Protocol compliance in a multicentric phase III trial investigating scheduled adaptive radiotherapy and dose painting in head and neck cancer

Leeuw, A.L.M.P. de; Giralt, J.; Tao, Y.G.; Benavente, S.; Nguyen, T.V.F.; Hoebers, F.J.P.; ... ; Hamming-Vrieze, O.

Radiotherapy and Oncology 202 (2025) 110612

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Original Article

Protocol compliance in a multicentric phase III trial investigating scheduled adaptive radiotherapy and dose painting in head and neck cancer

Anna Liza M.P. de Leeuw [a], Jordi Giralt [b,c], Yungan Tao [d], Sergi Benavente [b,c], Thanh-Vân F Nguyen [d], Frank J.P. Hoebers [e], Ann Hoeben [f], Chris H.J. Terhaard [g], Lip Wai Lee [h], Signe Friesland [i], Roel J.H.M. Steenbakkers [j], Lisa Tans [k], Simon R. van Kranen [a], Jeroen B. van de Kamer [a], Harry Bartelink [a], Coen R.N. Rasch [l], Jan-Jakob Sonke [a], Olga Hamming-Vrieze [a,*]

[a] Department of Radiation Oncology, The Netherlands Cancer Institute, Amsterdam, the Netherlands
[b] Department of Radiation Oncology, Hospital General Vall d'Hebron, Barcelona, Spain
[c] Vall d'Hebron Institute of Oncology, Barcelona, Spain
[d] Department of Radiation Oncology, Institut Gustave Roussy, Villejuif, France
[e] Department of Radiation Oncology (Maastro), GROW School for Oncology and Reproduction, Maastricht University Medical Centre+, Maastricht, the Netherlands
[f] Division of Medical Oncology, Department of Internal Medicine, GROW-School of Oncology and Developmental Biology Maastricht University Medical Center+, Maastricht, the Netherlands
[g] Department of Radiation Oncology, University Medical Center Utrecht, Utrecht, the Netherlands
[h] Department of Radiation Oncology, The Christie NHS Foundation Trust, Manchester, United Kingdom
[i] Department of Radiation Oncology, Karolinska Institute, Stockholm, Sweden
[j] Department of Radiation Oncology, University Medical Center Groningen, Groningen, the Netherlands
[k] Department of Radiation Oncology, Erasmus Medical Center, Rotterdam, the Netherlands
[l] Department of Radiation Oncology, Leiden University Medical Center, Leiden, the Netherlands

ARTICLE INFO

Keywords:
Protocol adherence
Quality assurance
Adaptive radiotherapy
Dose painting
Head and neck cancer

ABSTRACT

Purpose: To report on quality assurance (QA) and protocol adherence (PA) in a multicentre phase III trial for head and neck cancer, evaluate patterns of protocol deviations and investigate the effect of PA on study outcomes.
Methods: All 221 patients from the ARTFORCE trial (NCT01504815) were included in this study. Pre- and per-treatment QA measures included protocol guidelines, a dummy run, early case reviews and trial meetings. FDG-PET-guided dose painting and scheduled adaptive radiotherapy were reviewed in patients in the experimental arm (eRT). Patient and disease characteristics, as well as institutes' accrual rate and timing were examined for correlation with PA. Cox regression was used to determine the impact of PA on outcome.
Results: The dummy run was completed in all nine institutes and early case reviews were completed in five out of nine institutes that contributed 190 out of 221 patients. Among all patients randomized to eRT, 64 % had at least one deviation of the experimental trial components. Protocol deviations were significantly correlated with the institute patients were treated at (Cramer'sV 0.34–0.48). Despite early identification of institute-specific deviations in QA, these continued during the trial. No significant associations were seen between deviations and accrual timing or rate ($P \geq 0.26$). Within eRT, no significant relation was observed between experimental PA and locoregional control (LRC), the primary endpoint of the trial ($P \geq .15$).
Conclusions: Despite QA, protocol deviations persisted during the trial, which were mostly institute-specific. However, deviations of the experimental treatment strategy did not significantly impact LRC and therefore the trial conclusion.

* Corresponding author at: Department of Radiation Oncology, Netherlands Cancer Institute – Antoni van Leeuwenhoek Hospital, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands.
E-mail address: a.d.leeuw@nki.nl (O. Hamming-Vrieze).

https://doi.org/10.1016/j.radonc.2024.110612
Received 9 August 2024; Received in revised form 17 October 2024; Accepted 29 October 2024
Available online 9 November 2024

Introduction

Randomized controlled trials (RCTs) form a cornerstone in clinical research when investigating novel therapeutic strategies, because these are less susceptible to known and unknown bias than any other study designs. Potential bias, such as treatment adherence, is assumed to be similar between trial arms in RCTs, therefore washing out its absolute impact on comparative study outcomes. In radiotherapy trials, however, novel complex technologies are compared to technologies which are more routinely used in clinical practice. And despite careful design, major protocol deviations occur in 11–48 % of study participants, potentially causing bias and altering study results [1]. Planning quality and adherence to dose prescription have shown to significantly impact treatment efficacy [2]. Nevertheless, only half of RCTs in radiotherapy perform any pre-treatment quality assurance (QA), independent of technical complexity [3]. Studies evaluating protocol adherence within clinical trials treating head and neck cancer patients with definitive chemoradiation found that protocol deviations were not only associated with patient's outcome, but could also possibly mask a clinical benefit of the experimental treatment [4,5]. Considering the clinical impact of protocol adherence and the novelty of the technologies tested in radiotherapy RCTs, assuming comparable bias between trial arms in radiotherapy RCT studies might not be reasonable.

In the ARTFORCE RCT (NCT01504815) comparing conventional radiotherapy to a novel dose painting strategy with treatment adaptation, there are several factors that could contribute to a higher rate of protocol deviations in the experimental arm. Firstly, being at the forefront in biology-driven and adaptive radiotherapy, the experimental treatment strategy relied on novel planning strategies and additional labour efforts by including co-registration of FDG-PET/CT to planning CT for target volume delineation and demanding adaptation in the second week of treatment. Also, the trial took place across various institutes in Europe with different workflows, treatment planning systems and planning habits, which might have caused a heterogeneity in treatment given. Lastly, the unplanned extended accrual period from 3 to 7 years caused by slow accrual could have further compromised protocol adherence, especially in the more complex experimental arm.

Although QA has been incorporated in routine clinical practice to ensure treatment safety and effectiveness, QA procedures differ greatly in trial setting. To minimize heterogeneity within QA procedures and improve protocol adherence within trials, harmonized pre- and post-treatment QA procedures were defined by The Global Clinical Trial RTQA Harmonization Group [6]. These procedures include facility credentialing, technical audit, delineation exercise, dummy run, dosimetric evaluation, virtual phantom, case reviews and institutional visits [6]. Many of these harmonized QA procedures were utilized in our trial, but it remains uncertain if it has improved protocol compliance compared to historic protocol adherence rates. This paper will therefore report on utilized QA strategies and protocol adherence within the trial and identify contributing factors that may have led to protocol deviations, such as the extended accrual period. Lastly, the clinical impact of protocol deviations will be investigated.

Materials and methods

Study population and treatment protocol

A detailed description of the treatment protocol, patient characteristics and the primary outcome results were recently described by De Leeuw et al.[7] Between 2012 and 2019, 221 eligible patients were accrued at nine institutes in France, the Netherlands, Spain, Sweden and the United Kingdom. Patients eligible for registration had stage III-IV, T3-4, N0-1, M0 (6th edition AJCC) HNSCC in the oral cavity, oropharynx or hypopharynx and were randomized (1:1) to receive either conventional radiotherapy (cRT, n = 112) or the experimental adaptive and FDG-PET-guided dose painting (eRT, n = 109), delivered in 35 fractions over a course of seven weeks. Systemic treatment for both arms consisted of three-cycle high-dose cisplatin (100 mg/m^2).

Imaging and planning protocol

FDG-PET scans were scheduled within two weeks prior to treatment initiation and at the end of the second week of treatment for all patients independently of treatment arm. The pre-treatment FDG-PET imaging could be used clinically to guide delineation of tumor and pathological lymph nodes as is common clinical practice. In eRT, the pre-treatment FDG-PET scan was further used to guide dose painting. Replanning was performed using a planning CT scan, adjusting for clear anatomical changes using distinct landmarks, including vessels, bones and air. Although an FDG-PET scan was made at replanning as well, target volumes were not supposed to be adjusted for changes in metabolic activity. Gross target volumes of the primary tumor (GTVp) and metabolic subvolume (GTVp-PET) and corresponding planning target volumes (PTVp, PTVp-PET) of the initial plan were therefore rigidly propagated to the per-treatment planning CT scan and used for replanning. Although replanning was scheduled for the third week of treatment (fx 11–15), replaning performed in the third or fourth week (fx 11–20) of treatment and used for more than one fractions were considered scheduled replanning. Additional ad hoc replanning due to anatomical changes was allowed. In the cRT arm, the PTVp received a conventional homogeneous dose distribution and no replanning was scheduled.

Quality assurance measures

QA measures were introduced on multiple levels to improve uniformity, protocol adherence and timely detect relevant unforeseen variances between participating centers. First of all, in the protocol, several guidelines were stated to improve uniformity between institutes. FDG-PET systems had to be EANM compliant and pre-treatment FDG-PET scans were supposed to be less than two weeks old before the start of treatment. An extensive delineation protocol was described since no international delineation guidelines for organs at risk (OAR) were available at trial initiation. Planning guidelines, dose constraints and treatment course were standardized across institutes through detailed protocol descriptions. To maintain a treatment course of 47–49 days, it was advised to compensate fractions with a minimum of 6 h between fractions. As a next step, before the start of accrual, a dummy run using one FDG-PET/CT and planning CT was performed at the participating centers and evaluated by the principle investigator (PI), a radiation oncologist, and a colleague physicist to examine adherence to delineation protocols, evaluate co-registration of FDG-PET/CT with the planning CT and evaluate FDG-PET guided dose planning quality by means of dosimetry (Supplement 1). Furthermore, after treatment of the first three patients, an on-site early case review was organized to evaluate protocol adherence and data collection (Supplement 2). These reviews were also performed at the center of the PI with involvement of other participating centers. The review of patients treated at the PI's institute was performed by two other participating centers. Additional central monitoring of all patient data and serious adverse events was performed by an independent data center to ensure accurate data collection, log protocol deviations with their justifications and monitor trial safety. Lastly, regular physical and teleconference meetings with the principal investigators were conducted during the trial to evaluate study accrual progress and discuss uniformity.

Evaluated measures

In this study we evaluated if all institutes participated in the dummy run and early case review of the first three patients, as well as the results of these actions. Furthermore, a selection of the scored items in the case report forms were used to retrospectively score protocol adherence in the total cohort. Additionally, protocol adherence at replanning was

retrospectively evaluated for patients randomized to the experimental arm. All evaluated measures are listed in Table 1.

Statistical analysis

Protocol deviations conceivable in both arms were compared within arms using chi-square tests. To explore contributing factors to protocol deviations, Cramer's V correlation coefficients were calculated for deviations, treating center and disease characteristics. Statistical significance of these correlations were evaluated using chi-square tests. The relationship between protocol deviations and patients' outcome were evaluated using univariable Cox regression models. These outcome measures consisted of the trial's primary (locoregional control) and secondary endpoints (progression-free and overall survival). Lastly, frequency of deviations across participating centers and throughout the accrual period were explored using Pearsons correlation and linear models to evaluate the influence of accrual rate and passage of time on protocol adherence. No multiple testing correction was performed due to the exploratory nature of this study and the primary objective to identify protocol deviations potentially affecting study outcomes and needing further investigation.

Results

Six institutes were involved at trial initiation. Throughout the study period, three additional European institutes participated in the trial, resulting in a total of nine participating institutes. The dummy run was successfully completed in all nine institutes. The observed deviations primarily concerned delineations of OAR and the anatomical correction of clinical target volumes.

Early case reviews of the first three patients were performed in the first five institutes that started patient accrual, these institutes contributed 190 out of 221 patients. In these early case reviews, protocol deviations in at least one of the three evaluated patients were observed at three institutes, which were discussed as feedback (Fig. 1A). Pre-treatment FDG-PET was outdated in two institutes, prolonged treatment course was observed in two institutes, late scheduled replanning was observed in two institutes and the non-rigid adaptation of GTVp-PET based on per-treatment FDG-PET was observed in one institute.

Protocol deviations are stated in Table 2. Evaluated measures for which the treatment arms can be compared showed equal deviations across arms. Renal and hematological toxicities were most common causes for incomplete systemic treatment (Supplement 3). In total, 12 % of patients experienced prolonged treatment course with a median delay of two days (min–max;1–8). In only 30 % of these early cases a specific reason for delay was reported; 15 % due to adverse events and 15 % for logistical or technical reasons.

Protocol prescribed that replanning should have been scheduled within fraction 11–15 for patients allocated to the eRT arm, but it was performed beyond this timeframe in 26 % of the eRT arm (Fig. 1C). Additional ad hoc replanning was performed in 17 % of patients in the eRT arm (Fig. 2). In the cRT arm, ad hoc replanning was performed in 36 % of patients. 5 % of patients in the cRT had more than one ad hoc adaptations.

The retrospective replanning review included 86 of 98 (88 %) patients with scheduled replanning and for whom planning DICOM files were available for evaluation. In total, 27 % of patients in the eRT arm had non-compliant adaptation of the target volumes. Non-adherent adaptation of GTVp was observed in 17 % of patients, while non-adherent GTVp-PET adaptations were observed in 15 % of patients. 6 % of the eRT arm had their GTVp-PET clearly biologically adapted to a per-treatment FDG-PET/CT scan.

Correlation maps between disease characteristics and protocol adherence revealed significant, although limited, correlation between premature end of radiotherapy (<35 fractions) and disease site (0.24, $P \leq 0.01$) and between N-stage and timing of scheduled adaptation (0.17, $P = 0.04$). All other significant correlations were found involving study institute (Fig. 3), meaning that deviations rates significantly varied within institutes. The most common protocol deviation within institutes was an incomplete systemic treatment with an average of 47 % among all patients, but the range within institutes varied between 25–100 %. Among the experimental protocol items, non-compliant adaptation of target volumes was the most common deviation (27 % total cohort, range 0–36 %). From all three institutions that were found to have deviated from protocol during the early case reviews of the first three patients, all continued to deviate from the same items throughout the trial period (Fig. 1).

Overall, 64 % of patients in the eRT arm had at least one deviation from the experimental elements of the protocol. The number of deviations did not significantly change with year of accrual (Pearson correlation 0.18, $P = 0.70$) or with institutes' accrual rate (0.32, $P = 0.26$).

None of the protocol items were significantly correlated to all three outcome variables (Table 3). Administering ≤ 1 of the intended cisplatin course (hazard ratio (HR) 2.66, 95 % confidence interval (CI) (1.01–7.00), $p = 0.05$) and prolonged treatment duration (HR 2.39, 95 %CI 1.25–4.58, $P \leq 0.01$) were deviations that were associated with inferior locoregional control. Prolonged treatment duration was also significantly associated with poor progression-free survival (PFS). PFS (2.01 ,1.07–3.77, $P = 0.03$) and overall survival (OS, 2.28, 1.14–4.54, $P = 0.02$) were significantly inferior for patients with late scheduled replanning. Lastly, having an FDG-PET/CT scan made ≥ 2 weeks prior to treatment was significantly correlated with poor OS (2.14, 1.08–4.23, $P = 0.03$).

Table 1
Protocol adherence measures.

Dummy run performed	Per institute
Early case review first three included patients performed	Per institute
FDG-PET pre-treatment $\leq$ 2 weeks old	Both arms
Treatment time of < 50 days for conventional and < 43 days for accelerated radiotherapy *	Both arms
Received 35 fractions	Both arms
Complete systemic treatment	Both arms
Scheduled replanning performed	Experimental arm
Scheduled replanning performed in time (week 3, fraction 11–15)	Experimental arm
Target volumes at replanning were only adapted for clear boundaries i.e. in air, vessels or bone.	Experimental arm

* In the early phase of the trial, patients received six fractions (accelerated) instead of five fractions per week (conventional).

Discussion

In a randomized phase III trial exploring FDG-PET-guided dose painting and scheduled replanning, this study reported on quality assurance (QA) measures and protocol adherence (PA). This study further investigated factors contributing to protocol deviations and the potential impact of these on study outcomes. We observed that the chance of having a deviation was primarily determined by the institute patients were treated at. We further observed that institute-specific variation in PA could in part already be detected in early case reviews and these deviations mostly persisted during treatment. The portion of patients within the experimental arm with protocol deviations of any of the experimental component therefore remained high (64 %). However, none of the experimental protocol deviations significantly influenced the primary endpoint, locoregional control. Out of the secondary outcomes, late scheduled replanning after the third week of treatment was associated with inferior progression-free and overall survival compared to scheduled replanning. Also, a delay of more than two weeks between the FDG-PET/CT scan acquisition and the start of treatment resulted in significantly poorer OS.

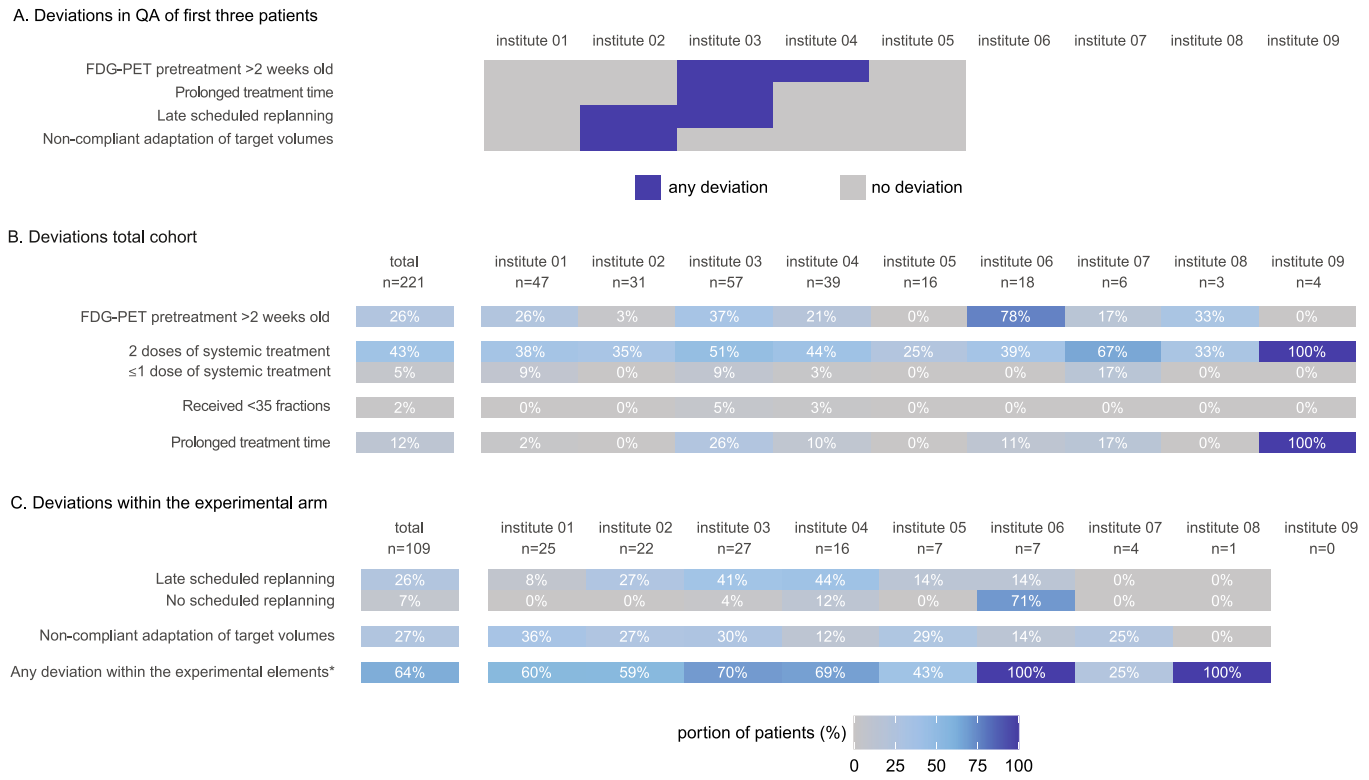

Fig. 1. A. Deviations (dark blue) that were observed in at least 1 patient at the quality assurance (QA) evaluations of the first three patients. This QA was not performed at institute 6 until 9. B. Frequency of protocol deviations (illustrated with a color gradient) across institutes within the total cohort. C. Frequency of protocol deviations across institutes within the experimental arm. No patients were allocated to the experimental arm at institute 9, which is why these cells are missing. Portion of deviations within the experimental elements (*) included outdated pretreatment FDG-PET, late or no scheduled replanning and non-compliant adaptation of target volumes. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Protocol deviations according to trial arm. eRT; experimental trial arm, cRT; conventional trial arm.

	eRT (N = 109)	cRT (N = 112)	Chi square P-value
FDG-PET pre-treatment > 2 weeks old	25 (25 %)	31 (28 %)	0.62
Prolonged treatment time of ≥ 50 days for conventional and ≥ 43 days for accelerated radiotherapy	10 (9 %)	17 (15 %)	0.40
Received < 35 fractions	2 (2 %)	2 (2 %)	0.98
Incomplete systemic treatment			
• *2 out of 3 doses*	48 (44 %)	47 (42 %)	0.67
• *≤1 out of 3 doses*	4 (4 %)	7 (6 %)	
No scheduled replanning	8 (7 %)	*n/a*	*n/a*
Late scheduled replanning (after week 3, fraction > 15)	28 (26 %)	*n/a*	*n/a*
Non-compliant adaptation of target volumes	29 (27 %)	*n/a*	*n/a*

Previous studies investigating PA in radiotherapy trials treating head and neck cancer patients found that deviations have the potential to compromise the results of a clinical trial [4,5]. Similar to our findings, incomplete chemotherapy treatment[8] and prolongued treatment duration[9,10] were associated with poorer tumor control and survival. For the latter, the repopulation effect has been considered as a possible underlying mechanism. However, adverse events or poor patient condition could potentially be a confounder as they may both lead to treatment delay and also impact patient's outcome. In this study, in a minority of cases with prolongued treatment a clear reason for delay was reported, of which half for technical or logistic reasons. Although the confounding effect of adverse events or patients' condition remains uncertain, these findings suggest that the repopulation effect might have contributed to the found association between prolongued treatment and clinical outcome. These results therefore stress the importance of compensating missed treatment days.

Other notable findings in this study were that late scheduled replanning and an extended period between pre-treatment FDG-PET/CT scan and treatment initiation were associated with inferior PFS and OS. These associations are most likely a result of patient selection. The delay in treatment initiation after FDG-PET/CT or in scheduled replanning might have been a consequence of more complex anatomy and/or tumor changes requiring additional examination that depreciate outcome, such as weight loss[11,12] or tumor progression. Another possible explanation for the inferior outcomes in patients with an increased period between FDG-PET/CT and treatment is understaging and consequent undertreatment of patients, leading to inferior disease response and outcome.

In general, QA remains poorly addressed in radiotherapy trials and high levels of protocol deviations are observed [1,2,13,14]. The strategy of combining dummy runs with prospective case reviews has shown to improve patient' outcome cost-effectively [15]. Although both dummy runs and case reviews were utilized in the present study, high levels of protocol deviations were observed with a majority of patients in this trial having at least one protocol deviation. A factor contributing to these protocol deviations may be challenges in implementing novel radiotherapeutic techniques[16], but also the high additional burden on staff caused by replanning [17]. Peters[4] and Zhong et al.[5] observed that deviations were more common in institutes with low patient accrual rates, presumably due to less experience with the experimental protocol. Contrary to these findings, we did not find a significant correlation between accrual rate and protocol deviations. Nevertheless, we did identify a noteworthy pattern of recurring institute-specific deviations

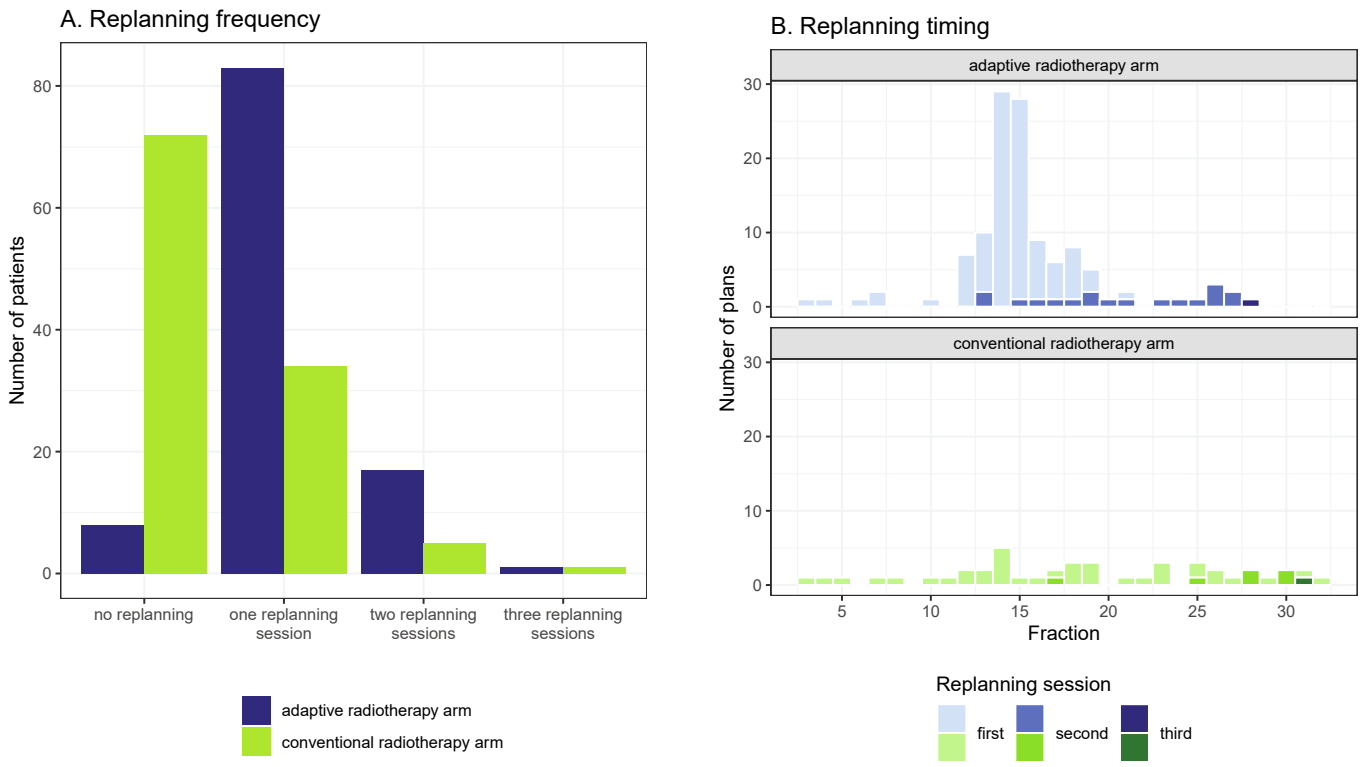

Fig. 2. Replanning frequency across trial arm (A) and replanning timing (B). In the conventional radiotherapy arm (green) only ad hoc replanning was performed. In the adaptive radiotherapy arm, treatment adaptation was scheduled during the third week (within fraction 10–15) of treatment, but additional ad hoc replanning was allowed. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

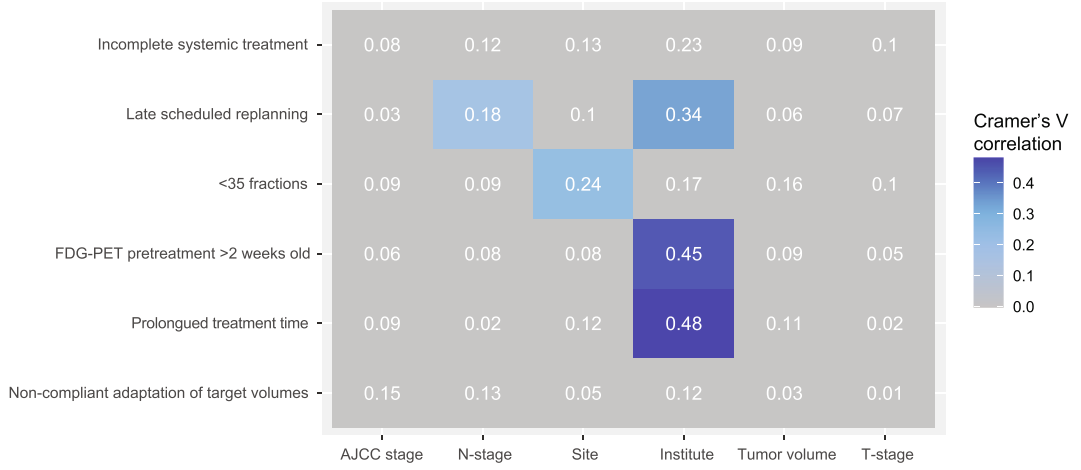

Fig. 3. Cramer's V correlation matrix of disease characteristics and institute (x-axis) and protocol adherence items (y-axis). Gradient color as well as the values within the cells indicate the correlation coefficient of significant correlations. A value close to 1 indicates a correlation and a value closer to 0 indicates no correlation. The correlations of grey cells were insignificant (Chi-square test).

already being apparent in early QA. Therefore, we believe that QA could have been leveraged by conducting follow-up on these early deviations within the institutes.

We also found that institutes differed in their approach to account for the changing anatomy during treatment, as deformable or FDG-PET-driven approaches were considered protocol deviations in this study. The variation of target volume delineation and an absence of clear disease borders allowing for individualized localization of disease during treatment and consequent adjustments of the target volumes are unresolved issues in the adaptive radiotherapy field [18–20]. Currently, most QA concern a dosimetric assessment of established delineations as there is much less consensus on target volume delineation, let alone to what extent they should be adapted to changing anatomy and biology during treatment. Given the variation of additional or unplanned adaptations between institutes and its impact on patient outcome, it is crucial for future studies in radiotherapy to establish consensus on how treatment adaptation should be managed and report on variations.

A notable limitation of this study is that in four out of nine institutes the early case review was omitted, which can partly be explained by the slow and low accrual rate in these institutes. We expect this to have a minimal effect on the trial results as the large majority (86 %) of patients were treated at the institutes that did perform early case reviews. Another limitation of this study was the exclusive collection of digital imaging and planning data, and the absence of planned dose data

Table 3
Associations between protocol deviations and disease or overall outcome. Analyses independent of arm are performed against the per-protocol reference. Experimental elements (*) are evaluated against the reference of per-protocol scheduled replanning within the experimental arm. HR; hazard ratio, CI; confidence interval.

		Locoregional control		Progression-free survival		Overall survival	
	n	*HR (95 %CI)*	*P-value*	*HR (95 %CI)*	*P-value*	*HR (95 %CI)*	*P-value*
FDG-PET pre-treatment < 2 weeks old*	25	1.22 (0.47–3.16)	0.68	1.55 (0.82–2.91)	0.18	2.14 (1.08–4.23)	0.03
Incomplete systemic treatment							
• ≤1 out of 3 doses	13	2.66 (1.01–7.00)	0.05	1.63 (0.70–3.82)	0.26	1.27 (0.45–3.56)	0.65
• 2 out of 3 doses	95	1.11 (0.62–1.98)	0.73	0.97 (0.63–1.48)	0.87	0.97 (0.61–1.55)	0.90
Treatment time of ≥ 50 days for conventional and ≥ 43 days for accelerated radiotherapy	27	2.39 (1.25–4.58)	≤0.01	1.77 (1.03–3.04)	0.04	1.59 (0.85–2.95)	0.15
No scheduled replanning*	8	1.75 (0.39–7.82)	0.47	1.99 (0.69–5.78)	0.20	2.31 (0.67–7.96)	0.18
Late scheduled replanning (after week 3, fraction > 15)*	28	1.76 (0.69–4.48)	0.24	2.01 (1.07–3.77)	0.03	2.28 (1.14–4.54)	0.02
Non-compliant adaptation of target volumes*	29	2.02 (0.78–5.23)	0.15	1.68 (0.86–3.28)	0.13	1.52 (0.70–3.29)	0.29

calculated on an institutional level. This resulted in dose prescription and dose-volume calculation differences across planning systems and institutes, hampering our ability to strictly evaluate dosimetric PA [11,21,22]. Additionally, this trial encountered challenges implementing an imaging acquisition system, which lead to delayed data collection causing missing imaging data. Future trials should aim to overcome these challenges by prospective collection of digital imaging and planning data, as well as planning data as calculated by institutes. It should be noted, however, that this trial was initiated in 2012 and suffered from limited data collection possibilities.

The observed institute-specific variation in PA necessitate more active and continuous QA to avoid unreliable data. However, additional QA in trials can be costly and increase workload, especially when these lead to interventions in the workflow [15]. Directing QA frequency or focus based on early QA results could help manage workload and costs. While this strain on resources may have limited the level of QA in previous trials, innovations in automating QA procedures has opened opportunities to improve QA efficiency. Ideally, with automated and accelerated QA processes, QA should be performed before clinical use for all patients in technical challenging trials to prevent protocol violations [23]. Anticipating on deviations, determining their impact on study results, and predefining criteria will, in that case, become of increasing importance. In addition to the existing predefined dose parameters and advancements in adaptive radiotherapy, future trials should establish a consensus on treatment adaptation with their participating institutes. Acknowledging and describing treatment variation caused by protocol deviations or institutional differences and evaluating their impact on outcome can guide future investigators in establishing evidence based criteria for trial protocols.

In conclusion, protocol deviations persisted during the trial despite being evident in early QA and seemed mostly correlated with the institute patients were treated at, highlighting the need for more guidance and directed QA. Within the ARTFORCE trial specifically, the impact of protocol deviations on trial results are expectedly limited considering the absence of a significant relation between experimental protocol deviations and locoregional control.

Funding information

This work was funded by the European commission FP7 project (Nr 257144).

CRediT authorship contribution statement

Anna Liza M.P. de Leeuw: Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation. **Jordi Giralt:** Writing – review & editing, Resources, Investigation. **Yungan Tao:** Writing – review & editing, Resources, Investigation. **Sergi Benavente:** Writing – review & editing, Resources, Investigation. **Thanh-Vân F Nguyen:** Writing – review & editing, Resources, Investigation. **Frank J. P. Hoebers:** Writing – review & editing, Resources, Investigation. **Ann Hoeben:** Writing – review & editing, Resources, Investigation. **Chris H. J. Terhaard:** Writing – review & editing, Resources, Investigation. **Lip Wai Lee:** Writing – review & editing, Resources, Investigation. **Signe Friesland:** Writing – review & editing, Resources, Methodology. **Roel J. H.M. Steenbakkers:** Writing – review & editing, Resources, Investigation. **Lisa Tans:** Writing – review & editing, Resources, Investigation. **Simon R. van Kranen:** Writing – review & editing, Supervision, Resources. **Jeroen B. van de Kamer:** Writing – review & editing, Investigation. **Harry Bartelink:** Writing – review & editing, Supervision, Resources, Project administration, Investigation, Funding acquisition, Conceptualization. **Coen R.N. Rasch:** Writing – review & editing, Supervision, Resources, Investigation, Funding acquisition, Conceptualization. **Jan-Jakob Sonke:** Writing – review & editing, Supervision, Methodology, Investigation, Conceptualization. **Olga Hamming-Vrieze:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: [Anna Liza M.P. de Leeuw; PhD-student at the Antoni van Leeuwenhoek hospital. Jordi Giralt; Radiation oncologist at the Vall d'Hebron Unniversity Hospital. Yungan Tao; Radiation oncologist at the Institut Gustave roussy. Sergi Benavente; Radiation oncologist at the Vall d'Hebron Unniversity Hospital. Thanh-Vân France Nguyen; Radiation oncologist at the Institut Gustave roussy. Frank J.P. Hoebers; Radiation oncologist at the Maastro Center. Ann Hoeben; Clinical oncologist at the Maastro Center. Chris H.J. Terhaard; retired radiation oncologist at the University Medical Center Utrecht. Lip Wai Lee; Radiation oncologist at the Christie NHS Foundation Trust. Signe Friesland; Radiation oncologist at the Karolinska Comprehensive Cancer Center. Roel J.H.M. Steenbakkers; Radiation oncologist at the University Medical Center Groningen. Lisa Tans; Radiation oncologist at the Erasmus Medical Center. Simon R. van Kranen; Postdoc at the Antoni van Leeuwenhoek hospital. Jeroen B van de Kamer; Medical physicist at the Antoni van Leeuwenhoek hospital. Harry Bartelink; retired radiation oncologist at the Antoni van Leeuwenhoek hospital. Coen R.N. Rasch; Professor at Leiden University; Radiation oncologist at the Leiden University Medical Center; Jan-Jakob

Sonke; Professor by special appointment at University of Amsterdam; senior group leader at Antoni van Leeuwenhoek hospital; ownership of patent/license fees/copyright at Elekta AB; member of the international advisory board at Physics, medicine and biology; editorial board member of Physics & Imaging in Radiation Oncology. Olga Hamming-Vrieze; Radiation oncologist at Antoni van Leeuwenhoek hospital].

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.radonc.2024.110612.

References

[1] Weber DC, Tomsej M, Melidis C, Hurkmans CW. QA makes a clinical trial stronger: Evidence-based medicine in radiation therapy. Radiother Oncol 2012;105:4–8. https://doi.org/10.1016/J.RADONC.2012.08.008.

[2] Moran JM, Molineu A, Kruse JJ, et al. Executive summary of AAPM Report Task Group 113: Guidance for the physics aspects of clinical trials. J Appl Clin Med Phys 2018;19:335–46. https://doi.org/10.1002/acm2.12384.

[3] Corrigan KL, Kry S, Howell RM, et al. The radiotherapy quality assurance gap among phase III cancer clinical trials. Radiother Oncol 2022;166:51–7. https://doi.org/10.1016/j.radonc.2021.11.018.

[4] Peters LJ, O'Sullivan B, Giralt J, et al. Critical impact of radiotherapy protocol compliance and quality in the treatment of advanced head and neck cancer: Results from TROG 02.02. J Clin Oncol 2010;28:2996–3001. https://doi.org/10.1200/JCO.2009.27.4498.

[5] Zhong H, Men K, Wang J, et al. The impact of clinical trial quality assurance on outcome in head and neck radiotherapy treatment. *Front.* Oncol 2019;9:1–7. https://doi.org/10.3389/fonc.2019.00792.

[6] Melidis C, Bosch WR, Izewska J, et al. Global harmonization of quality assurance naming conventions in radiation therapy clinical trials. Int J Radiat Oncol Biol Phys 2014;90:1242–9. https://doi.org/10.1016/j.ijrobp.2014.08.348.

[7] De LA, Giralt J, Tao Y, et al. A multicentric randomized controlled phase III trial of adaptive and 18F-FDG-PET-guided dose-redistribution in locally advanced head and neck squamous cell carcinoma (ARTFORCE). Radiother Oncol 2024;196:110281. https://doi.org/10.1016/j.radonc.2024.110281.

[8] Lacas B, Carmel A, Landais C, et al. Meta-analysis of chemotherapy in head and neck cancer (MACH- NC): An update on 107 randomized trials and 19,805 patients, on behalf of MACH-NC Group HHS Public Access. Radiother Oncol 2021;156:281–93. https://doi.org/10.1016/j.radonc.2021.01.013.

[9] Xiang M, Gensheimer MF, Pollom EL, et al. Prolongation of definitive head and neck cancer radiotherapy: Survival impact and predisposing factors. Radiother Oncol 2021;156:201–8. https://doi.org/10.1016/j.radonc.2020.12.025.

[10] Wuthrick EJ, Zhang Q, Machtay M, et al. Institutional clinical trial accrual volume and survival of patients with head and neck cancer. J Clin Oncol 2015;33:156–64. https://doi.org/10.1200/JCO.2014.56.5218.

[11] Langius JAE, Bakker S, Rietveld DHF, et al. Critical weight loss is a major prognostic indicator for disease-specific survival in patients with head and neck cancer receiving radiotherapy. Br J Cancer 2013;109:1093–9. https://doi.org/10.1038/bjc.2013.458.

[12] Leoncini E, Ricciardi W, Cadoni G, et al. Adult height and head and neck cancer: A pooled analysis within the INHANCE Consortium. Head Neck 2014;36:1391. https://doi.org/10.1002/HED.

[13] Draeger E, Sawant A, Johnstone C, et al. A dose of reality: how 20 years of incomplete physics and dosimetry reporting in radiobiology studies may have contributed to the reproducibility crisis. Int J Radiat Oncol Biol Phys 2020;106:243–52. https://doi.org/10.1016/j.ijrobp.2019.06.2545.

[14] McDowell LJ, Corry J. A call to arms: radiation therapy quality assurance in the next generation of clinical trials. Int J Radiat Oncol Biol Phys 2018;102:1590–1. https://doi.org/10.1016/j.ijrobp.2018.07.2001.

[15] Weber DC, Hurkmans CW, Melidis C, et al. Outcome impact and cost-effectiveness of quality assurance for radiotherapy planned for the EORTC 22071–24071 prospective study for head and neck cancer. Radiother Oncol 2014;111:393–9. https://doi.org/10.1016/j.radonc.2014.04.015.

[16] Eaton DJ. Why all radiotherapy planning studies are wrong but some are useful. Med Dosim 2021;46:209–11. https://doi.org/10.1016/j.meddos.2020.12.002.

[17] Gouw ZAR, La Fontaine MD, Vogel WV, van de Kamer JB, Sonke JJ, Al-Mamgani A. Single-center prospective trial investigating the feasibility of serial FDG-PET guided adaptive radiation therapy for head and neck cancer. Int J Radiat Oncol Biol Phys 2020;108:960–8. https://doi.org/10.1016/j.ijrobp.2020.04.030.

[18] Cox S, Cleves A, Clementel E, Miles E, Staffurth J, Gwynne S. Impact of deviations in target volume delineation – Time for a new RTQA approach? Radiother Oncol 2019;137:1–8. https://doi.org/10.1016/j.radonc.2019.04.012.

[19] Hamming-Vrieze O, van Kranen SR, Heemsbergen WD, et al. Analysis of GTV reduction during radiotherapy for oropharyngeal cancer: Implications for adaptive radiotherapy. Radiother Oncol 2017;122:224–8. https://doi.org/10.1016/j.radonc.2016.10.012.

[20] Ligtenberg H, Jager EA, Caldas-Magalhaes J, et al. Modality-specific target definition for laryngeal and hypopharyngeal cancer on FDG-PET CT and MRI. *Radiother Oncol* 2017;123:63–70. https://doi.org/10.1016/j.radonc.2017.02.005.

[21] Hansen CR, Tadic T, McNiven A, et al. 191 dose prescription variability in oropharynx cancer radiotherapy. Radiother Oncol 2024;192:S225–8. https://doi.org/10.1016/s0167-8140(24)00538-3.

[22] Bosse C, Narayanasamy G, Saenz D, et al. Dose calculation comparisons between three modern treatment planning systems. J Med Phys 2020;45:143–7. https://doi.org/10.4103/jmp.JMP_111_19.

[23] Vandewinckele L, Claessens M, Dinkla A, et al. Overview of artificial intelligence-based applications in radiotherapy: Recommendations for implementation and quality assurance. Radiother Oncol 2020;153:55–66. https://doi.org/10.1016/j.radonc.2020.09.008.