Quality of Life Research (2021) 30:117–127
https://doi.org/10.1007/s11136-020-02628-w

Randomized Clinical Trial on Reduction of Radiotherapy Dose to the Elective Neck in Head and Neck Squamous Cell Carcinoma: Results on the Quality of Life

S. Deschuymer[1] · D. Nevens[1,5] · F. Duprez[2] · J. F. Daisne[1,3] · M. Voordeckers[4] · W. De Neve[2] · S. Nuyts[1]

Accepted: 2 September 2020 / Published online: 13 September 2020

Abstract

Purpose A randomized trial was initiated to investigate whether a reduction of the dose to the elective nodal sites would result in less toxicity and improvement in Quality of Life (QoL) without compromising tumor control. This paper aimed to compare QoL in both treatment arms.

Methods Two-hundred head and neck cancer patients treated with radiotherapy (RT) or chemo-RT were randomized (all stages, mean age: 60 years, M/F: 82%/18%). The elective nodal volumes of patients randomized in the experimental arm were treated up to a 40 Gy equivalent dose. In the standard arm, the elective nodal volumes were treated up to a 50 Gy equivalent dose. The QoL data were collected using The European Organization for Research and Treatment of Cancer (EORTC) core questionnaire QLQ-C30 and the EORTC Head and Neck Cancer module (H&N35).

Results A trend toward less decline in QoL during treatment was observed in the 40 Gy arm compared to the 50 Gy arm. Statistically significant differences for global health status, physical functioning, emotional functioning, speech problems, and trouble with social eating in favor of the 40 Gy arm were observed. A clinically relevant better outcome in the 40 Gy arm was found for physical functioning at the end of therapy.

Conclusion QoL during RT for head and neck cancer tends to be less impaired in the 40 Gy arm. However, reducing the dose only on the elective neck does not result in clinically relevant improvement of QoL. Therefore, additional treatment strategies must be examined to further improve the QoL of HNSCC patients.

Keywords Quality of life · Head and neck cancer · Radiotherapy · Elective nodes · Dose reduction

Introduction

Radiotherapy (RT) for head and neck squamous cell carcinoma (HNSCC) has made enormous progress over the past decades and treatment intensification led to improved clinical outcome [1]. Unfortunately, intensification also increased both acute and late toxicity, heavily compromising the quality of life (QoL) of the surviving HNSCC patients with xerostomia and late swallowing disorders of particular concern [2–9]. From a dosimetric point of view, the dose delivered to the pharyngeal constrictor muscles plays a crucial role in the development of severe late dysphagia [9–12]. A possible means to limit the dose to these organs at risk is by delivering a lower RT dose to the elective nodal volume. Historically, the advised dose to achieve microscopic sterilization is 45–50 Gy in 1.8–2 Gy fractions [13]. To this end, a multicenter, randomized clinical trial was initiated, reducing the dose to the elective nodal sites from 50 to 40 Gy and

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11136-020-02628-w) contains supplementary material, which is available to authorized users.

✉ S. Nuyts
sandra.nuyts@uzleuven.be

1 Present Address: Department of Radiation Oncology, KU Leuven - University of Leuven, University Hospitals Leuven, Herestraat 49, 3000 Leuven, Belgium

2 Department of Radiotherapy-Oncology, Ghent University Hospital, Ghent, Belgium

3 Department of Radiation Oncology, Université Catholique de Louvain, CHU-UCL-Namur, Site Ste-Elisabeth, Namur, Belgium

4 Department of Radiation Oncology, UZ Brussel, Vrije Universiteit Brussel, Brussel, Belgium

5 Present Address: Department of Radiation Oncology, Iridium Kanker Netwerk, Faculty of Medicine and Health Sciences, University of Antwerp, Antwerp, Belgium

off-target swallowing apparatus using intensity-modulated RT (IMRT). Previously, we have demonstrated that this treatment strategy results in significantly less dose to the pharyngeal constrictor muscles leading to significantly less severe physician-scored dysphagia at 3 months following treatment and to a trend toward less dysphagia at 6 months [14, 15]. Furthermore, dose de-escalation to the elective nodal volume in HNSCC leads to less physician-scored salivary gland toxicity without significant differences in disease control or survival [15]. Following 40 Gy, the actuarial rate of recurrence in the electively irradiated lymph node regions was comparable to recurrence rates observed after a standard dose of 50 Gy [15–17]. In this trial, QoL was prospectively registered as a secondary outcome with the hypothesis that reducing the dose to the elective neck would result in less decline of QoL during RT and better long-term QoL compared to the standard treatment.

Material and methods

Study design and patients

A prospective multicenter, randomized phase III study was set up in 6 Belgian centers. Inclusion criteria were previously untreated, histologically proven squamous cell carcinoma of the oral cavity, oropharynx, hypopharynx, larynx, or cervical lymph node metastases with unknown primary cancer (CUP). Patients were older than 18 years with a Karnofsky performance status ≥ 70%. Concurrent chemotherapy was allowed, as well as pretreatment lymph node dissection. Local ethics committee approval (Ethische Commissie Onderzoek UZ/KU Leuven) was obtained before the start of the study, and all patients gave written informed consent. Patients were randomized to the treatment arms (experimental arm versus standard arm) in a 1:1 design. Randomization was performed centrally (UZ Leuven) using permuted blocks (block size of 4) stratified per center to minimize the influence of center-specific parameters [14]. (NCT01812486).

Endpoints and power calculation

The primary endpoint of the study was the rate of physician-scored dysphagia at 6 months of follow-up. Assuming for the primary endpoint, a 70% rate of late dysphagia ≥ grade 1 will be unacceptable and a 50% rate will be expected ($\alpha = 0.05$, $p = 0.8$, two-tailed test), the power calculation for the study resulted in 200 patients (100 patients per arm). Secondary endpoints were clinical outcome, acute and late toxicities, and QoL [14]. This paper addresses QoL, which is one of the secondary endpoints of the trial.

Treatment

After randomization, a planning computed tomography study was performed. The primary tumor and enlarged lymph nodes were contoured separately and defined as gross tumor volume. All macroscopically affected tumor sites were treated up to an equivalent dose in 2 Gy fractions (EQD2) of 70 Gy with IMRT. The elective nodal volumes were in both arms delineated according to the guidelines of Grégoire et al. [18]. These volumes were treated up to an EQD2 of 40 Gy in patients in the experimental arm. For the standard arm, the elective nodal volumes were treated up to an EQD2 of 50 Gy. Fractionation scheme, PTV margins, and quality assurance were conducted according to the institutional guidelines.

QoL questionnaires

The QoL data were collected using The European Organization for Research and Treatment of Cancer (EORTC) core questionnaire QLQ-C30 and the EORTC Head and Neck Cancer module (QLQ-H&N35) [19, 20].

The cancer-specific QLQ-C30 questionnaire includes 30 questions organized into five functional scales (physical, role, emotional, cognitive, and social) and three symptom scales (nausea/vomiting, pain, and fatigue). It also includes a global health/QoL scale and six single additional symptom items (constipation, diarrhea, loss of appetite, insomnia, dyspnea, and financial difficulties) [19].

The QLQ-H&N35 questionnaire measures symptoms and problems specific to HNSCC and HNSCC treatment-related side effects. The questionnaire includes 35 questions organized into seven symptom scales (local pain, problems with: swallowing, senses, speech, social eating, social contact, and sexuality), six single symptom items (problems with: teeth, opening the mouth, dry mouth, sticky saliva, coughing and feeling ill), and 5 yes/no questions (use of painkillers, nutritional supplements, feeding tube, losing or gaining weight) [20]. Both QLQ-C30 and QLQ-H&N35 questionnaires were available in Dutch and French and had robust psychometric properties established by rigorous testing and external validation [21, 22]. Both instruments were assessed on paper at baseline (before the start of the treatment), at the end of treatment (EOT) and at 1, 2, 3, 6, 12, 18, and 24 months following the EOT. If the patient's disease progressed, QoL was no longer assessed.

Statistical analysis

Patient, disease, and treatment characteristics were compared with two-sided Fisher's exact test for categorical

variables. Compliance levels were calculated following a standard EORTC procedure, and the number of received forms was divided by the number of expected forms (total number of patients minus the number of patients with tumor recurrence or deceased patients) at each assessment point.

The questions of both QLQ-C30 and QLQ-H&N35 questionnaires were scaled and scored using the recommended EORTC Quality of Life Group procedures [23]. Raw scores were transformed to a linear scale ranging from 0 to 100. Provided that at least half of the questions in a multi-item scale were completed, the scale score was calculated with only the existing values [23]. Higher scores on functional and global health/QoL scales indicate a higher level of functioning, while a higher score on the symptom scales or single items indicates more prominent symptoms or problems. For the 5 yes/no questions, the scores indicate the percentage of "yes" answers [19]. The internal consistency of multi-item scales was investigated through Cronbach's alpha, with > 0.70 considered as high internal consistency [22].

Changes of QoL over time from baseline values were analyzed, including only data from patients responding to all questionnaires in the follow-up period up to 12 months and who were tumor-free at 12 months. To interpret improvement or deteriorations over time as trivial, small, medium, or large, separate thresholds were used for the different subscales of the QLQ-C30 according to Cocks et al. [24]. The changes in H&N35 scales were interpreted according to Osaba et al.: the mean differences from 0 to 5 points was considered as trivial, from 5 to 10 points as small, from 10 to 20 points as medium and more than 20 points as large [25]. Medium and large differences were considered to be a clinically meaningful improvement or deterioration.

Next, a longitudinal analysis was performed to assess differences between the two groups over time, including data from all disease-free patients at each specific time point. For this analysis, a Generalized Estimating Equations (GEE) proportional odds model was used, with an independent correlation matrix to account for the clustering of the data. The longitudinal proportional odds model includes a factor for randomized treatment, visit, and treatment-visit interaction. The inclusion of the interaction allows for the treatment effect to differ between visits. The effect of randomized treatment was estimated at each visit, and analysis was adjusted for baseline values. For both questionnaires, differences between the two arms of at least 10 points (on a scale of 0–100) were classified as the minimum clinically meaningful difference in the mean value of the QoL parameter. All tests were 2-sided and assessed at a significance level of 5%. Due to the exploratory nature of the study, no adjustments were made to the significance level to account for multiple testing. All analyses were using SAS System for Windows.

Results

Patient characteristics and compliance

Between 2008 and 2011, a total of 200 patients were included in the study (100 for each arm). Data from 1 center, which included seven patients, were not retrieved after randomization. The patient, disease, and treatment characteristics are depicted in Table 1. No grouping according to survival, site, gender, age, or other clinical parameters was done, since both groups were well balanced at baseline. The frequency of surgery, upfront neck dissections, and concomitant chemotherapy was not significantly different between both groups at baseline and randomization was done per center [14]. This assures that the differences found in the QoL are related to the difference in treatment between both study arms, namely the different RT dose to the elective nodal volumes.

The compliance at every time point is given in Table 2. Compliance was 92% at baseline, 86% at 3 months of follow-up, 82% at 6 months, 83% at 1 year of follow-up, and 74% at 2 years of follow-up.

One hundred and three patients filled in all seven questionnaires up to 12 months after the EOT (77% of the expected 134 forms). At 24 months, only 68 patients (56% of the expected 121) completed the questionnaires at all nine time points. No significant differences in patient and treatment characteristics could be observed between the patients who completed all questionnaires and those who did not.

Baseline assessment

An overview of the baseline mean QLQ-C30 cores and comparison with reference data (Head and neck cancer: all stages & Head and neck cancer: stage III–IV) is given in Table 3 [26]. We did not observe clinically relevant differences (> 10 points of difference, observed difference ranges from 0 to 8) between the mean baseline values from the QLQ-C30 and the reference data [26]. However, we did observe clinically relevant differences between our mean baseline values from the QLQ-HN35 and the reference data, namely for teeth problems, dry mouth, sticky saliva, the use of painkillers, the use of a feeding tube, weight loss, and weight gain [26]. Furthermore, in both questionnaires, no clinically relevant differences (range 0–7) were seen between both treatment arms at baseline except for the use of painkillers (absolute difference in mean (Δ) = 13) and weight loss (Δ = 12).

Cronbach's alpha of the scales was > 0.70 in 13 of the 16 multi-item scales (Table 3). The values ranged from 0.94 to 0.51. The scales with alpha < 0.70 were cognitive functioning, nausea and vomiting, and senses.

Table 1 Patient, disease and treatment characteristics at baseline

	Arm A (n = 96) (%)	Arm B (n = 97) (%)	P-value
Age	60 (8)	59 (8)	0.42
Mean (SD)	80 (83.3)	85 (87.6)	
< 70	16 (17.7)	12 (12.4)	
≥ 70			
Gender	75 (78.1)	84 (86.6)	0.14
M	21 (21.9)	13 (13.4)	
F			
Karnofsky performance status	70 (72.9)	67 (69.1)	0.47
> 80	26 (27.1)	28 (28.9)	
≤ 80	0 (0.0)	2 (2.1)	
Unknown			
Tumor subsite	4 (4.2)	5 (5.2)	0.99
CUP	18 (18.8)	18 (18.6)	
Larynx	11 (11.5)	9 (9.3)	
Oral Cavity	22 (22.9)	23 (23.7)	
Hypopharynx	41 (42.7)	42 (43.3)	
Oropharynx	7	10	
HPV +	31	29	
HPV –	3	3	
HPV unknown			
AJCC stage	1 (1.0)	0 (0.0)	0.53
I	10 (10.4)	12 (12.4)	
II	19 (19.8)	25 (25.8)	
III	66 (68.8)	60 (61.9)	
IV			
T-stage	4 (4.2)	5 (5.2)	0.39
0	1 (1.0)	4 (4.1)	
1	32 (33.3)	40 (41.2)	
2	34 (35.4)	30 (30.9)	
3	25 (26.0)	18 (18.6)	
4			
N-stage	1 (1.0)	0 (0.0)	0.77
x	22 (22.9)	26 (26.8)	
0	16 (16.7)	14 (14.4)	
1	6 (6.3)	5 (5.2)	
2a	32 (33.3)	29 (29.9)	
2b	16 (16.7)	22 (22.7)	
2c	3 (3.1)	1 (1.0)	
3			
Concurrent systemic treatment	62 (64.6)	68 (70.1)	0.45
Yes	34 (35.4)	29 (29.9)	
No			
Neo-adjuvant systemic treatment	5 (5.2)	2 (2.1)	0.28
Yes	91 (94.8)	95 (97.9)	
No			
Pretreatment neck dissection	17 (17.7)	19 (19.6)	0.85
Yes	79 (82.3)	78 (80.4)	
No			

KI Karnofsky Index, *CUP* carcinoma of unknown primary, *HPV* Human Papillomavirus, *AJCC* American Joint Committee on Cancer 7th edition

Changes over time from baseline to 12 months of follow-up

In the changes over time, 19 of the 27 scales and single items had a clinically meaningful deterioration (6 medium and 13 large deteriorations) in the 40 Gy group from baseline to EOT. (Supplementary data) While in the 50 Gy arm, 23 of the 27 scales showed clinically meaningful deterioration of which were 4 medium and 19 large. At 3 months, the majority of scores improved toward baseline levels with only 3 and 4 scales/items in the 40 Gy arm and 50 Gy arm, respectively, showing clinically meaningful deterioration from baseline. In both study arms, a large deterioration from baseline remained for dry mouth ($\Delta = 26$) and sticky saliva ($\Delta = 23$)

Table 2 Compliance for both study arms at every time point for QLQ-C30 and QLQ-H&N35

Time point	40 Gy arm (n)	50 Gy arm (n)	Total for both arms (n)	Expected n for both arms[b] (n)	Compliance
Baseline	88	89	177	193[a]	92%
End of RT	77	82	159	191	83%
Month 1	77	80	157	189	83%
Month 2	80	80	160	188	85%
Month 3	80	75	155	179	86%
Month 6	69	66	135	164	82%
Month 12	57	54	111	134	83%
Month 18	49	50	99	126	79%
Month 24	42	48	90	121	74%

[a]Information of 1 participating center (7 patients) was not retrieved after randomization, yielding 193 patients for analysis (96 in the experimental arm and 97 in control arm)

[b]Expected n = total number of patients minus the number of patients with tumor recurrence of deceased patients

at 12 months of follow-up. A medium to large improvement from baseline could only be measured in the 40 Gy arm as follows: global health status ($\Delta = 11$), insomnia ($\Delta = 17$), and emotional functioning ($\Delta = 16$).

Difference between the two treatment arms: QLQ-C30

Regarding *global health status* (Fig. 1)*, physical functioning* (Fig. 2)*, and nausea and vomiting,* the interaction between treatment and visit was found to be significant ($p < 0.01$, $p = 0.03$, and $p = 0.02$, respectively), indicating that the treatment effect differs at the different time points (Table 4). An overall effect toward better global health status, better physical functioning, and less nausea and vomiting was observed in the 40 Gy arm ($p < 0.01$, $p = 0.01$ and $p = 0.03$). Furthermore, there was a significant difference for *global health status* in favor of the 40 Gy arm at month 6 ($p = 0.04$; estimated effect: 6.76 95% CI 0.44 to 13.07) and month 18 ($p < 0.01$; estimated effect: 14.31 95% CI 7.03 to 21.60). However, the differences ($\Delta = 3$ and $\Delta = 7$) were not more than the clinically relevant threshold of ten points between both treatment arms. For *physical functioning*, the estimated effect favors the 40 Gy group at the EOT ($p < 0.01$; 9.53 95% CI 3.08–15.99) and at 1 month ($p = 0.02$; 6.77 95% CI 0.90–12.64) (Table 4). At the EOT, the difference between the mean values in both groups was more than 10 points (73.2 vs 62.2; $\Delta = 11$) in favor of the 40 Gy arm, suggesting that there was not only a statistically significant difference but also a clinically relevant difference (Table 4). After 1 month of follow-up, this difference between both treatment arms became fewer than ten points but was still in favor of the 40 Gy arm (76.1 vs 71.2; $\Delta = 5$). For *nausea and vomiting*, there was a significant ($p < 0.01$) but not clinically relevant difference ($\Delta = 9$) in favor of the 40 Gy group, at the EOT.

Regarding *emotional functioning,* the interaction between treatment and visit was found to be not significant ($p = 0.24$). This means the treatment effect was similar at all time points. When we removed the interaction from the model, the estimated effect between the 40 Gy group and the 50 Gy group was in favor of the 40 Gy arm for all time points ($p = 0.02$; 4.94 95% CI 0.73–9.15). However, the difference between the two arms never reached clinical relevance.

We did not observe significant differences between both groups regarding role functioning, cognitive functioning, social functioning, fatigue, pain, dyspnea, insomnia, appetite loss, constipation, diarrhea, and financial difficulties.

Differences between the two treatment arms: QLQ-H&N35

Regarding *swallowing problems and trouble with social eating,* the interaction between treatment and visit was not significant ($p = 0.08$ and $p = 0.45$, respectively). When we removed the interaction from the model, we only observed a trend toward fewer swallowing problems in the 40 Gy group ($p = 0.06$; -5.00 95% CI − 10.23 to 0.24), while there was a statistically significant effect toward less trouble with social eating in the 40 Gy arm for all time points ($p < 0.01$; − 8.15 95% CI − 13.89 to − 2.41) (Table 4). For both scales, we did not observed any clinically relevant difference between both groups.

We observed a significant interaction effect between treatment and visit regarding *speech problems* and *senses problems* ($p = 0.04$ and 0.03, respectively). Moreover, we observed a statistically significant overall effect of fewer speech problems and fewer senses problems in the 40 Gy arm ($p = 0.03$ and $p = 0.02$). We observed significant but not clinically relevant differences between both groups for speech problems at 24 months ($p = 0.02$; − 6.93 95%

Table 3 Baseline score of QLQ-C30 and QLQ-H&N35 for both treatment arms compared to the reference data and Cronbach's alpha for multi-item scale [26]

	40 Gy arm (N = 88), Mean (SD)	50 Gy arm (N = 89), Mean (SD)	Reference Data: Head and Neck Cancer at All Stages (N = 2929), Mean (SD)	Reference Data: Head and Neck Cancer Stage III-IV (N = 1722), Mean (SD)	Cronbach's alpha
QLQ-C30					
Global health status	56.4 (20.2)	62.6 (19.1)	64.1 (22.7)	63.1 (22.4)	0.85
Physical Functioning	86.1 (16.7)	84.1 (19.9)	81.2 (20.4)	81.2 (20.2)	0.80
Role Functioning	76.8 (26.5)	80.3 (25.9)	78.9 (28.1)	78.8 (27.9)	0.87
Emotional Functioning	71.2 (22.3)	73.2 (24.3)	72.5 (24.1)	71.2 (24.1)	0.83
Cognitive Functioning	88.4 (19.1)	88.0 (19.3)	85.9 (19.7)	86.4 (19.1)	0.61*
Social Functioning	86.5 (21.3)	85.5 (24.9)	82.6 (24.7)	82.2 (24.7)	0.82
Fatigue	30.5 (26.2)	24.0 (23.7)	26.9 (24.9)	27.6 (25)	0.88
Nausea and Vomiting	6.3 (14.0)	3.9 (10.9)	5.3 (13.7)	5.2 (13.3)	0.51*
Pain	23.6 (25.7)	20.4 (25.4)	23.2 (26.1)	24.9 (26.3)	0.74
Dyspnea	20.6 (25.2)	14.6 (25.0)	18.2 (26.9)	18.0 (26.6)	–
Insomnia	31.4 (31.3)	29.2 (31.3)	27.3 (31.8)	28.5 (32.4)	–
Appetite loss	21.6 (31.8)	18.4 (31.6)	17.7 (28.2)	19.4 (29.3)	–
Constipation	13.1 (25.3)	10.1 (22.2)	11.1 (22.6)	11.7 (23.2)	–
Diarrhea	4.5 (12.5)	3.9 (12.9)	6.1 (16.9)	6.1 (16.7)	–
Financial difficulties	15.2 (26.1)	13.7 (25.7)	18.2 (29.6)	18.8 (30.2)	–
QLQ-H&N35					
Local pain	30.4 (25.1)	26.0 (25.4)	27.1 (24)	29.9 (25.1)	0.80
Swallowing	29.0 (27.9)	27.0 (25.9)	23.9 (25.3)	27.5 (26.1)	0.85
Senses problems	11.7 (19.5)	10.7 (20.2)	19.3 (28.8)	20.0 (30.0)	0.60*
Speech problems	21.8 (25.2)	19.9 (26.5)	28.0 (27.6)	27.1 (27.2)	0.78
Social eating	24.2 (26.7)	20.9 (24.1)	20.9 (25.1)	23.9 (26.7)	0.85
Social contact	10.7 (15.2)	10.3 (20.0)	13.0 (18.9)	13.2 (19.1)	0.87
Sexuality	34.7 (34.6)	28.5 (34.9)	31.3 (35.2)	32.3 (36.1)	0.94
Teeth problems	**17.3 (31.2)**	**13.9 (27.4)**	**25.5 (33.2)**	**27.8 (35.0)**	–
Opening mouth	22.6 (32.3)	18.8 (30.6)	19.5 (29.5)	22.4 (31.9)	–
Dry mouth	**20.5 (27.7)**	**19.2 (27.0)**	**30.7 (33.4)**	**31.1 (34.2)**	–
Sticky Saliva	**21.2 (29.4)**	**19.4 (27.1)**	**30.5 (33.9)**	**32.4 (35.4)**	–
Coughing	31.1 (29.2)	28.7 (31.2)	33.9 (32.2)	34.9 (32.1)	–
Felt ill	16.1 (26.2)	16.9 (26.2)	21.6 (28.9)	21.7 (29.2)	–
Painkillers	**64.0 (48.0)**[$]	50.6 (50.0)[$]	**49.5 (50)**	**52.8 (49.9)**	–
Nutritional supplementation	18.2 (38.6)	21.2 (40.9)	26.7 (44.2)	27.0 (44.4)	–
Feeding tube	**5.7 (23.3)**	**7.1 (25.8)**	**19.7 (39.8)**	**18.3 (38.7)**	–
Weight loss	**51.8 (50.0)**[$]	39.8 (48.9)[$]	**38.9 (48.8)**	**41.3 (49.2)**	–
Weight gain	**12.9 (33.6)**	**13.4 (34.1)**	**27.3 (44.6)**	**25.9 (43.8)**	–

QLQ-C30 Quality of Life Questionnaire C30; *SD* standard deviation; *H&N35* European Organization for Research and Treatment of Cancer Quality of Life Head and Neck cancer-specific questionnaire

Bold type indicates clinically relevant difference between trial data and reference data (> 10 points)

[$]Indicates clinically relevant difference between the two treatment arms (> 10 points)

*Cronbach's alpha < 0.70 reflecting poor internal consistency of the multi-item scale

CI − 12.68 to − 1.19) and for senses problems at the EOT ($p < 0.01$; − 2.34 95% CI − 4.15 to − 1.32).

The other scales of the H&N35 questionnaire did not show any significant nor clinically relevant differences between both arms at any follow-up time point.

Discussion

This paper reports on QoL in HNSCC patients treated with RT and compares a RT equivalent dose of 40 Gy to the elective nodal neck with the standard RT dose of 50 Gy. The

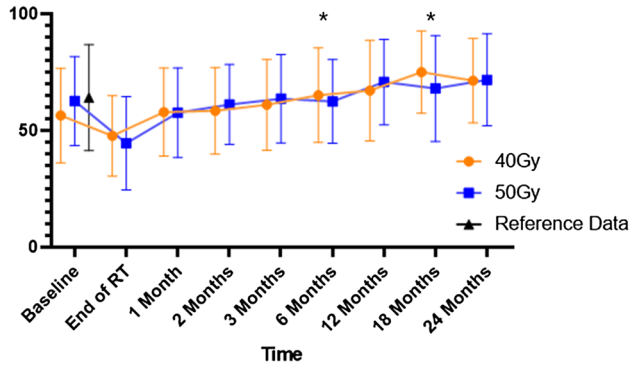

Fig. 1 Quality of Life/Global Health Status of the 2 treatment arms. Lines connect mean values (+standard deviation (SD)) at each visit. At baseline, reference data (mean + SD) are projected. Higher scores on the scale indicate a higher level of functioning. *Statistically significant differences between the two treatment arms ($p < 0.05$)

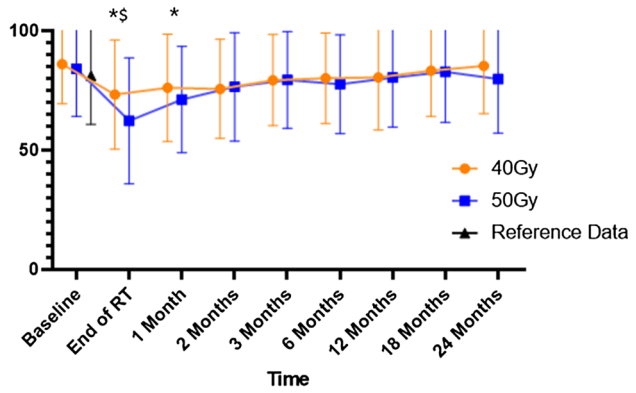

Fig. 2 Physical functioning of the 2 treatment arms. Lines connect mean values (+standard deviation(SD)) at each visit. At baseline, reference data (mean + SD) are projected. Higher scores on the scale indicate a higher level of functioning. *Statistically significant differences between the two treatment arms ($p < 0.05$). $^{\$}$Clinically meaningful difference between the two treatment arms (> 10 points difference)

primary endpoint of this randomized, phase III trial study was to detect a difference in physician-scored dysphagia at 6 months. Oncological outcome and QoL were secondary endpoints. We demonstrated in a previous paper that a dose de-escalation to the elective lymph nodes in HNSCC results in significantly less dose to the swallowing structures and less severe dysphagia at 3 months following treatment [14]. Furthermore, we observed a trend toward less dysphagia at 6 months and less moderate salivary gland toxicity without significant differences in disease control or survival [15–17].

QLQ-C30 scores were comparable to the reference values in patients with head and neck cancer. However, the baseline values of some single items of the QLQ-H&N35 questionnaire were lower than the reference values reflecting fewer prominent symptoms in our population. The baseline values of two items, the use of painkillers and weight loss, were relevantly higher in the 40 Gy arm compared to the 50 Gy arm and to the reference values. Since patient, tumor, and treatment characteristics of both arms were well balanced, no clear explanation for this baseline difference between both arms can be given. The differences observed between the reference and current data can be attributed to the substantial differences in sample size between our study population and the reference data and were also observed in a previous study [19, 27].

Compliance in our study is good to high and reached at least 80% up to 12 months of follow-up and 77% of the patients filled in all questionnaires up to 12 months. At 24 months, compliance was still good but only 56% filled in all questionnaires. To minimize bias and to have a representative sample, pattern and changes over time from baseline values were only calculated up to 12 months. Overall, a similar pattern in QoL can be observed in both treatment arms with medium to large deteriorations during the RT, which generally recovered to baseline levels early after the end of treatment, by month 2–3 of follow-up. This point is important and also supported by other literature showing that initial treatment can have short-term negative effects on QoL [28–30]. However, for dry mouth and sticky saliva, the mean results in our patient population never returned to baseline levels, and after 2 years of follow-up stayed clinically inferior when compared to baseline in both groups. Although the pattern of both arms is similar, there appears to be a trend toward a more considerable deterioration in QoL in the 50 Gy arm compared to the 40 Gy arm at the end of RT. Even more, global health status and emotional functioning were only in the 40 Gy arm improved compared to baseline, after one year of follow-up.

Focusing more on the difference between the two treatment arms, the results of the present study are in favor of the 40 Gy arm. We see overall better results in the 40 Gy arm, with statistically significant differences in favor of the 40 Gy arm for global health status, physical functioning, nausea and vomiting, emotional functioning, senses and speech problems, and trouble with social eating. The scales, nausea and vomiting, and senses problems should be interpreted with caution based on the low Cronbach's alpha reflecting poor internal consistency of these multi-item scales. Unfortunately, clinically relevant differences were only found for physical functioning in favor of the 40 Gy arm. The better physical functioning in the 40 Gy arm at the EOT might be related to the lower dose on the elective neck and the organs at risk with fewer side effects in the 40 Gy arm, as was demonstrated in the first paper [14, 15]. As stated before, a difference in physician-scored salivary function and swallowing dysfunction was observed between the 40 Gy and 50 Gy groups. Differences are not seen in patient-scored swallowing dysfunction nor dry mouth/sticky saliva. These

Table 4 Statistically significant differences between both treatment groups for QLQ-C30 and QLQ-H&N35

Scale	Time point	40 Gy arm Mean value (SD)	50 Gy arm Mean value (SD)	Estimated effect (95% CI)	p-value
QLQ-C30					
Global health status	Interaction				p < 0.01
	Overall effect				p < 0.01
	Month 6	65.1 (20.3)	62.4 (17.9)	6.73 (0.44 to 13.07)	p = 0.04
	Month 18	75.0 (17.6)	68.0 (22.7)	14.31 (7.03 to 21.60)	p < 0.01
Physical functioning	Interaction				p = 0.03
	Overall effect				p = 0.01
	End of RT	73.2 (22.9)	62.2 (26.4)	9.53 (3.08 to 15.99)	p < 0.01
	Month 1	76.1 (22.4)	71.2 (22.2)	6.77 (0.90 to 12.64)	p = 0.02
Nausea and	Interaction				p = 0.02
Vomiting	Overall effect				p = 0.03
	End of RT	20.1 (26.1)	28.9 (28.4)	− 2.24 (− 4.10 to − 1.23)	p < 0.01
Emotional functioning	Interaction				p = 0.24
	Without interaction			4.94 (0.73 to 9.15)	p = 0.02
QLQ-H&N35					
Swallowing problems	Interaction				P = 0.08
	Without interaction			− 5.00 (− 10.23 to 0.24)	P = 0.06
Trouble with social eating	Interaction				p = 0.45
	Without interaction			− 8.15 (− 13.89 to − 2.14)	p < 0.01
Speech problems	Interaction				p = 0.04
	Overall				p = 0.03
	Month 24	11.3 (14.4)	11.6 (16.2)	− 6.93 (− 12.68 to − 1.19)	p = 0.02
Senses problems	Interaction				p = 0.03
	Overall				p = 0.02
	End of RT	51.7 (26.8)	60.6 (28.5)	− 2.34 (− 4.15 to − 1.32)	p < 0.01

Bold type indicates clinically relevant differences between both treatment groups

QLQ-C30 Quality of Life Questionnaire C30; H&N35: European Organization for Research and Treatment of Cancer Quality of Life Head and Neck cancer-specific questionnaire; SD: standard deviation; 95% CI 95% confidence interval

findings are in agreement with other publications showing a poor correlation between patient and physician-scored toxicity [31, 32]. Kaae et al. reported that patients tend to report higher scores of xerostomia than the physician [32].

Although the dose to the swallowing muscles was reduced by lowering the dose to the elective nodal neck to 40 Gy, the RT volume of the elective neck and the dose to the macroscopic affected tumor sites (70 Gy) remained unchanged [14]. In addition, the head and neck region is a complex network of muscles, nerves, vasculature, salivary gland, and other organs at risk. All these structures influence the QoL of the patients. Probably, the acquired dose reduction in our study is insufficient to improve the Qol notably. Nonetheless, de-escalating the dose to the elective neck in combination with other treatment strategies such as volume reduction of the elective neck (NCT01287390, [33–35]), reduction of the high-dose region in HPV-positive oropharyngeal cancer (NRG-HN002), or protontherapy [36] still has great potential in improving the QoL of our patients.

The strengths of this trial are that, to our knowledge, there is no earlier report on QoL after de-escalating the dose to the elective neck, and the results of the current study are presented in accordance with guidelines for the reporting of QoL studies [37]. However, currently, there is a lack of standardization of endpoint definitions, the analysis and presentation of patient-reported outcome data [38–40]. In the future, this problem will be solved by new guidelines from the SISAQOL consortium (Setting International Standards in Analyzing Patient-Reported Outcomes and Quality of Life Endpoints data) [38, 39]. Limitations of this manuscript, include the high chance of type I error. No correction for multiple testing was performed, and these results should, therefore, be considered hypothesis-generating. Furthermore, the limited sample size does not allow subgroup analysis between HPV-positive and HPV-negative oropharyngeal cancers.

To conclude, we analyzed the secondary endpoint, QoL, between 2 arms of a radiotherapy randomized trial

in HNSCC comparing 40 Gy versus 50 Gy to the elective lymph nodes. Unfortunately, our results do not show definite clinically relevant differences and improvement in QOL with reducing the dose to the elective neck. However, a trend toward less decline in QoL by the end of RT was visible. In view of this, dose de-escalation and/or volume reduction strategies should be further explored in order to improve the QoL of HNSCC patients after treatment.

Acknowledgements Study organization and data collection were performed by Rita Aerts. Statistical analysis was done in collaboration with Ann Belmans.

Funding This project is supported by a grant from Stand up to Cancer (Kom op tegen Kanker), the Flemish Cancer Society.

Compliance with ethical standards

Conflict of interest All authors declare that they have no conflicts of interest.

Consent to Participate All patients gave written informed consent.

Ethical Approval This study was approved by the local ethics committee (Ethische Commissie Onderzoek UZ/KU Leuven) and was conducted in accordance with the ethical standards as laid down in the 1964 Declaration of Helsinki and its later amendments.

References

1. Grégoire, V., Langendijk, J. A., & Nuyts, S. (2015). Advances in radiotherapy for head and neck cancer. *Journal of Clinical Oncology, 33*(29), 3277–3284. https://doi.org/10.1200/JCO.2015.61.2994.
2. Bourhis, J., Overgaard, J., Audry, H., Ang, K. K., Saunders, M., Bernier, J., et al. (2006). Hyperfractionated or accelerated radiotherapy in head and neck cancer: A meta-analysis. *The Lancet, 368*(938), 843–854. https://doi.org/10.1016/S0140-6736(06)69121-6.
3. Blanchard, P., Baujat, B., Holostenco, V., Bourredjem, A., Baey, C., Bourhis, J., et al., …, MACH-CH Collaborative Group. (2011). Meta-analysis of chemotherapy in head and neck cancer (MACH-NC): A comprehensive analysis by tumour site. *Radiotherapy and Oncology, 100*(1), 33–40. https://doi.org/10.1016/j.radonc.2011.05.036.
4. Nuyts, S., Dirix, P., Clement, P. M. J., Vander Poorten, V., Delaere, P., Schoenaers, J., et al. (2009). Impact of adding concomitant chemotherapy to hyperfractionated accelerated radiotherapy for advanced head-and-neck squamous cell carcinoma. *International Journal of Radiation Oncology, Biology, Physics, 73*(4), 1088–1095. https://doi.org/10.1016/j.ijrobp.2008.05.042.
5. Machtay, M., Moughan, J., Trotti, A., Garden, A. S., Weber, R. S., Cooper, J. S., …, Ang, K. K. (2008). Factors associated with severe late toxicity after concurrent chemoradiation for locally advanced head and neck cancer: An RTOG analysis. *Journal of Clinical Oncology, 26*(21), 3582–3589. https://doi.org/10.1200/JCO.2007.14.8841.
6. Langendijk, J. A., Doornaert, P., Verdonck-de Leeuw, I. M., Leemans, C. R., Aaronson, N. K., & Slotman, B. J. (2008). Impact of late treatment-related toxicity on quality of life among patients with head and neck cancer treated with radiotherapy. *Journal of Clinical Oncology, 26*(22), 3770–3776. https://doi.org/10.1200/JCO.2007.14.6647.
7. Nutting, C. M., Morden, J. P., Harrington, K. J., Urbano, T. G., Bhide, S. A., Clark, C., et al. (2011). Parotid-sparing intensity modulated versus conventional radiotherapy in head and neck cancer (PARSPORT): A phase 3 multicentre randomised controlled trial. *The Lancet Oncology, 12*(2), 127–136. https://doi.org/10.1016/S1470-2045(10)70290-4.
8. Langendijk, J. A., Doornaert, P., Rietveld, D. H. F., Verdonck-de Leeuw, I. M., Leemans, C. R., & Slotman, B. J. (2009). A predictive model for swallowing dysfunction after curative radiotherapy in head and neck cancer. *Radiotherapy and Oncology, 90*(2), 189–195. https://doi.org/10.1016/j.radonc.2008.12.017.
9. Dirix, P., Abbeel, S., Vanstraelen, B., Hermans, R., & Nuyts, S. (2009). Dysphagia after chemoradiotherapy for head-and-neck squamous cell carcinoma: Dose-effect relationships for the swallowing structures. *International Journal of Radiation Oncology, Biology, Physics, 75*(2), 385–392. https://doi.org/10.1016/j.ijrobp.2008.11.041.
10. Levendag, P. C., Teguh, D. N., Voet, P., van der Est, H., Noever, I., de Kruijf, W. J. M., …, Heijmen, B. J. (2007). Dysphagia disorders in patients with cancer of the oropharynx are significantly affected by the radiation therapy dose to the superior and middle constrictor muscle: A dose-effect relationship. *Radiotherapy and Oncology, 85*(1), 64–73. https://doi.org/10.1016/j.radonc.2007.07.009.
11. Christianen, M. E. M. C., Schilstra, C., Beetz, I., Muijs, C. T., Chouvalova, O., Burlage, F. R., …, Langendijk, J. A. (2012). Predictive modelling for swallowing dysfunction after primary (chemo)radiation: Results of a prospective observational study. *Radiotherapy and Oncology, 105*(1), 107–114. https://doi.org/10.1016/j.radonc.2011.08.009.
12. Feng, F. Y., Kim, H. M., Lyden, T. H., Haxer, M. J., Worden, F. P., Feng, M., …, Eisbruch, A. (2010). Intensity-modulated chemoradiotherapy aiming to reduce dysphagia in patients with oropharyngeal cancer: clinical and functional results. *Journal of Clinical Oncology, 28*(16), 2732–2738. https://doi.org/10.1200/JCO.2009.24.6199.
13. Maccomb, W. S., & Fletcher, G. H. (1957). Planned combination of surgery and radiation in treatment of advanced primary head and neck cancers. *The American Journal of Roentgenology, Radium Therapy, and Nuclear Medicine, 77*(3), 397–414.
14. Nuyts, S., Lambrecht, M., Duprez, F., Daisne, J.-F., Van Gestel, D., Van den Weyngaert, D., … De Neve, W. (2013). Reduction of the dose to the elective neck in head and neck squamous cell carcinoma, a randomized clinical trial using intensity modulated radiotherapy (IMRT). Dosimetrical analysis and effect on acute toxicity. *Radiotherapy and Oncology*, *109*(2), 323–329. https://doi.org/10.1016/j.radonc.2013.06.044
15. Nevens, D., Duprez, F., Daisne, J. F., Dok, R., Belmans, A., Voordeckers, M., … Nuyts, S. (2017). Reduction of the dose of radiotherapy to the elective neck in head and neck squamous cell carcinoma; a randomized clinical trial. Effect on late toxicity and tumor control. *Radiotherapy and Oncology*, *122*(2), 171–177. https://doi.org/10.1016/j.radonc.2016.08.009
16. Nevens, D., Duprez, F., Daisne, J.-F., Schatteman, J., Van der Vorst, A., De Neve, W., et al. (2017). Recurrence patterns after a decreased dose of 40Gy to the elective treated neck in head and neck cancer. *Radiotherapy and Oncology, 123*(3), 419–423. https://doi.org/10.1016/j.radonc.2017.03.003.
17. Deschuymer, S., Nevens, D., Duprez, F., Daisne, J.-F., Dok, R., Laenen, A., …, Nuyts, S. (2020). Randomized clinical trial on reduction of radiotherapy dose to the elective neck in head and neck squamous cell carcinoma; update of the long-term tumor

outcome. *Radiotherapy and Oncology, 143*, 24–29. https://doi.org/10.1016/j.radonc.2020.01.005.
18. Grégoire, V., Eisbruch, A., Hamoir, M., & Levendag, P. (2006). Proposal for the delineation of the nodal CTV in the node-positive and the post-operative neck. *Radiotherapy and Oncology, 79*(1), 15–20. https://doi.org/10.1016/j.radonc.2006.03.009.
19. Aaronson, N. K., Ahmedzai, S., Bergman, B., Bullinger, M., Cull, A., Duez, N. J., …, de Haes, J. C. (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*(5), 365–376.
20. Bjordal, K., Hammerlid, E., Ahlner-Elmqvist, M., de Graeff, A., Boysen, M., Evensen, J. F., …, Kaasa, S. (1999). Quality of life in head and neck cancer patients: Validation of the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-H&N35. *Journal of Clinical Oncology, 17*(3), 1008–1019. https://doi.org/10.1200/JCO.1999.17.3.1008.
21. Singer, S., Wollbrück, D., Wulke, C., Dietz, A., Klemm, E., Oeken, J., …, Schwarz, R. (2009). Validation of the EORTC QLQ-C30 and EORTC QLQ-H&N35 in patients with laryngeal cancer after surgery. *Head & Neck, 31*(1), 64–76. https://doi.org/10.1002/hed.20938.
22. Singer, S., Arraras, J. I., Chie, W.-C., Fisher, S. E., Galalae, R., Hammerlid, E., …, Hofmeister, D. (2013). Performance of the EORTC questionnaire for the assessment of quality of life in head and neck cancer patients EORTC QLQ-H&N35: A methodological review. *Quality of Life Research, 22*(8), 1927–1941. https://doi.org/10.1007/s11136-012-0325-1.
23. Fayers, P. M., Aaronson, N. K., Bjordal, K., Groenvold, M., Curran, D., Bottomley, A. & On behalf of the EORTC Quality of Life Group. (2001). The EORTC QLQ-C30 Scoring Manual (3rd Edition). Brussels: European Organisation for Research and Treatment of Cancer.
24. Cocks, K., King, M. T., Velikova, G., Martyn St-James, M., Fayers, P. M., & Brown, J. M. (2011). 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. *Journal of Clinical Oncology, 29*(1), 89–96. https://doi.org/10.1200/JCO.2010.28.0107.
25. Osoba, D., Rodrigues, G., Myles, J., Zee, B., & Pater, J. (1998). Interpreting the significance of changes in health-related quality-of-life scores. *Journal of Clinical Oncology, 16*(1), 139–144. https://doi.org/10.1200/JCO.1998.16.1.139.
26. Scott, N. W., Fayers, P. M., Aaronson, N. K., Bottomley, A., Graeff, A., Groenvold, M., Gundy, C., Koller, M., Petersen, M. A., Mirjam A. G., & Sprangers on behalf of the EORTC Quality of Life Group. (2008). EORTC QLQ-C30 Reference values manual. Brussels. *EORTC Quality of Life Group Publications*.
27. Bottomley, A., Tridello, G., Coens, C., Rolland, F., Tesselaar, M. E. T., Leemans, C. R., …, Lefebvre, J.-L. (2014). An international phase 3 trial in head and neck cancer: quality of life and symptom results: EORTC 24954 on behalf of the EORTC Head and Neck and the EORTC Radiation Oncology Group. *Cancer, 120*(3), 390–398. https://doi.org/10.1002/cncr.28392.
28. de Graeff, A., de Leeuw, J. R., Ros, W. J., Hordijk, G. J., Blijham, G. H., & Winnubst, J. A. (2000). Long-term quality of life of patients with head and neck cancer. *The Laryngoscope, 110*(1), 98–106. https://doi.org/10.1097/00005537-200001000-00018.
29. Curran, D., Giralt, J., Harari, P. M., Ang, K. K., Cohen, R. B., Kies, M. S., …, Bonner, J. A. (2007). Quality of life in head and neck cancer patients after treatment with high-dose radiotherapy alone or in combination with cetuximab. *Journal of Clinical Oncology, 25*(16), 2191–2197. https://doi.org/10.1200/JCO.2006.08.8005.
30. Vainshtein, J. M., Moon, D. H., Feng, F. Y., Chepeha, D. B., Eisbruch, A., & Stenmark, M. H. (2015). Long-term quality of life after swallowing and salivary-sparing chemo-intensity modulated radiation therapy in survivors of human papillomavirus-related oropharyngeal cancer. *International Journal of Radiation Oncology, Biology, Physics, 91*(5), 925–933. https://doi.org/10.1016/j.ijrobp.2014.12.045.
31. Meirovitz, A., Murdoch-Kinch, C. A., Schipper, M., Pan, C., & Eisbruch, A. (2006). Grading xerostomia by physicians or by patients after intensity-modulated radiotherapy of head-and-neck cancer. *International Journal of Radiation Oncology, Biology, Physics, 66*(2), 445–453. https://doi.org/10.1016/j.ijrobp.2006.05.002.
32. Kaae, J. K., Johnsen, L., Hansen, C. R., Kristensen, M. H., Brink, C., & Eriksen, J. G. (2019). Relationship between patient and physician-rated xerostomia and dose distribution to the oral cavity and salivary glands for head and neck cancer patients after radiotherapy. *Acta oncologica (Stockholm, Sweden), 58*(10), 1366–1372. https://doi.org/10.1080/0284186X.2019.1627413.
33. de Veij Mestdagh, P. D., Jonker, M. C. J., Vogel, W. V., Schreuder, W. H., Donswijk, M. L., Klop, W. M. C., et al. (2018). SPECT/CT-guided lymph drainage mapping for the planning of unilateral elective nodal irradiation in head and neck squamous cell carcinoma. *European Archives of Oto-Rhino-Laryngology, 275*(8), 2135–2144. https://doi.org/10.1007/s00405-018-5050-0.
34. de Veij Mestdagh, P. D., Schreuder, W. H., Vogel, W. V., Donswijk, M. L., van Werkhoven, E., van der Wal, J. E., …, Al-Mamgani, A. (2019). Mapping of sentinel lymph node drainage using SPECT/CT to tailor elective nodal irradiation in head and neck cancer patients (SUSPECT-2): A single-center prospective trial. *BMC Cancer, 19*(1), 1110. https://doi.org/10.1186/s12885-019-6331-8.
35. Daisne, J.-F., Installé, J., Bihin, B., Laloux, M., Vander Borght, T., Mathieu, I., et al. (2014). SPECT/CT lymphoscintigraphy of sentinel node(s) for superselective prophylactic irradiation of the neck in cN0 head and neck cancer patients: A prospective phase I feasibility study. *Radiation Oncology, 9*, 121. https://doi.org/10.1186/1748-717X-9-121.
36. Blanchard, P., Garden, A. S., Gunn, G. B., Rosenthal, D. I., Morrison, W. H., Hernandez, M., …, Frank, S. J. (2016). Intensity-modulated proton beam therapy (IMPT) versus intensity-modulated photon therapy (IMRT) for patients with oropharynx cancer—a case matched analysis. *Radiotherapy and Oncology, 120*(1), 48–55. https://doi.org/10.1016/j.radonc.2016.05.022.
37. Bottomley, A., Flechtner, H., Efficace, F., Vanvoorden, V., Coens, C., Therasse, P., et al. (2005). Health related quality of life outcomes in cancer clinical trials. *European Journal of Cancer, 41*(12), 1697–1709. https://doi.org/10.1016/j.ejca.2005.05.007.
38. Bottomley, A., Pe, M., Sloan, J., Basch, E., Bonnetain, F., Calvert, M., et al., …, Setting International Standards in Analyzing Patient-Reported Outcomes and Quality of Life Endpoints Data (SISAQOL) consortium. (2016). Analysing data from patient-reported outcome and quality of life endpoints for cancer clinical trials: A start in setting international standards. *The Lancet Oncology, 17*(11), e510–e514. https://doi.org/10.1016/S1470-2045(16)30510-1.
39. Pe, M., Dorme, L., Coens, C., Basch, E., Calvert, M., Campbell, A., Campbell, A., …, Setting International Standards in Analyzing Patient-Reported Outcomes and Quality of Life Endpoints Data Consortium (SISAQOL). (2018). Statistical analysis of patient-reported outcome data in randomised controlled trials

of locally advanced and metastatic breast cancer: A systematic review. *The Lancet Oncology, 19*(9), e459–e469. https://doi.org/10.1016/S1470-2045(18)30418-2.

40. Bottomley, A., Pe, M., Sloan, J., Basch, E., Bonnetain, F., Calvert, M., …, Coens, C. (2018). Moving forward toward standardizing analysis of quality of life data in randomized cancer clinical trials. *Clinical Trials, 15*(6), 624–630. https://doi.org/10.1177/1740774518795637.