Physics and Imaging in Radiation Oncology 28 (2023) 100489

Contents lists available at ScienceDirect

Physics and Imaging in Radiation Oncology

journal homepage: www.sciencedirect.com/journal/physics-and-imaging-in-radiation-oncology

Original Research Article

A robust semi-automatic delineation workflow using denoised diffusion weighted magnetic resonance imaging for response assessment of patients with esophageal cancer treated with neoadjuvant chemoradiotherapy

Robin den Boer [a,1], Kelvin Ng Wei Siang [b,c,1], Mandy Yuen [a], Alicia Borggreve [a], Ingmar Defize [a], Astrid van Lier [a], Jelle Ruurda [d], Richard van Hillegersberg [d], Stella Mook [a], Gert Meijer [a,*]

[a] University Medical Center Utrecht, Department of Radiation Oncology, Utrecht, The Netherlands
[b] Erasmus MC Cancer Institute, University Medical Center Rotterdam, Department of Radiotherapy, Rotterdam, The Netherlands
[c] Holland Proton Therapy Center, Department of Medical Physics & Informatics, Delft, The Netherlands
[d] University Medical Center Utrecht, Department of Surgery, Utrecht, The Netherlands

ARTICLE INFO

Keywords:
diffusion weighted MRI
Esophageal cancer
Response prediction
Neoadjuvant chemoradiotherapy
Automatic workflow
Imaging biomarker

ABSTRACT

Background and Purpose: Diffusion weighted magnetic resonance imaging (DW-MRI) can be prognostic for response to neoadjuvant chemotherapy (nCRT) in patients with esophageal cancer. However, manual tumor delineation is labor intensive and subjective. Furthermore, noise in DW-MRI images will propagate into the corresponding apparent diffusion coefficient (ADC) signal. In this study a workflow is investigated that combines a denoising algorithm with semi-automatic segmentation for quantifying ADC changes.

Materials and Methods: Twenty patients with esophageal cancer who underwent nCRT before esophagectomy were included. One baseline and five weekly DW-MRI scans were acquired for every patient during nCRT. A self-supervised learning denoising algorithm, Patch2Self, was used to denoise the DWI-MRI images. A semi-automatic delineation workflow (SADW) was next developed and compared with a manually adjusted workflow (MAW). The agreement between workflows was determined using the Dice coefficients and Brand Altman plots. The prognostic value of ADC_{mean} increases (%/week) for pathologic complete response (pCR) was assessed using c-statistics.

Results: The median Dice coefficient between the SADW and MAW was 0.64 (interquartile range 0.20). For the MAW, the c-statistic for predicting pCR was 0.80 (95% confidence interval (CI):0.56–1.00). The SADW showed a c-statistic of 0.84 (95%CI:0.63–1.00) after denoising. No statistically significant differences in c-statistics were observed between the workflows or after applying denoising.

Conclusions: The SADW resulted in non-inferior prognostic value for pCR compared to the more laborious MAW, allowing broad scale applications. The effect of denoising on the prognostic value for pCR needs to be investigated in larger cohorts.

1. Introduction

Esophageal cancer is the 8th most common cancer type in the world and it is associated with a poor five year overall survival rate of 35% [1,2]. Currently, neoadjuvant chemoradiotherapy (nCRT) followed by esophagectomy is the standard of curative care for locally advanced, resectable esophageal cancer without distant metastases [3,4].

Pathologic complete response (pCR) occurs in 16–30% of the patients, which means no viable tumor cells are observed in the resection specimen [5]. If these patients could be identified prior to surgery, a wait-and-see approach might be considered, excluding them from the considerable risks associated with an esophagectomy. Preoperative imaging techniques to diagnose pCR are yet unsatisfactory. For instance, FDG-PET scanning demonstrates a sensitivity and specificity of around 60–70%, which is deemed insufficient to alter treatment strategy [6]. Endoscopic biopsy is a specific but not sensitive method to assess pCR (sensitivity 35%, specificity 91%) [7].

Diffusion weighted MRI (DW-MRI) is an imaging technique that

* Corresponding author at: Heidelberglaan 100, 3584 CX Utrecht, The Netherlands.
E-mail address: G.j.meijer@umcutrecht.nl (G. Meijer).
[1] Both authors contributed equally.

https://doi.org/10.1016/j.phro.2023.100489
Received 21 April 2023; Received in revised form 24 August 2023; Accepted 25 August 2023
Available online 30 August 2023

measures variation of diffusion of water molecules in different tissues. Tumor tissue is generally denser than healthy tissue due to the high density of tumor cells, which results in a more restricted diffusion of water molecules [8]. The apparent diffusion coefficient (ADC) is utilized as a quantitative measure for assessing the extent of water diffusion motion within tissue. Administration of chemoradiotherapy can result in the loss of cell membrane integrity and apoptosis and thereby increase the diffusion of water molecules, which can be detected as an increase in the mean tumor ADC [9]. Various previous studies demonstrated that ADC changes (e.g. between prior and during treatment) are prognostic for pCR in patients with esophageal cancer with area under the receiver operating characteristic curve (AUC) ranging from 0.71 to 1.00 [10–14].

Delineation of the region of interest is necessary to extract ADC values from the tumor. Manual tumor delineation on high b-value DW-MRI scans is labor intensive, time consuming, and is often subject to interpretation, especially at later intervals where signal intensities might have dropped. This in turn affects delineation accuracy and ADC readings. Semi-automatic delineation could potentially mitigate these issues, which facilitates the use of ADC as a biomarker in clinical decision-making.

Furthermore, fluctuations from multiple sources (e.g. breathing, B0 inhomogeneity, gradient artefacts) can create significant additive noise in DW-MRI scans especially in the thoracic region, which complicates tumor delineations and are also propagated in the resulting ADC values [15]. Denoising algorithms have been shown to reduce the additive noise in DW-MRI scans, e.g. for brain imaging [16]. The effect of denoising on ADC values, obtained using a semi-automatic workflow of DW-MRI scans of patients with esophageal cancer is however, yet unknown.

The aim of this study was to present the application of a denoising algorithm and our own developed semi-automated delineation workflow (SADW) for delineating and analyzing the DW-MRI signal of esophageal tumors during nCRT. The resulting ADC measurements and subsequent prognostic values are benchmarked against the results from a manually adjusted workflow (MAW) of a previous study [10].

2. Material and methods

2.1. Study population

This study was approved by the institutional review board of the University Medical Center Utrecht (protocol ID 15–340). All participants gave informed consent. Included patients corresponded with a previous publication of our department using manually adjusted tumor delineations [10]. Patients underwent one pre-nCRT baseline DW-MRI scan around 5 days before the start of nCRT and five weekly DW-MRI scans during the course of nCRT.

Patients diagnosed with squamous cell carcinoma or adenocarcinoma of the esophagus or gastroesophageal junction, who were scheduled for nCRT followed by esophagectomy from December 2015 to April 2018, were eligible for inclusion. Exclusion criteria for the study were patients under 18 years of age, previous thoracic surgery or radiotherapy, and contraindications for MRI. The current analysis excluded patients with unexpected distant metastatic disease after nCRT (n = 2), poor tumor visibility on DW-MRI (n = 3), initial small tumor volume < 7 mL (n = 3), tumor histology other than adenoacarcinoma or squamous cell carcinoma based on the resection specimen (n = 1), or those who withdrew from the study (n = 1).

2.2. Image acquisition

All DW-MRI scans were acquired on a 1.5-T Philips Ingenia (Best, The Netherlands). This included respiratory-triggered transversal T2-weighted scans (tT2W) and DW-MRI scans. All patients were scanned with 13b-values between 0 and 800 s/mm^2 (0, 10, 20, 30, 40, 50, 75, 100, 200, 350, 500, 650, 800 s/mm^2). Image acquisition details are provided in Table S1.

2.3. Denoising of DW-MRI images

A self-supervised denoising method (Patch2Self) was used to improve signal to noise ratio without having to calibrate a noise model [15]. Patch2Self uses the entire volume to learn a full-rank locally linear denoiser for that volume. By taking advantage of the oversampled DW-MRI data (13b-values), this algorithm can separate structure from noise without requiring an explicit model for either. For each voxel we employed a 3D-patch of all neighbouring voxels and itself (27 voxels) to train the denoiser.

2.4. Manually adjusted delineation workflow

In the MAW, the primary tumor, excluding the lumen, was identified using baseline DW-MRI scans with a b-value of 800 s/mm^2. This was done using ITK-SNAP software, with a semi-automated delineation method. Two readers reviewed extensively and made manual edits to reach a consensus. The initial contouring was then rigidly propagated to subsequent weeks, with further adjustments made by one reader based on signal reduction and tumor regression on b = 800 s/mm^2 DW-MRI and tT2W scans, respectively. The contouring was performed cautiously to avoid unreliable boundaries due to motion or image distortions. The tT2W images were used to verify the apparent tumor bed. The readers were unaware of patient characteristics and clinical outcomes regarding pathological response.

2.5. Semi-automatic delineation and propagation tool

Semi-automatic delineation of the baseline DW-MRI scans was performed in MIM version 7.0.1. (MIM Software Inc., Cleveland, Ohio, USA) by a clinical PhD student (RdB). Here, the PET Edge® functionality, an active contour algorithm based on intensity and gradient information was used to semi-automatically delineate the tumor on the b = 500 s/mm^2 denoised images. The rationale behind the use of b = 500 s/mm^2 scans for delineation was to achieve better signal to noise ratio and improved performance of PET Edge®. No manual editing was performed afterwards. This delineation was propagated to the DW-MRI scans in subsequent weeks by using deformable registration techniques of VoxAlign Deformation Engine® in MIM version 7.0.1. Here, the deformation vector fields extracted from the b0-b0 registrations were used, as these images contained most anatomical information. To avoid the edges of the tumor boundaries and to remove possible inclusion of surrounding structures, a 3-mm isotropic shrinkage margin of the delineation was applied (Fig. 1).

ADC maps were generated based on a mono-exponential model fitted on b-values of 0, 200 and 800 s/mm^2, as commonly used in literature [101217]. For each ADC map the average ADC value was calculated within both the semi-automatic segmentation and the manually adjusted segmentation [10]. To separate the impact of the denoising, ADC values were calculated both for the raw and the denoised data.

2.6. Statistical analysis

Differences between median tumor volumes of the SADW and MAW delineations were assessed using the Wilcoxon signed-rank test. Moreover, the overlap of the SADW and MAW delineations was quantified using Dice coefficients. The Hausdorff distance was used to indicate the largest distance between the annotations of the two workflows. Mean tumor ADC values were extracted from the DW-MRI volumes of interest. The relative changes in mean ADC values between the baseline scans and the scans during nCRT were calculated and included in the analyses, as based on previous literature: $\Delta ADC(\%) = [\text{mean } ADC_{week(n)} - \text{mean } ADC_{baseline}] / \text{mean } ADC_{baseline} * 100\%$ [12]. For each patient the ADC_{mean} increase (%/week) was assessed by applying a linear regression

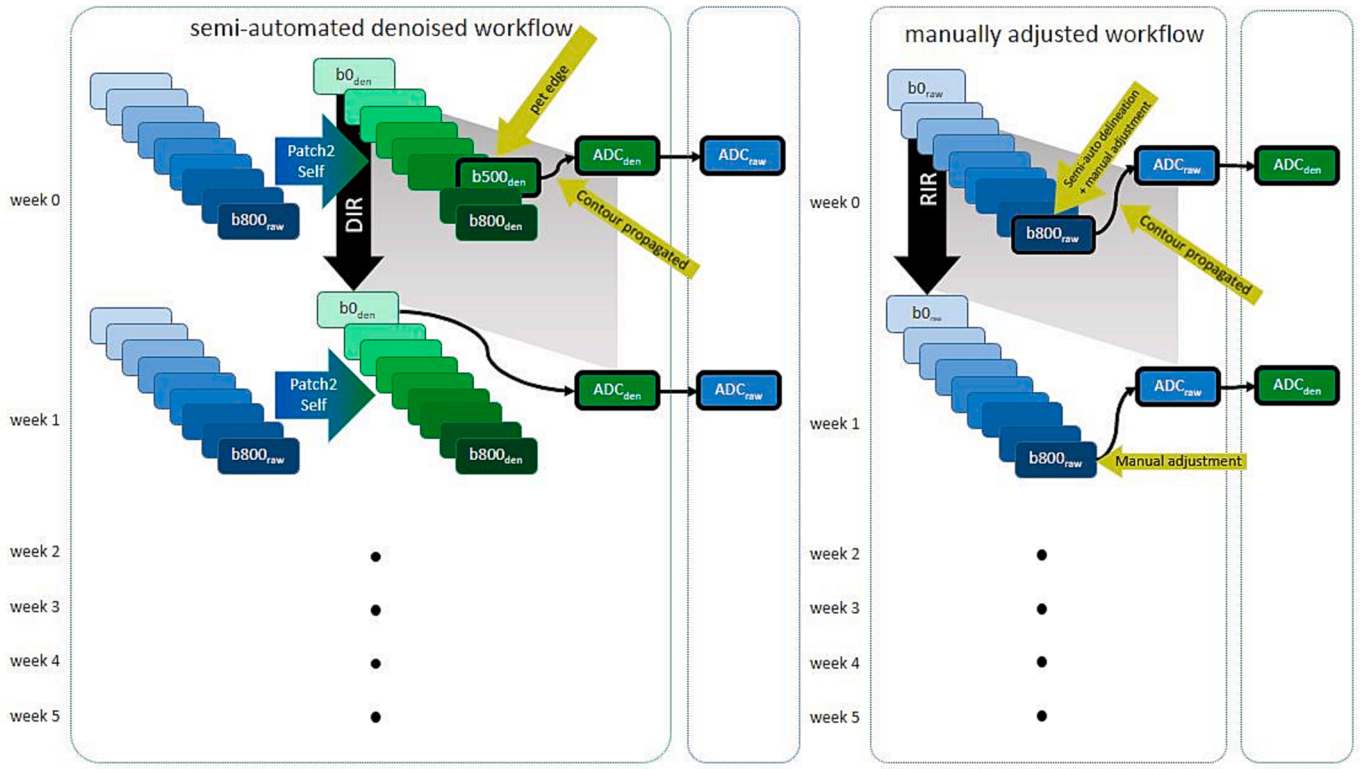

Fig. 1. Schematic overview of the denoised (den) semi-automatic delineation workflow (SADW, left) with b0-b0 deformable image registration (DIR) and of the manually adjusted workflow (right) to obtained apparent diffusion coefficient (ADC) values. The narrow columns are used to analyse the effect of denoising.

fit through all ΔADC_{mean} data points for the SADW and the MAW method. Bland Altman analysis and the Pearson correlation coefficient were used to compare ADC values of the SADW and MAW.

The ability to discriminate between pCR and non-pCR was quantified using logistic regression for both delineation methods based on the variable of ADC_{mean} increase (%/week). To study the impact of

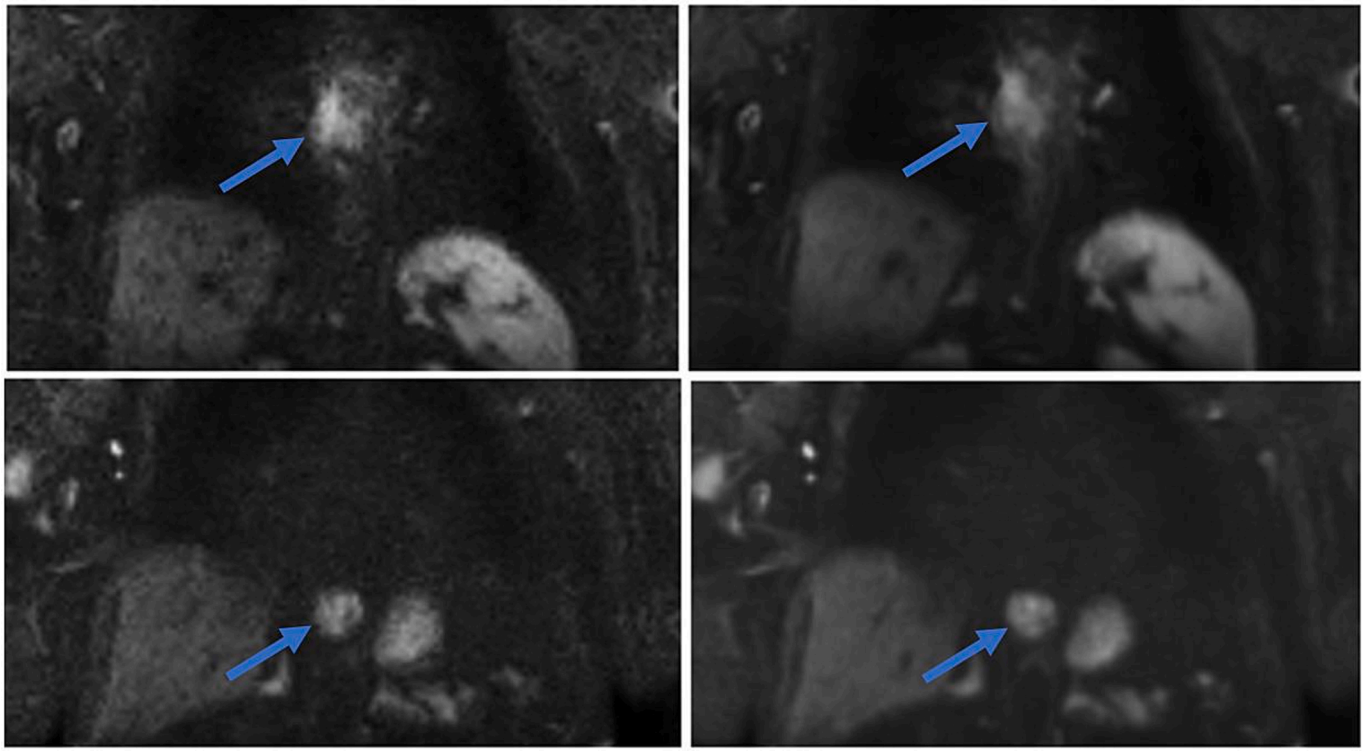

Fig. 2. DW-MRI (b = 500 s/mm^2) images of two of the included patients without denoising (left) and with denoising (right) in the coronal plane. The tumour is indicated with an arrow.

denoising, the c-statistic of the logistic regression models for prediction of pCR of ADC_{mean} increase (%/week) during nCRT was calculated for both the raw and the denoised data. To calculate the 95% confidence interval (CI) for the difference in the c-statistics, the DeLong method was used. The c-statistics of the prediction models were compared using de one-sided Delong test for non-inferiority. Statistical analyses were performed using SPSS version 25.0 (IBM, Armonk, New York, USA) and R software for statistical computing version 1.4.1106 ('glmnet', 'pROC', and 'Hmisc' packages, https://www.R-project.org). P values < 0.05 were deemed statistically significant. Data visualisation was performed using GraphPad Prism version 6.04 (GraphPad Software, La Jolla, California, USA).

3. Results

3.1. Patients

Patient characteristics are presented in the supplementary material Table S2. A total of 119 DW-MRI scans were analysed. Baseline DW-MRI scans were available in all of the included patients. However, one DW-MRI scan during nCRT at week 4 was missing due to the patient's refusal.

3.2. Denoising of the DW-MRI images

Denoising was applied to all DW-MRI scans and visual results before and after denoising are displayed in Fig. 2. Denoising resulted in visually superior images, compared with the raw DW-MRI scans. Additionally, ADC maps before and after denoising of two included patients are displayed in Figure S1.

3.3. Comparison of delineations of the SADW and MAW

Median volumes of the delineations were 11.4 mL and 10.8 mL for the SADW and MAW, respectively (Wilcoxon signed-rank test, $p < 0.001$). Median Dice coefficient between the SADW and MAW delineations was 0.64 (interquartile range (IQR) 0.20, Table S3). The median Hausdorff distance between the two workflows was 13.37 mm (IQR 12.98). An example of SADW and MAW delineations on the b = 500 s/mm^2 DW-MRI scans is displayed in Fig. 3, which shows comparable tumor delineations.

3.4. ADC changes during nCRT using the SADW and the MAW

Pearson regression analysis showed high accordance between ADC values obtained from the SADW and MAW in all included patients (r =

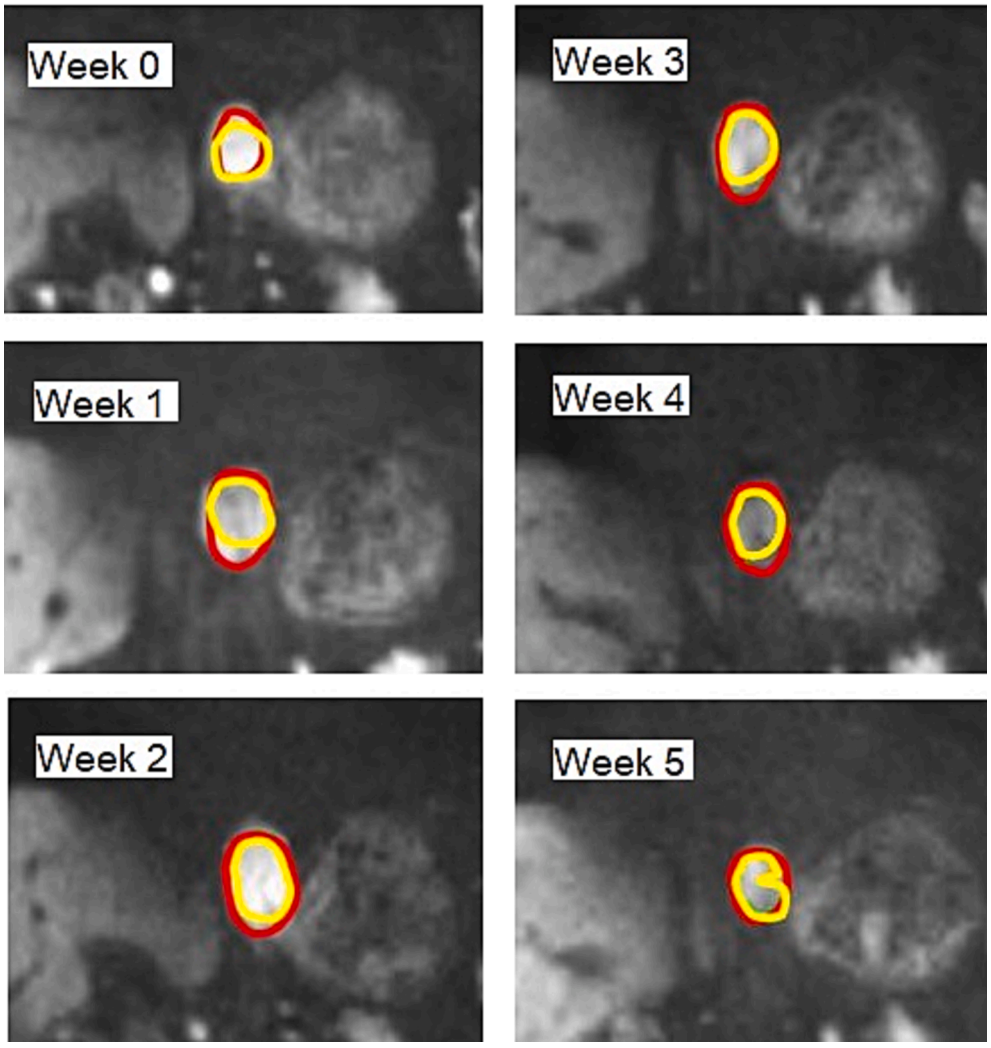

Fig. 3. Example of a semi-automatic delineation workflow (red) and manually adjusted delineation (yellow), on the b = 500 s/mm^2 DW-MRI scans in coronal plane of week 0, 1, 2, 3, 4, and 5 for a single patient. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

0.90, Fig. 4). Bland Altman analysis demonstrated that absolute mean ADC values were systematically higher for the SADW, compared with the MAW (mean difference in mean ADC 220 * 10^{-6} mm^2/s, $p < 0.001$). The contours of the two most evident outliers, labled as number 1 and 2 in Fig. 4, are presented in the supplementary material Figure S2.

Relative changes in ADC values for the SADW and MAW are displayed in Fig. 5, including a linear regression line fitted through the mean ΔADC values for pCR and non-pCR patients. SADW and MAW showed similar trends for identifying pCR patients. A larger increase in mean ΔADC(%) per week during nCRT for patients having pCR was observed for SADW compared with MAW (16.6% vs. 12.4% per week).

3.5. *Prognostic value for pCR of ADCmean increase using SADW and the effect of denoising*

For the MAW with raw data, the c-statistic for predicting pCR with ADC_{mean} increase (%/week) was 0.80 (95% confidence interval (CI): 0.56 – 1.00). After denoising, the MAW c-statistic improved to 0.83 (95% CI: 0.61 – 1.00). Similarly, the SADW with raw data had a c-statistic of 0.77 (95% CI: 0.49 – 1.00), which increased to 0.84 (95% CI 0.62 – 1.00) after denoising. The effect of denoising for the SADW and MAW is displayed in Fig. 6. There were no statistically significant differences observed between the two delineation workflows or after applying the denoising procedure (Table S4).

4. Discussion

DW-MRI appears to have prognostic value for identifying patients with a pCR after nCRT for esophageal cancer. However, significant additive noise in DW-MRI scans may complicate tumor delineation in DW-MRI workflows and this will affect the corresponding ADC signal. Moreover, manual tumor delineation on high b-value DW-MRI scans is labor intensive and often subject to interpretations. This study presents the use of a denoising algorithm and a SADW for the assessment of ADC and its prognostic value for pCR, using weekly DW-MRI scans during nCRT in patients with esophageal cancer. The use of a SADW can increase general applicability in larger studies and in clinical settings [13]. The SADW presents a more standardized way to determine tumor segmentations for ADC analysis compared with manual delineation. Semi-automatic delineations are less likely to be dependent on interobserver variations [18,19].

Semi-automated and manually adjusted tumor delineations showed a moderate overlap with a median Dice coefficient of 0.64. It is important to note that comparing the two delineation approaches is challenging due to the lack of a reference standard. Additionally, Dice coefficients tend to be moderate or low for small structures, such as esophageal carcinomas. The volume of the SADW was found to be significantly larger than the MAW (median volume 11.4 mL vs. 10.8 mL). Possibly, the manual corrections made for tumor response was the reason for the smaller volumes. It should be noted that these differences in volume may not have a substantial impact, as relative changes in mean ADC values are used to identify patients with pCR.

Bland Altman analysis showed that the SADW exhibited higher absolute mean ADC values compared to the MAW, potentially due to the larger delineations, containing tissue with higher apparent diffusion. Variation in mean ADC values between the MAW and SADW were more pronounced during the final week of nCRT, when tumors may be less visible on DW-MRI. Ultimately, relative ADC changes per week during nCRT using the SADW demonstrated a prognostic value for pCR comparable to the MAW (c-statistic of 0.84 vs. 0.80, $p = 0.69$). Denoising showed improved tumor visualisation compared with the original DW-MRI images. No significant improvement in prognostic value of pCR was observed for denoising in the SADW and MAW groups. While improved tumor visualization was observed, the impact of denoising on prognostic accuracy for pCR is not strongly substantiated here using the mean ADC values. Nonetheless, the potential augmentation from denoising in assessing treatment response remains promising and should be studied using larger cohorts.

As the sample size in this study was small, the difference in the prognostic value between the SADW and MAW is best illustrated by comparing the fitted ADC increases per week for the pCR and the non-pCR patients (Fig. 6). Smaller datasets may exhibit larger variations in the c-statistic, making it challenging to draw definitive conclusions. Minor changes in c-statistics were observed comparing the MAW and SADW, including a slight decrease in c-statistic from 0.80 to 0.77 in the raw data using the SADW. The uncertainty is highlighted by the wide confidence intervals and superiority of any workflow cannot be claimed. However, the SADW and MAW showed comparable results concerning the prognostication of pathologic complete response in this study cohort.

In this study, we reported the use of a semi-automatic delineation method and the effect of using denoised DW-MRI scans for the assessment of ADC changes and its prognostic value for response to nCRT for esophageal cancer. Studies evaluating (semi-)automatic delineations of DW-MRI scans are published for other cancer subtypes, for example rectal malignancies [18]. One study in rectal cancer patients found that

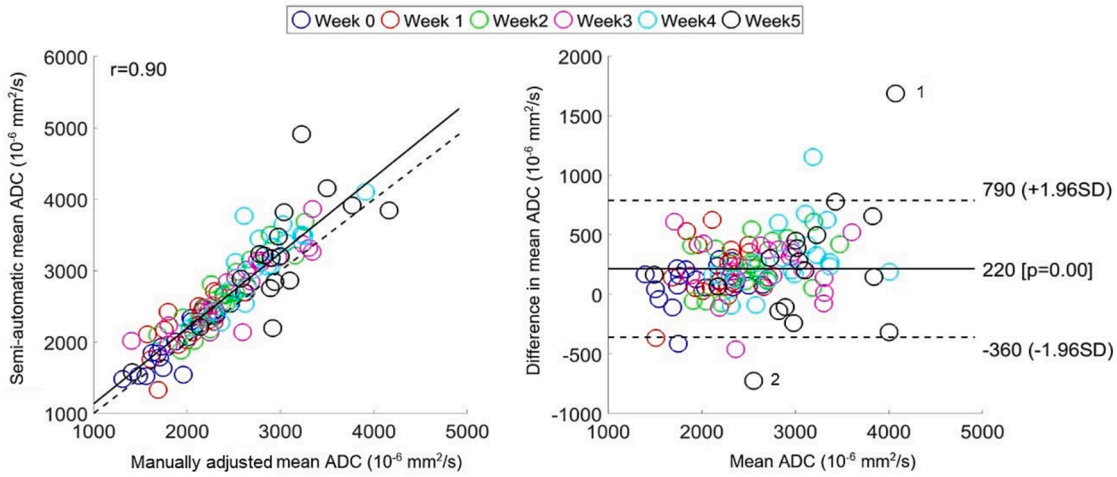

Fig. 4. Pearson regression analysis (left, r = 0.90) and Bland Altman (right) plot comparing apparent diffusion coefficient (ADC) values with the semi-automatic delineation workflow and manually adjusted delineation for the included cohort of 20 patients. Mean ADC values were significantly higher for the SADW ($p < 0.001$). In the Pearson regression analysis the dashed line indicates perfect correspondence, the solid line is the linear regression line of the observed data points. In the Bland Altman plot, the dashed lines indicate 1.96 standard deviations of the mean difference in mean ADC between the two workflows (solid line).

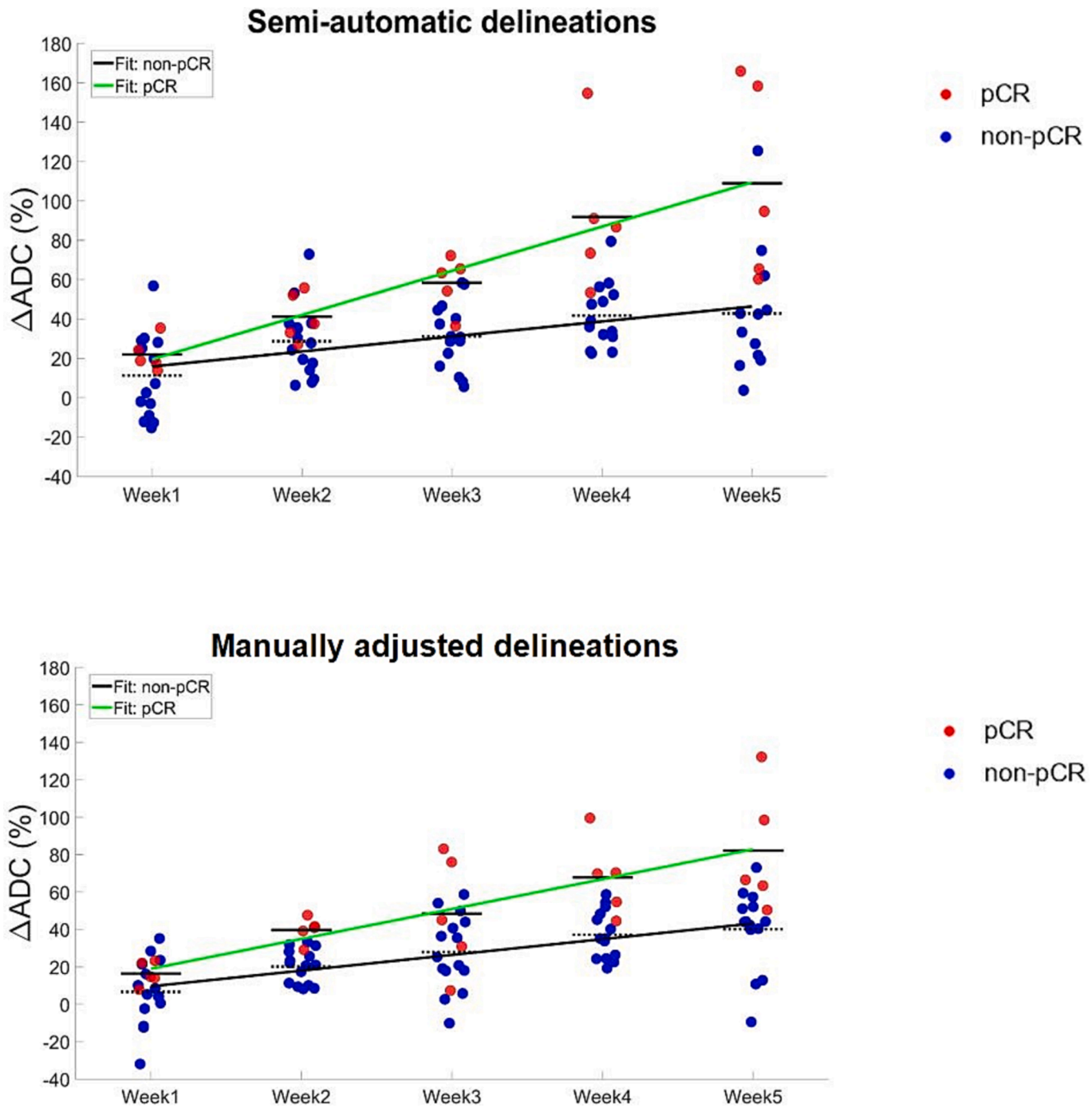

Fig. 5. Relative changes in apparent diffusion coefficient (ADC) values between baseline scans and weekly scans during neoadjuvant chemoradiotherapy between pathologic complete responders and non-pathologic complete responders (pCR in red, non-pCR in blue) using semi-automatic workflow (above) and manually adjusted delineation (below). The group mean per week for pCR and non-pCR is indicated as a solid and dashed horizontal line, respectively. A linear regression line is fitted through the points for both pCR (green) and non-pCR (black) separately. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

DW-MRI volumetry using a semi-automated delineation approach is promising and a potentially time-saving alternative to manual tumor delineation, particularly for assessing primary tumor volumetry [20]. Another study concluded that deep learning can perform an accurate automated localisation and delineation of rectal cancer on MRI [21]. Semi-automatic contouring based on gradient is shown to have consistent results for target volume contouring in thorax regions [22]. Denoised images have also been found to improve delineation performance of deep learning algorithms in breast cancer radiotherapy planning [23]. To our best knowledge, this is the first study that indicates that semi-automatic segmentations of DW-MRI can result in non-inferior prediction of treatment response.

Strengths of this study include consistent usage of the same nCRT regimens in all included patients. Only one DW-MRI scan during treatment was missing, which resulted in a nearly complete study cohort with a large number of DW-MRI scans. An important limitation is the relatively small number of patients included in the study, especially with regard to analysis of pCR, more reliable conclusions can be expected in

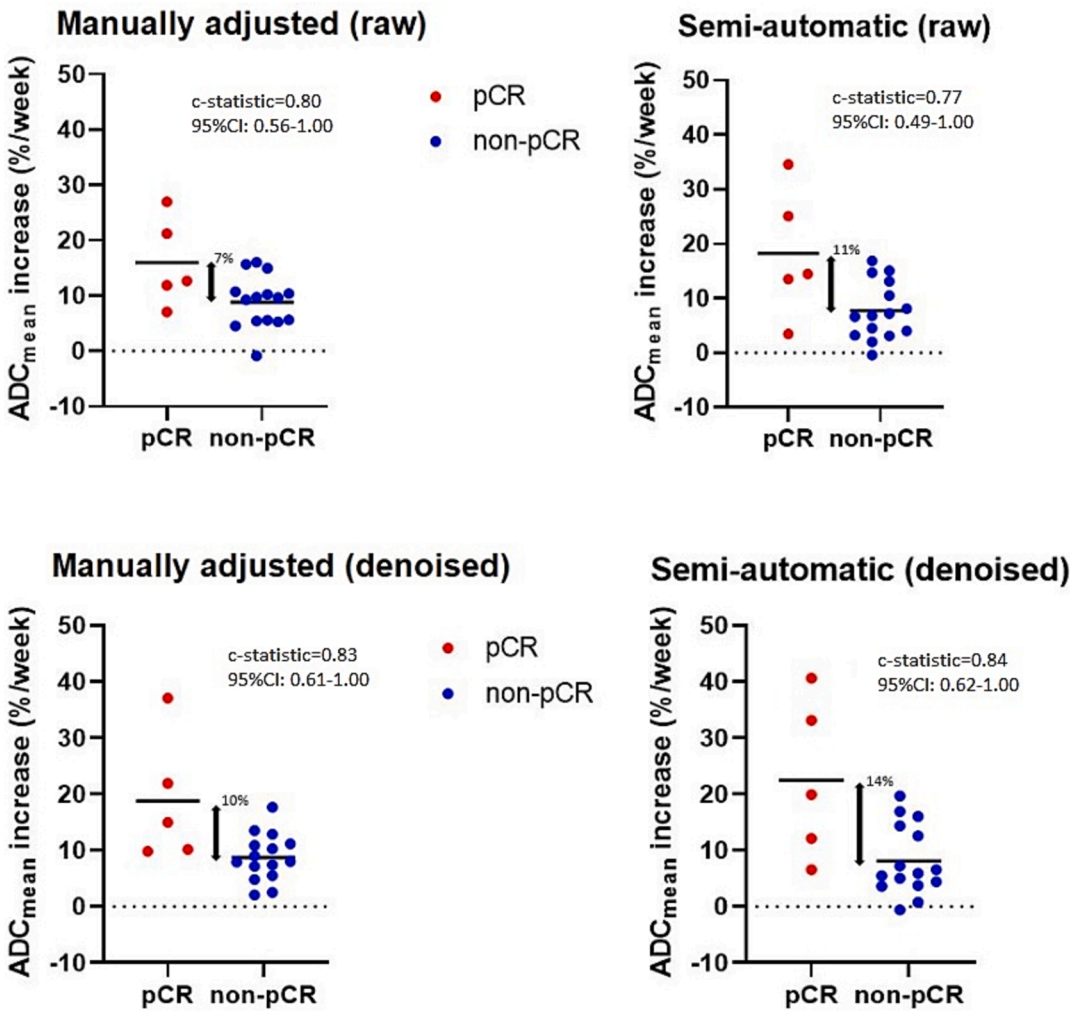

Fig. 6. ADC_{mean} increase (%/week) during neoadjuvant chemoradiotherapy between pathologic complete responders (pCR, red dots) and non-pCR patients (blue dots). Results are presented for the semi-automatic delineation workflow (SADW) and manually adjusted workflow (MAW) with and without denoising of the DW-MRI scans. The black lines indicate the mean of the pCR and non-pCR group. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

larger patient cohorts. At last, uncertainty may remain in this workflow due to the semi-automatic contouring and the deformable registration strategy, introducing potential sources of variability.

Future research should aim at larger patient cohorts to validate the effect of denoising and the SADW. The role of intravoxel incoherent motion (IVIM) MRI in evaluating treatment response must be investigated as well. Also, the prognostic value for DW-MRI in combination with other imaging modality parameters must be assessed, preferably in a comparable semi-automatic manner. The choice of delineation strategy, including the selection of the appropriate b value for tumor annotation, can significantly impact the calculated apparent diffusion coefficient (ADC) values. Therefore, it is essential to report the standards for delineation on DWI to ensure more consistent comparisons in future studies [24].

Results of the PRIDE study, a multicentre observational cohort study in the Netherlands, will provide further insight of the prognostic value of DW-MRI for pCR in a larger cohort [13]. The aim of the PRIDE study is to develop a prediction model for pCR to nCRT in esophageal cancer, based on DW-MRI, dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI), and positron emission tomography with computed tomography (PET-CT). Also, the PRIDE study will give an opportunity to investigate the current SADW in a larger cohort.

In conclusion, in this study a simple SADW and a denoising algorithm of DW-MRI scans to assess relative ADC values was applied. The SADW showed non-inferior prognostic accuracy for pCR compared with manually adjusted delineations, is less labour intensive and possibly less operator dependent. Denoising can improve tumor visualisation, but its effect on ADC values to predict pCR needs to be further explored. This allows for widespread application of the SADW in larger patient cohorts and in centers with limited experience with DW-MRI.

Funding information

No financial support was received nor requested for the research, authorship, and/or publication of this article.

CRediT authorship contribution statement

Robin den Boer: Conceptualization, Formal analysis, Methodology, Writing – original draft. **Kelvin Ng Wei Siang:** Data curation, Writing – original draft, Software, Methodology. **Mandy Yuen:** Software, Formal analysis. **Alicia Borggreve:** Data curation, Writing – review & editing. **Ingmar Defize:** Conceptualization, Supervision. **Astrid van Lier:** Conceptualization, Methodology, Writing – review & editing. **Jelle Ruurda:** Conceptualization, Supervision, Writing – review & editing. **Richard van Hillegersberg:** Conceptualization, Supervision, Writing – review & editing. **Stella Mook:** Conceptualization, Supervision, Writing – review & editing. **Gert Meijer:** Conceptualization, Methodology,

Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.phro.2023.100489.

References

[1] Omloo JMT, Lagarde SM, Hulscher JBF, Reitsma JB, Fockens P, van Dekken H, et al. Extended transthoracic resection compared with limited transhiatal resection for adenocarcinoma of the mid/distal esophagus: five-year survival of a randomized clinical trial. Ann Surg 2007;246:992–1000. https://doi.org/10.1097/SLA.0b013e31815c4037.

[2] Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. CA Cancer J Clin 2015;65:87–108. https://doi.org/10.3322/caac.21262.

[3] Van Hagen P, Hulshof MCCM, Van Lanschot JJB, Steyerberg EW, Van Berge Henegouwen MI, Wijnhoven BPL, et al. Preoperative chemoradiotherapy for esophageal or junctional cancer. N Engl J Med 2012;366:2074–84. https://doi.org/10.1056/NEJMoa1112088.

[4] Shapiro J, van Lanschot JJB, Hulshof MCCM, Van Hagen P, Van Berge Henegouwen MI, Wijnhoven BPL, et al. Neoadjuvant chemoradiotherapy plus surgery versus surgery alone for oesophageal or junctional cancer (CROSS): long-term results of a randomised controlled trial. Lancet Oncol 2015;16:1090–8. https://doi.org/10.1016/S1470-2045(15)00040-6.

[5] Donahue JM, Nichols FC, Li Z, Schomas DA, Allen MS, Cassivi SD, et al. Complete pathologic response after neoadjuvant chemoradiotherapy for esophageal cancer is associated with enhanced survival. Ann Thorac Surg 2009;87:392–8. https://doi.org/10.1016/j.athoracsur.2008.11.001.

[6] De Gouw DJJM, Klarenbeek BR, Driessen M, Bouwense SAW, Van Workum F, Fütterer JJ, et al. Detecting pathological complete response in esophageal cancer after neoadjuvant therapy based on imaging techniques : a diagnostic systematic review and meta-analysis. J Thorac Oncol 2019;14:1156–71. https://doi.org/10.1016/j.jtho.2019.04.004.

[7] Van Rossum PSN, Goense L, Meziani J, Reitsma JB, Siersema PD, Vleggaar FP, et al. Endoscopic biopsy and EUS for the detection of pathologic complete response after neoadjuvant chemoradiotherapy in esophageal cancer: a systematic review and meta-analysis. Gastrointest Endosc 2016;83:866–79. https://doi.org/10.1016/j.gie.2015.11.026.

[8] Koh DM, Collins DJ. Diffusion-weighted MRI in the body: applications and challenges in oncology. Am J Roentgenol 2007;188:1622–35. https://doi.org/10.2214/AJR.06.1403.

[9] Thoeny HC, Ross BD. Predicting and monitoring cancer treatment response with diffusion-weighted MRI. J Magn Reson Imaging 2010;32:2–16. https://doi.org/10.1002/jmri.22167.

[10] Borggreve AS, Heethuis SE, Boekhoff MR, Goense L, van Rossum PSN, Brosens LAA, et al. Optimal timing for prediction of pathologic complete response to neoadjuvant chemoradiotherapy with diffusion-weighted MRI in patients with esophageal cancer. Eur Radiol 2020;30:1896–907. https://doi.org/10.1007/s00330-019-06513-0.

[11] Fang P, Musall BC, Son JB, Moreno AC, Hobbs BP, Carter BW, et al. Multimodal imaging of pathologic response to chemoradiation in esophageal cancer. Int J Radiat Oncol Biol Phys 2018;102:996–1001. https://doi.org/10.1016/j.ijrobp.2018.02.029.

[12] Van Rossum PS, Van Lier AL, Van Vulpen M, Reerink O, Lagendijk JJ, Lin SH, et al. Diffusion-weighted magnetic resonance imaging for the prediction of pathologic response to neoadjuvant chemoradiotherapy in esophageal cancer. Radiother Oncol 2015;115:163–70. https://doi.org/10.1016/j.radonc.2015.04.027.

[13] Borggreve AS, Mook S, Verheij M, Mul VEM, Bergman JJ, Bartels-Rutten A, et al. Preoperative image-guided identification of response to neoadjuvant chemoradiotherapy in esophageal cancer (PRIDE): a multicenter observational study. BMC Cancer 2018;18:1006. https://doi.org/10.1186/s12885-018-4892-6.

[14] Li QW, Qiu B, Wang B, Wang DL, Yin SH, Yang H, et al. Prediction of pathologic responders to neoadjuvant chemoradiotherapy by diffusion-weighted magnetic resonance imaging in locally advanced esophageal squamous cell carcinoma: a prospective study. Dis Esophagus 2018:31. https://doi.org/10.1093/dote/dox121.

[15] Fadnavis S, Batson J, Garyfallidis E. Patch2Self: Denoising diffusion MRI with self-supervised learning. Advances in Neural Information Processing Systems 2020. https://doi.org/10.48550/arXiv.2011.01355.

[16] Cheng H, Vinci-Booher S, Wang J, Caron B, Wen Q, Newman S, et al. Denoising diffusion weighted imaging data using convolutional neural networks. PLoS One 2022:17. https://doi.org/10.1371/journal.pone.0274396.

[17] Maffazzioli L, Zilio MB, Klamt AL, Duarte JA, Mazzini GS, Campos VJ, et al. ADC as a predictor of pathologic response to neoadjuvant therapy in esophageal cancer: a systematic review and meta-analysis. Eur Radiol 2020;30:3942-42. https://doi.org/10.1007/s00330-020-06723-x.

[18] Bisgaard ALH, Brink C, Fransen ML, Schytte T, Behrens CP, Vogelius I, et al. Robust extraction of biological information from diffusion-weighted magnetic resonance imaging during radiotherapy using semi-automatic delineation. Phys Imaging Radiat Oncol 2022;21:146–52. https://doi.org/10.1016/j.phro.2022.02.014.

[19] Pfaehler E, Mesotten L, Kramer G, Thomeer M, Vanhove K, de Jong J, et al. Repeatability of two semi-automatic artificial intelligence approaches for tumor segmentation in PET. EJNMMI Res 2021;11:4. https://doi.org/10.1186/s13550-020-00744-9.

[20] Van Heeswijk MM, Lambregts DMJ, Van Griethuysen JJM, Oei S, Rao SX, De Graaff CAM, et al. Automated and semiautomated segmentation of rectal tumor volumes on diffusion-weighted MRI: can it replace manual volumetry? Int J Radiat Oncol Biol Phys 2016;94:824–31. https://doi.org/10.1016/j.ijrobp.2015.12.017.

[21] Trebeschi S, Van Griethuysen JJM, Lambregts DMJ, Lahaye MJ, Parmar C, Bakers FCH, et al. Deep learning for fully-automated localization and segmentation of rectal cancer on multiparametric MR. Sci Rep 2017;7:5301. https://doi.org/10.1038/s41598-017-05728-9.

[22] Werner-Wasik M, Nelson AD, Choi W, Arai Y, Faulhaber PF, Kang P, et al. What is the best way to contour lung tumors on PET scans? multiobserver validation of a gradient-based method using a NSCLC digital PET phantom. Int J Radiat Oncol Biol Phys 2012;82:1164–71. https://doi.org/10.1016/j.ijrobp.2010.12.055.

[23] Im JH, Lee IJ, Sung CY, J, Ha JS, Lee H.. Impact of denoising on deep-learning-based automatic segmentation framework for breast cancer radiotherapy planning. Cancers 2022;14:3581. https://doi.org/10.3390/cancers14153581.

[24] Mahmood F, Johannesen HH, Geertsen P, Opheim GF, Hansen RH. The effect of region of interest strategies on apparent diffusion coefficient assessment in patients treated with palliative radiation therapy to brain metastases. Acta Oncol 2015;54:1529–34. https://doi.org/10.3109/0284186X.2015.1061211.