clinical Genitourinary Cancer*, 18, e21–e27. <https://doi.org/10.1016/j.clgc.2019.10.026>
- Zenger, M., Brix, C., Borowski, J., Stolzenburg, J., & Hinz, A. (2010). The impact of optimism on anxiety, depression and quality of life in urogenital cancer patients. *Psycho-Oncology*, 19, 879–886. <https://doi.org/10.1002/pon.1635>
- Zenger, M., Glaesmer, H., Höckel, M., & Hinz, A. (2011). Pessimism predicts anxiety, depression and quality of life in female cancer patients. *Japanese Journal of Clinical Oncology*, 41, 87–94. <https://doi.org/10.1093/jjco/hyq168>