

Original Research Article

Validation of quantitative magnetic resonance imaging techniques in head and neck healthy structures involved in the salivary and swallowing function: Accuracy and repeatability

F. Guerreiro^a, P.J. van Houdt^b, R.J.M. Navest^b, N. Hoekstra^a, M. de Jong^a, B.J. Heijnen^e, S.E. Zijlema^b, B. Verbist^{c,d}, U.A. van der Heide^{a,b}, E. Astreimidou^{a,*}

^a Department of Radiation Oncology, Leiden University Medical Center, Leiden, the Netherlands

^b Department of Radiation Oncology, The Netherlands Cancer Institute, Amsterdam, the Netherlands

^c Department of Radiology, Leiden University Medical Center, Leiden, the Netherlands

^d HollandPTC, Delft, the Netherlands

^e Department of Otorhinolaryngology and Head and Neck Surgery, Leiden University Medical Center, Leiden, the Netherlands



ARTICLE INFO

Keywords:

Head and Neck cancer
Radiotherapy
Xerostomia
Salivary glands
Dysphagia
Swallowing muscles
Quantitative MRI
Repeatability
T2 mapping
Fat fraction
DWI
ADC

ABSTRACT

Background and Purpose: Radiation-induced damage to the organs at risk (OARs) in head-and-neck cancer (HNC) patient can result in long-term complications. Quantitative magnetic resonance imaging (qMRI) techniques such as diffusion-weighted imaging (DWI), DIXON for fat fraction (FF) estimation and T₂ mapping could potentially provide a spatial assessment of such damage. The goal of this study is to validate these qMRI techniques in terms of accuracy in phantoms and repeatability in-vivo across a broad selection of healthy OARs in the HN region.

Materials and Methods: Scanning was performed at a 3 T diagnostic MRI scanner, including the calculation of apparent diffusion coefficient (ADC) from DWI, FF and T₂ maps. Phantoms were scanned to estimate the qMRI techniques bias using Bland-Altman statistics. Twenty-six healthy subjects were scanned twice in a test-retest study to determine repeatability. Repeatability coefficients (RC) were calculated for the parotid, submandibular, sublingual and tubarial salivary glands, oral cavity, pharyngeal constrictor muscle and brainstem. Additionally, a linear mixed-effect model analysis was used to evaluate the effect of subject-specific characteristics on the qMRI values.

Results: Bias was $0.009 \times 10^{-3} \text{ mm}^2/\text{s}$ for ADC, -0.7% for FF and -7.9 ms for T₂. RCs ranged $0.11\text{--}0.25 \times 10^{-3} \text{ mm}^2/\text{s}$ for ADC, $1.2\text{--}6.3\%$ for FF and $2.5\text{--}6.3 \text{ ms}$ for T₂. A significant positive linear relationship between age and the FF and T₂ for some of the OARs was found.

Conclusion: These qMRI techniques are feasible, accurate and repeatable, which is promising for treatment response monitoring and/or differentiating between healthy and unhealthy tissues due to radiation-induced damage in HNC patients.

1. Introduction

Radiation-induced damage to the salivary glands and swallowing muscles of head and neck cancer (HNC) patients can lead to late toxicities such as xerostomia and dysphagia, which severely impair the patients' quality of life (QoL) [1,2,3]. Techniques such as salivary gland scintigraphy [4], salivary flow measurements [5] and video fluoroscopy [6] are employed to measure functional impairment. However, these techniques are invasive to some extent, do not correlate with patient-

reported outcomes [7] and most importantly lack spatial information on the radiation-induced damage location. In the era of advanced photon and proton therapy techniques, steering steep dose gradients while sparing important organs at risk (OARs) is feasible [8]. Therefore, an objective quantification of the OARs radiation-induced damage may contribute to improving radiotherapy planning strategies and ultimately refining the current late toxicity prediction models [9].

Magnetic resonance imaging (MRI) is a non-invasive technique with superior soft-tissue contrast that plays an important role in the

* Corresponding author.

E-mail addresses: p.v.houdt@nki.nl (P.J. van Houdt), B.J.Heijnen@lumc.nl (B.J. Heijnen), B.M.Verbist@lumc.nl (B. Verbist), u.vd.heide@nki.nl (U.A. van der Heide), e.astreimidou@lumc.nl (E. Astreimidou).

<https://doi.org/10.1016/j.phro.2024.100608>

Received 17 May 2024; Received in revised form 20 June 2024; Accepted 27 June 2024

Available online 3 July 2024

2405-6316/© 2024 Published by Elsevier B.V. on behalf of European Society of Radiotherapy & Oncology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

diagnosis, prognosis, radiation treatment planning [10] and more recently in the position verification with MRI-guided systems [11]. Furthermore, quantitative MRI (qMRI) has shown potential in predicting and monitoring tumor response during treatment and to some limited extent evaluating OAR toxicity [12,13]. Diffusion-weighted imaging (DWI), for tissue cellularity quantification [14–16], chemical shift-encoded sequences such as DIXON, for fat fraction (FF) measurement [17] and T_2 mapping, for tissue edema characterization [17], have been showing promising value in HNC. These studies reported an apparent diffusion coefficient (ADC) and T_2 values increase in the salivary glands post-radiotherapy suggesting that these techniques could potentially be employed for late toxicity quantification.

Before using information derived from qMRI in clinical practice, their quality must be validated in terms of accuracy, repeatability and reproducibility as defined by the Radiological Society of North America's Quantitative Imaging Biomarkers Alliance (QIBA) [18]. Estimating qMRI repeatability is essential to define the threshold between changes due to true biological events and measurement errors. Recent studies have validated the acquisition of DWI and/or T_2 mapping in different tumor sites like HN [19–21], prostate [22,23] and liver [24]. Due to motion (swallowing) and susceptibility (air/tissue interfaces) artifacts, the HN region represents a challenging location for the qMRI acquisition. Thus, investigating qMRI repeatability for this region is crucial. Repeatability depends not only on the MR field strength, gradient and coil hardware but also on the structure composition, size and anatomical region [20,21,23]. Therefore, different repeatability outcomes might be expected between diverse structures (tumor and/or different OARs). To date, qMRI repeatability in the HN region with the above mentioned sequences has only been evaluated for DWI acquisitions and only for the submandibular and parotid glands [20,21].

The lack of repeatability literature in HNC creates a barrier for the clinical translation of qMRI measurements. Thus, the main goal of this study was to evaluate the accuracy in phantoms and the repeatability in healthy subjects of the following qMRI techniques at a diagnostic 3 T MRI scanner: DWI with split acquisition of fast spin echo signal for diffusion imaging (SPLICE) [25], FF and T_2 mapping. The repeatability analysis was performed for a wide selection of healthy structures shown to be potentially associated with HNC late toxicity: sublingual [26], submandibular, parotid [26–28] and tubarial salivary glands [29], oral cavity [28], pharyngeal constrictor muscle [27,28] and the brainstem [30]. Additionally, qMRI reference values for the abovementioned techniques and healthy OARs were reported and the effect of subject-specific characteristics on the qMRI measurements was evaluated.

2. Material and Methods

2.1. Imaging details

MRI acquisitions were performed at a 3 T diagnostic scanner with a 16-channel HN coil (Ingenia, Philips Healthcare, Best, The Netherlands), including: T_1 -weighted DIXON for anatomical definition, DWI-SPLICE for ADC calculation, mDIXON Quant for FF estimation and T_2 mapping for measurement of the T_2 relaxation time. All qMRI maps were calculated on the scanner. DWI-SPLICE acquisition was performed using three b-values: 0, 200 and 800 s/mm^2 . ADC maps were calculated by fitting a linear function to the log of the signal decay at two b-values, 200 and 800. The DWI at 0-value was excluded in order to mitigate perfusion contamination [31]. For FF the mDIXON Quant method with a six-echo 3D gradient echo sequence was employed and maps were reconstructed using a 7-peak fat spectral model and correction for T_2^* decay. For T_2 mapping, an accelerated multi-echo spin echo sequence with 12 echoes was used [32]. To avoid bias from stimulated echoes, the first echo was discarded by skipping the acquisition during scanning [33]. T_2 maps were calculated by fitting a mono-exponential decay function. The same scanner, receive coil and sequence parameters were used in both phantoms and healthy subjects examinations (Table 1).

Table 1

qMRI sequence parameters: (1) DWI-SPLICE, (2) mDIXON Quant and (3) T_2 mapping. Abbreviations: turbo spin echo (TSE), fast field echo (FFE), field of view (FOV), repetition time (TR), echo time (TE), Compressed Sensing–Sensitivity Encoding (CS-SENSE) and anterior-posterior (AP).

Parameter	DWI-SPLICE	mDIXON Quant	T_2 mapping
Technique	multi-slice TSE	3D FFE	multi-slice TSE
Orientation	transverse	–	transverse
Shot mode	single-shot	–	multi-shot
FOV [mm^3]	280 x 222 x 140	400 x 350 x 180	250 x 205 x 180
Acquired voxel size [mm^3]	3 x 3 x 4	2.5 x 2.5 x 3	3 x 3 x 4
Number of echoes	1	6	12
TR / TE / Δ TE [ms]	4322 / 67 / –	5.6 / 0.97 / 0.7	7650 / 20 / 10
b-values (averages)	0(2), 200(3), 800(6)	–	–
Flip angle [degree]	90	3	90
Acceleration factor (direction)	SENSE, 1.8 (AP)	SENSE, 2 (AP)	CS-SENSE, 4
Acquisition time [min: sec]	05:01	00:24	05:52

2.2. Phantom

Measurements with different phantoms were performed to evaluate the qMRI techniques accuracy. All phantoms were positioned such that samples were aligned with the B_0 . For ADC, the QIBA diffusion phantom (High Precision Devices, Inc., Boulder, Colorado) consisting of thirteen vials with aqueous solutions of 0–50 % w/w polyvinylpyrrolidone (true ADC values between 0.1–1.1 x 10^{-3} mm^2/s) was used. The vials were surrounded by ice water to ensure measurements at 273.15 K per phantom instructions. The DWI phantom was positioned with the center vial at the scanner isocenter to minimize the effect of gradient nonlinearities [18]. For FF, a commercially available proton-density FF phantom (Calimetrix, Model 300) was used with twelve agar gel-based vials with true FF values between 0–100 %. The phantom was placed in the scanner room one hour before examination to adjust to the room temperature. For T_2 mapping, the Eurospin TO5 phantom (Diagnostic Sonar, Livingston, Scotland) was used. T_2 analysis was performed in twelve gel samples with relaxation times between 49–156 ms at 3 T and 296 K. The temperature was measured before and after acquisition in a water tube kept near the phantom. Temperature corrections on the measured T_2 values were performed using the phantom's manufacturer tables. For data analysis, circular regions of interest (ROIs) with 20 mm diameter were delineated manually in the center slice of each vial on the qMRI maps using RayStation treatment planning system (version 10B, RaySearch Laboratories, Stockholm, Sweden).

2.3. In-vivo

Twenty-six healthy subjects were scanned twice with a minimum of 3 days apart in a test–retest fashion to assess long-term repeatability. The protocol was approved by the institutional review board and written informed consent was obtained from all participants before the first session. Subjects were instructed not to drink/eat 1 h before scanning so that the salivary glands were at rest during acquisition. All scans were performed between 17:00–19:00 h to account for the circadian rhythm. A head support placed inside the HN coil was used to stabilize the subject's head [34].

The following ROIs were defined by a board-certified HN radiation oncologist and a radiologist using RayStation on the anatomical T_1 -weighted DIXON test and retest images: all pairs of the salivary glands (parotid, submandibular, sublingual, tubarial), extended oral cavity (OC), the total, superior, middle and inferior pharyngeal constrictor

muscle (PCM, PCM_S, PCM_M, PCM_I) and the inferior section of the brainstem (Fig. 1). In addition, the retromandibular vein was manually removed from the parotid gland contours and air and dental artifacts from the OC contour. To minimize motion uncertainties, a rigid registration between the anatomical image and the DWI (b = 0) image, the FF map and the first acquired echo image (TE = 20 ms) was performed. The DWI (b = 0) and (TE = 20 ms) echo images were defined in the same frame of reference as the ADC and T₂ maps, respectively. For both test and retest sessions, the ROIs were rigidly propagated to each map and reviewed slice-by-slice for any misalignments. Manual corrections on the propagated ROIs were performed when needed. To assess the impact of delineation variability and ROI propagation uncertainties the same analysis was performed on both delineated and contracted ROIs. Large ROIs (OC, parotid and submandibular glands) were contracted by 1 mm and small ROIs (sublingual and tubarial glands and all PCM volumes) by 0.5 mm.

2.4. Statistical analysis

For the phantom data, the agreement between measured and reference qMRI values was assessed by a Bland-Altman analysis. Bias and limits of agreement (LoA) were computed using the difference between measured and reference qMRI values (Δ qMRI) and the reference value among all phantom vials.

For the subjects data, mean qMRI values were extracted for all ROIs and test–retest sessions. Mean qMRI values \pm 1 standard deviation (SD) among all subjects were reported for the test session. The repeatability coefficient (RC) was calculated by estimating the within-subject SD (wSD) (Eq.1). The RC is the value under which the difference between any two repeat measurements on the same subject acquired under identical conditions should fall with 95 % probability. In addition, the within-subject coefficient of variation (wCV%) was also calculated (Eq.2) [18].

$$RC = 1.96 \times \sqrt{2} \times wSD = 2.77 \times \sqrt{\frac{1}{n} \sum_{i=1}^n \sigma_i^2} \quad (1)$$

$$wCV (\%) = \sqrt{\frac{1}{n} \sum_{i=1}^n \frac{\sigma_i^2}{\mu_i^2}} \times 100 \quad (2)$$

with $i = 1, 2, 3 \dots n$ where n is the total number of ROIs, σ^2 the sample variance between test–retest values, μ the mean of test–retest values.

To estimate the effect of the MRI session and subject-characteristics on the qMRI values, a linear mixed-effect model analysis was carried-out using SPSS (version 29, IBM SPSS Inc, Chicago, Illinois). The dependent variable in the model was the qMRI value of the specified ROI, subject id was included as random effect. Age, body mass index (BMI), gender (female/male) and MRI session (test/retest) were included as fixed effects (significance < 0.05). Age and BMI were defined as continuous variables while sex and session as categorical variables (with male and test session as default).

3. Results

3.1. Phantom

The Δ qMRI ranged from -0.032 to $0.028 \times 10^{-3} \text{ mm}^2/\text{s}$ for ADC, from -1.6 to 0.1 % for FF and from -15.3 to -1.8 ms for T₂. The bias \pm LoA, as determined by the Bland-Altman analysis,

were $0.009 \pm 0.042 \times 10^{-3} \text{ mm}^2/\text{s}$ for ADC, -0.7 ± 0.9 % for FF and -7.9 ± -9.3 ms for T₂ (Fig. 2).

3.2. In-vivo

Twenty-six subjects, 12 males and 14 females, with median age of 34 years (range 23–61 years), median BMI 23.2 (range 18.2–31.1) were recruited. Re-test session was on average 17 ± 12 days apart from test session. An overview of the delineated ROIs volumes and mean qMRI values for the test session and repeatability analysis is given in Table 2. Fig. 3 shows the Δ qMRI values per ROI.

The range difference of the RC between the delineated ROIs and the contracted ROIs was from -0.09 to $0.00 \times 10^{-3} \text{ mm}^2/\text{s}$ for ADC, from -1.3 to 0.8 % for FF and from -1.7 to 0.7 ms for T₂. Thus for all qMRI techniques and the majority of the ROIs, the contracted ROIs had the

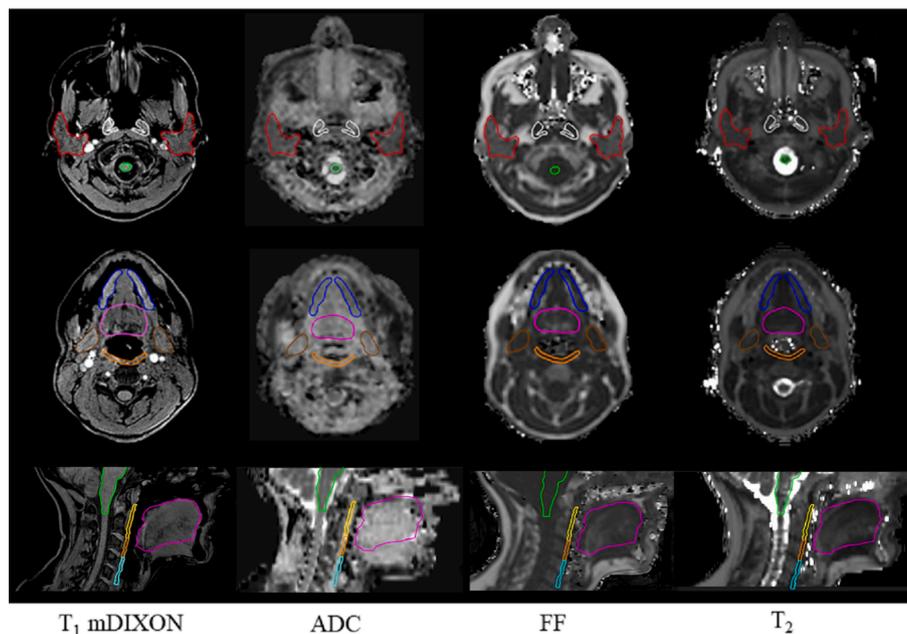


Fig. 1. Representative images of the study protocol sequences and the delineations for one healthy subject. The parotid glands are shown in red, the submandibular glands in brown, the sublingual glands in dark blue, the tubarial glands in white, the OC in pink, the PCM_S in yellow, the PCM_M in orange, the PCM_I in light blue and the brainstem in green. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

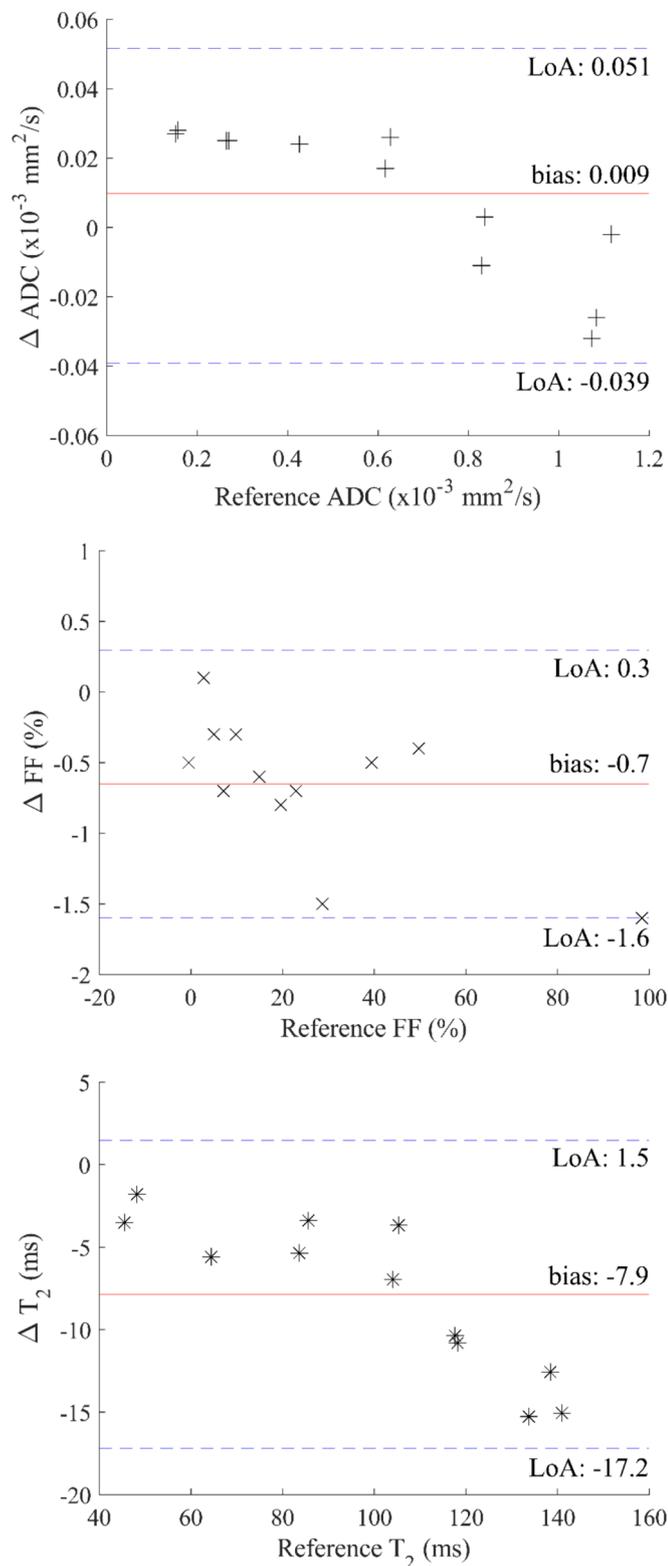


Fig. 2. Bland-Altman analysis for the three qMRI techniques with phantoms. The solid red line indicate the bias whereas the dashed blue lines represent the upper and lower limits of agreement (LoA). Each individual point represents a vial. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

same or slightly higher RC.

The linear mixed-effect model analysis results are given in Table 3. A significant positive linear relationship between age and the FF and T₂ values was found for the majority of the ROIs, suggesting that older subjects will denote higher qMRI values. A significant effect on the ADC values for age, BMI and sex was seen for a minority of the structures, however, the effect size was negligible. No significant differences were found between test–retest measurements, confirming a good repeatability for all qMRI techniques.

4. Discussion

With the potential of improving treatment response monitoring and the assessment of radiation-induced toxicity, the interest for qMRI in radiation oncology has been rapidly growing. We assessed the accuracy in phantoms and the repeatability in healthy subjects of ADC, FF and T₂ measurements on a 3 T diagnostic scanner. All qMRI techniques demonstrated good accuracy. To date, this is the first study evaluating the in-vivo repeatability of multiple qMRI techniques across a broad selection of OARs in the HN region. Calculated RCs were lower than $0.25 \times 10^{-3} \text{ mm}^2/\text{s}$ for ADC, 6.2 % for FF and 6.3 ms for T₂.

The qMRI techniques were evaluated in terms of accuracy with phantoms. The bias from the Bland-Altman analysis was close to zero for ADC and FF, thus considered negligible. These results align with prior reports on ADC [35] and FF [36] accuracy at 3 T diagnostic scanners. For T₂, a small negative bias was found over the entire range of T₂ values, which was also previously reported in studies using a multi-echo approach [22,23,35]. Nevertheless, as this bias was seen consistently for multiple vials with the same reference T₂, its impact in comparing differences within the same tissue is expected to be minimal.

For the in-vivo analysis, mean qMRI values for some organs are available in the literature in healthy subjects and/or HNC patients before radiotherapy. For ADC, mean values ranged from 0.95 to $1.26 \times 10^{-3} \text{ mm}^2/\text{s}^2$ [14,15,16,21], 1.13 – 1.40 [14,15,16] and 1.52 – $1.58 \times 10^{-3} \text{ mm}^2/\text{s}^2$ [16] for the parotid, submandibular and sublingual glands were reported, respectively. The mean ADC values calculated in this study (0.85 , 1.05 and $1.03 \times 10^{-3} \text{ mm}^2/\text{s}^2$, respectively) were lower, which can be potentially attributed to differences in the choice of b-values, scanner (field strength and type) and acquisition protocols [31]. In this work, only the non-zero b-values were used for the ADC estimation, excluding the signal attenuation from perfusion in contrast to the majority of the studies using the 0b-value [14,15,16,21]. In addition, we used a SPLICE acquisition since it has been shown that SPLICE has higher geometrical accuracy at tissue-air interfaces compared to EPI [37]. However, SPLICE requires longer acquisition times which may cause larger uncertainties due to motion (swallowing and coughing) and lower signal to noise ratio. McDonald et al. [21] showed that for a group of HNC and healthy subjects, slightly different mean ADC values in the parotid glands were obtained for a SPLICE compared to an EPI acquisition at a 1.5 T MRI-Linac. For FF, mean values ranged from 29.4 % to 38.4 % for the parotid [17,38,39,40] and from 4.7 % to 13.8 % for the submandibular glands [37,38,39] were reported. T₂ values were only available for the parotid glands (72 ms) [17]. We found mean FF values of 29.3 % and 10.7 % for the parotid and submandibular glands, respectively and mean T₂ values of 77.5 ms for the parotid glands, which are consistent with the literature.

Furthermore, a significant effect of age in the FF and T₂ values was found. The effects of gender and BMI in all qMRI techniques were negligible. The observed trend of increasing qMRI values with advancing age may be attributed to age-related histopathological changes in the salivary glands and muscles, such as fatty infiltration, squamous and mucous metaplasia, atrophy and regeneration [41,42]. Our results suggest that the influence of age needs to be taken into account when evaluating changes of FF and T₂ values between subjects.

In-vivo repeatability qMRI analysis in HNC are rare. Nevertheless, with the introduction of MRI-guided systems more repeatability studies

Table 2

Mean and standard deviation (\pm SD) volume (vol) and qMRI values among all healthy subjects for the test session, within-subject coefficient of variation (wCV %) and repeatability coefficient (RC) for all techniques and ROIs. Results for the parotid, submandibular, sublingual and tubarial salivary glands include right and left glands (n = 52).

ROI	Vol (cc)		ADC ($10^{-3}\text{mm}^2/\text{s}$)		FF (%)			T ₂ (ms)		
	Mean \pm SD	Mean \pm SD	wCV (%)	RC	Mean \pm SD	wCV (%)	RC	Mean \pm SD	wCV (%)	RC
Parotid	18.5 \pm 5.2	0.85 \pm 0.11	5.5	0.13	30.2 \pm 11.0	5.0	3.6	78.0 \pm 7.6	1.5	3.3
Submandibular	7.3 \pm 1.6	1.05 \pm 0.10	6.7	0.20	12.6 \pm 7.5	11.1	3.6	69.5 \pm 6.0	2.4	4.6
Sublingual	2.6 \pm 0.7	1.03 \pm 0.11	8.4	0.24	14.1 \pm 6.9	11.9	3.9	69.1 \pm 7.3	3.3	6.3
Tubarial	2.5 \pm 0.6	1.08 \pm 0.11	5.4	0.11	7.4 \pm 3.0	25.8	4.8	65.2 \pm 5.0	3.5	6.2
OC	84.5 \pm 15.4	1.20 \pm 0.11	6.9	0.23	11.8 \pm 3.3	4.2	1.4	57.7 \pm 4.4	1.6	2.5
PCM	11.8 \pm 2.7	1.03 \pm 0.12	6.5	0.16	7.5 \pm 2.3	14.3	3.7	53.7 \pm 2.7	2.0	3.0
PCM _S	6.1 \pm 1.7	1.04 \pm 0.11	7.2	0.20	7.0 \pm 2.5	15.0	3.2	57.6 \pm 3.7	2.5	4.0
PCM _M	2.5 \pm 0.7	1.02 \pm 0.16	8.8	0.23	8.2 \pm 2.4	22.6	6.3	50.1 \pm 3.9	2.9	4.0
PCM _I	3.2 \pm 0.8	1.01 \pm 0.18	11.0	0.25	7.8 \pm 3.3	23.0	5.2	48.4 \pm 3.7	4.1	4.0
Brainstem	12.3 \pm 5.8	0.81 \pm 0.04	4.7	0.11	2.2 \pm 0.6	20.0	1.2	93.8 \pm 3.6	2.4	6.2

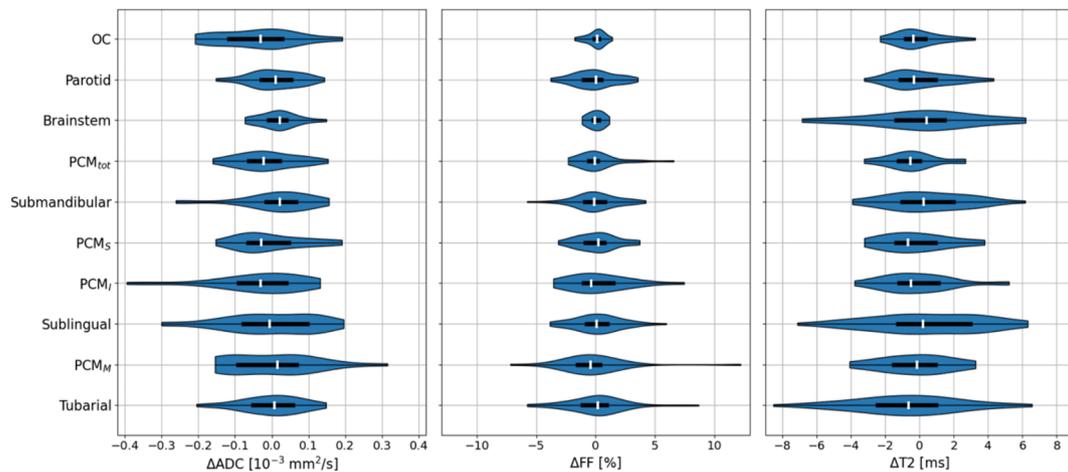


Fig. 3. Violin plots showing the distribution of the Δ qMRI values for all techniques and ROIs between test–retest session. Starting with the smallest ROI at the bottom, tubarial salivary glands, ascending to the largest volume at the top, oral cavity, OC. The black bar in the center of violin is the interquartile range (q1–q3), and the white vertical line represents the median.

Table 3

Fixed effects estimates per technique and ROI. Significant effects ($p < 0.05$) are denoted in bold with a *.

	ADC					FF					T ₂				
	Int.	Age	BMI	Sex	Session	Int.	Age	BMI	Sex	Session	Int.	Age	BMI	Sex	Session
Parotid	0.961	-0.001	-0.003	0.077*	-0.010	-5.79	0.49*	0.82	-4.12	0.24	56.15	0.39*	0.36	-2.40	-0.40
Submandibular	1.121	-0.000	-0.002	0.081*	-0.013	-13.60	0.35*	0.62	-3.06	-0.00	47.10	0.31*	0.50	-1.37	-0.44
Sublingual	0.785	-0.002	0.012*	0.003	0.007	-13.10	0.48*	0.41	-0.30	-0.22	52.63	0.48*	0.00	-0.90	-0.56
Tubarial	1.039	-0.003	0.002	-0.018	-0.002	3.90	0.13*	-0.07	1.50	-0.11	58.52	-0.06	0.37	-1.40	0.81
OC	1.061	-0.003	0.008	0.025	0.028	3.98	0.20*	0.03	-0.53	-0.02	46.50	0.25*	0.12	-1.73	0.15
PCM	1.090	0.003	-0.008	0.068	0.006	1.92	0.14*	0.03	-0.55	-0.25	47.97	0.03	0.16	0.59	0.37
PCM _S	0.964	0.005*	-0.003	-0.007	-0.007	0.83	0.15*	0.04	-0.18	-0.19	51.07	-0.42	0.32	0.45	0.01
PCM _M	1.168	0.001	-0.009	0.071	-0.004	2.98	0.09	0.06	0.39	0.03	41.23	0.03	0.27	1.30	0.33
PCM _I	1.133	0.001	-0.012	0.192*	0.032	3.40	0.17*	-0.01	-2.00	-0.44	38.81	0.20*	0.07	0.53	0.07
Brainstem	0.810	-0.001	0.001	0.033*	-0.021	1.28	0.00	0.04	-0.02	-0.02	91.05	-0.02	0.12	0.89	-0.33

Int. stands for intercept. For the continuous variables, age is incremented per year and BMI per unit. For the categorical variables, sex was set to male and session to test, as default.

during the course of radiotherapy are currently appearing in the literature. Paudyal et al. [19] reported a short-term wCV of 2.4 % for ADC in nine HN squamous cell carcinoma patients imaged pre-treatment with an EPI acquisition at a 3 T diagnostic scanner. At a 1.5 T MR-linac in eleven HNC patients using an EPI acquisition, Habrich et al. [20] denoted a RC range of 0.14–0.16 and 0.27–0.28 $\times 10^{-3}\text{mm}^2/\text{s}^2$ for the parotid and submandibular glands, respectively. Moreover, McDonald et al. [21] estimated a long-term wCV of 8.8 % for the parotid glands with multiple SPLICE acquisitions on consecutive days. The ADC repeatability estimated in this study was better than Habrich and

McDonald et al. and was slightly worse compared to Paudyal et al. In the latter, DWI acquisitions were performed twice on the same day compared to long-term measurements in this study (on average 17 days apart) which could potentially explain the larger ADC variations. Furthermore, no significant effect between test–retest measurements was estimated from the mixed-effect model analysis, confirming a good repeatability for all qMRI techniques in this work.

A limitation of this study is the absence of an immobilization mask for positioning, which enlarges the risk of motion and neck flexion during and in-between consecutive sessions. To minimize movement, a

head support was used during scanning [34]. Additionally, rigid registration between the image sets was performed before ROIs propagation. Afterwards, ROIs position was checked slice-by-slice on each qMRI map. In some cases, minor manual contour adjustments on the qMRI maps were necessary, especially for ROIs adjacent to air cavities, such as the PCM and the tubarial glands. The RC is an absolute value, however wCV % is a relative measure which depends on how large the ROI volume is. Therefore for small ROIs such as the PCM, low in fat content, the relative difference of the FF between test–retest session is considerably high (Table 2). Currently, strict cut-off values for repeatability do not exist resulting in uncertainties in the RCs assessment in terms of quality. As such, certain RCs can only be considered acceptable depending on the expected biological or physical change that is being quantified. In HNC, changes of ADC, FF and T_2 values in OARs during and after irradiation (≤ 1 year) up to 40 % for ADC [14–16] and 10 % for FF and T_2 [17] were reported in the literature. Considering the RC values determined in this study, all qMRI techniques demonstrated good repeatability in all evaluated structures.

To our knowledge, no studies regarding FF and T_2 mapping in-vivo repeatability for the HN region are currently available in the literature. Furthermore, for ADC the range of RC values for the structures examined in this study, excluding the parotid and submandibular glands, is currently lacking in the literature as well. This is the first work proving a range of RC and healthy qMRI values for a wide selection of OARs shown to be potentially associated with HNC late toxicity [26–30] in three different qMRI techniques. With the calculated RCs, the assessment of treatment response during the course of radiotherapy and the evaluation of late-toxicity with either ADC, FF and T_2 could be performed. Additionally, the observations obtained by qMRI in HNC patients, when correlated with the 3D dose distribution, could provide insights to the radiation-induced damage of all healthy structures surrounding the tumor simultaneously. This study brings a first-step towards this outcome. Nonetheless, validation studies across a larger patient cohort investigating sequential early and late qMRI changes versus reported RCs remain essential before clinical use of these qMRI techniques.

Funding statement

The research for this work was funded by Varian, a Siemens Healthineers Company (HollandPTC-Varian Consortium grant ID 2018021) and partly financed by the Surcharge for Top Consortia for Knowledge and Innovation (TKIs) from the Ministry of Economic Affairs and Climate, The Netherlands.

CRediT authorship contribution statement

F. Guerreiro: Validation, Formal analysis, Investigation, Visualization, Project administration, Data curation, Writing – original draft. **P.J. van Houdt:** Conceptualization, Investigation, Methodology, Resources, Formal analysis, Writing – review & editing. **R.J.M. Navest:** Investigation, Methodology, Resources, Writing – review & editing. **N. Hoekstra:** Data curation, Resources, Writing – review & editing. **M. de Jong:** Data curation, Resources, Writing – review & editing. **B.J. Heijnen:** Methodology. **S.E. Zijlema:** Investigation, Methodology, Resources, Writing – review & editing. **B. Verbist:** Data curation, Methodology, Funding acquisition, Resources, Writing – review & editing. **U.A. van der Heide:** Conceptualization, Investigation, Methodology, Writing – review & editing, Funding acquisition, Supervision. **E. Astreinidou:** Conceptualization, Methodology, Validation, Data curation, Formal analysis, Investigation, Resources, Writing – review & editing, Funding acquisition, Supervision.

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.

Acknowledgements

The authors would like to thank Bart Mertens, Associate professor, Biomedical Data Sciences at Leiden University for his expertise and assistance throughout aspects of the statistical analysis in this study.

References

- [1] Baudelet M, Van den Steen L, Tomassen P, Bonte K, Deron P, Huvenne W, et al. Very late xerostomia, dysphagia, and neck fibrosis after head and neck radiotherapy. *Head Neck* 2019;41(10):3594–603. <https://doi.org/10.1002/hed.25880>.
- [2] Vainshtein JM, Vainshtein JM, Samuels S, Tao Y, Lyden T, Haxer M, et al. Impact of xerostomia on dysphagia after chemotherapy- intensity-modulated radiotherapy for oropharyngeal cancer: Prospective longitudinal study. *Head Neck* 2016;38(1): 1605–12. <https://doi.org/10.1002/hed.24286>.
- [3] Langendijk JA, Doornaert P, Verdonck-de Leeuw IM, Leemans CR, Aaronson NK, Slotman BJ. Impact of late treatment related toxicity on quality of life among patients with head and neck cancer treated with radiotherapy. *J Clin Oncol* 2008; 26(22):3770–6. <https://doi.org/10.1200/JCO.2007.14.6647>.
- [4] Roesink JM, Moerland MA, Hoekstra A, Van Rijk PP, Terhaard CH. Scintigraphic assessment of early and late parotid gland function after radiotherapy for head-and-neck cancer: a prospective study of dose-volume response relationships. *Int J Radiat Oncol Biol Phys* 2004;58:1451–60. <https://doi.org/10.1016/j.ijrobp.2003.09.021>.
- [5] Roesink JM, Moerland MA, Battermann JJ, Hordijk GJ, Terhaard CH. Quantitative dose-volume response analysis of changes in parotid gland function after radiotherapy in the head-and-neck region. *Int J Radiat Oncol Biol Phys* 2001;51(4): 938–46. [https://doi.org/10.1016/S0360-3016\(01\)01717-5](https://doi.org/10.1016/S0360-3016(01)01717-5).
- [6] Bath SS, Caudell JJ, Chen AM. Practical considerations in reducing swallowing dysfunction following concurrent chemoradiotherapy with intensity-modulated radiotherapy for head and neck cancer. *Head Neck* 2014;36(2):291–8. <https://doi.org/10.1002/hed.23246>.
- [7] Jensen K, Lamberts K, Torkov P, Dahl M, Jensen AB, Grau C. Patient assessed symptoms are poor predictors of objective findings. Results from a cross sectional study in patients treated with radiotherapy for pharyngeal cancer. *Acta Oncol* 2007;46(8):1159–68. <https://doi.org/10.1080/02841860701491041>.
- [8] Håkansson K, Smulders B, Specht L, Zhu M, Friborg J, Rasmussen JH, et al. Radiation dose-painting with protons vs. photons for head-and-neck cancer. *Acta Oncol* 2020;59(5):525–33. <https://doi.org/10.1080/0284186X.2020.1714720>.
- [9] Landelijk Platform Protontherapie (LPPT) Landelijk Indicatieprotocol Protontherapie Hoofdhals. 2017. Available online: <http://www.nvro.nl/publicaties/rapporten.html>.
- [10] Lagendijk JJW, Raaymakers BW, Van den Berg CAT, Moerland MA, Philippens ME, van Vulpen M. MR guidance in radiotherapy. *Phys Med Biol* 2014;59(21):349–69. <https://doi.org/10.1088/0031-9155/59/21/R349>.
- [11] Ng J, Gregucci F, Pennell RT, Nagar H, Golden EB, Knisely JPS, et al. MRI-LINAC: A transformative technology in radiation oncology. *Front Oncol* 2023;13:1117874. <https://doi.org/10.3389/fonc.2023.1117874>.
- [12] Press RH, Shu H, Shim H, Mountz JM, Kurland BF, Wahl RL, et al. The use of quantitative imaging in radiation oncology: A quantitative imaging network (QIN) perspective. *Int J Radiat Oncol Biol Phys* 2018;102:1219–35. <https://doi.org/10.1016/j.ijrobp.2018.06.023>.
- [13] Hall WA, Paulson ES, van der Heide UA, Fuller CD, Raaymakers BW, Lagendijk JJW, et al. The transformation of radiation oncology using real-time magnetic resonance guidance: A review. *Eur J Cancer* 2019;122:42–52. <https://doi.org/10.1016/j.ejca.2019.07.02>.
- [14] Loimu V, Seppälä T, Kapanen M, Tuomikoski L, Nurmi H, Mäkitie A, et al. Diffusion-weighted magnetic resonance imaging for evaluation of salivary gland function in head and neck cancer patients treated with intensity- modulated radiotherapy. *Radiother Oncol* 2017;122(2):178–84. <https://doi.org/10.1016/j.radonc.2016.07.008>.
- [15] Zhang Y, Ou D, Gu Y, He X, Peng W. Evaluation of salivary gland function using diffusion-weighted magnetic resonance imaging for follow-up of radiation-induced xerostomia. *Korean J Radiol* 2018;19(4):758–66. <https://doi.org/10.3348/kjr.2018.19.4.758>.
- [16] Duffton A, Kemp O, Devlin L, Hay L, McLoone P, Paterson C. Feasibility of DW-MRI analysis of salivary glands during head and neck radiotherapy. *Tech Innov Patient Support Radiat Oncol* 2021;19:46–51. <https://doi.org/10.1016/j.tipsro.2021.07.002>.
- [17] Zhou N, Chu C, Dou X, Chen W, He J, Yan J, et al. Early evaluation of radiation-induced parotid damage in patients with nasopharyngeal carcinoma by T2 mapping and mDIXON Quant imaging: initial findings. *Radiat Oncol* 2018;13(1): 22. <https://doi.org/10.1186/s13014-018-0970-9>.
- [18] Shukla-Dave A, Obuchowski NA, Chenevert TL, Jambawalikar S, Schwartz LH, Malyarenko D, et al. Quantitative imaging biomarkers alliance (QIBA) recommendations for improved precision of DWI and DCE-MRI derived biomarkers in multicenter oncology trials. *J Magn Reson Imag* 2019;49:e101–21. <https://doi.org/10.1002/jmri.26518>.

- [19] Paudyal R, Konar AS, Obuchowski NA, Hatzoglou V, Chenevert TL, Malyarenko DI, et al. Repeatability of quantitative diffusion-weighted imaging metrics in phantoms, head-and-neck and thyroid cancers: preliminary findings. *Tomography* 2019;5(1):15–25. <https://doi.org/10.18383/j.tom.2018.00044>.
- [20] Habrich J, Boeke S, Nachbar M, Nikolaou K, Schick F, Gani C, et al. Repeatability of diffusion-weighted magnetic resonance imaging in head and neck cancer at a 1.5 T MR-Linac. *Radiother Oncol* 2022;174:141–8. <https://doi.org/10.1016/j.radonc.2022.07.020>.
- [21] McDonald BA, Salzillo T, Mulder S, Ahmed S, Dresner A, Preston K, et al. Prospective evaluation of in vivo and phantom repeatability and reproducibility of diffusion-weighted MRI sequences on 1.5T MRI-linear accelerator (MR-Linac) and MR simulator devices for head and neck cancers. *Radiother Oncol* 2023;185. <https://doi.org/10.1016/j.radonc.2023.109717>.
- [22] Kooreman ES, van Houdt PJ, Nowee ME, van Pelt VWJ, Tijssen RHN, Paulson ES, et al. Feasibility and accuracy of quantitative imaging on a 1.5 T MR-linear accelerator. *Radiother Oncol* 2019;133:156–62. <https://doi.org/10.1016/j.radonc.2019.01.011>.
- [23] van Houdt PJ, Agarwal HK, van Buuren LD, Heijmink SWTPJ, Haack S, van der Poel HG, et al. Performance of a fast and high-resolution multi-echo spin-echo sequence for prostate T2 mapping across multiple systems. *Magn Reson Med* 2017;79:1586–94. <https://doi.org/10.1002/mrm.26816>.
- [24] Bachtiar V, Kelly MD, Wilman HR, Jacobs J, Newbould R, Kelly CJ, et al. Repeatability and reproducibility of multiparametric magnetic resonance imaging of the liver. *PLoS One* 2019;14(4):e0214921. <https://doi.org/10.1371/journal.pone.0214921>.
- [25] Sakamoto J, Yoshino N, Okochi K, Imaizumi A, Tetsumura A, Kurohara K, et al. Tissue characterization of head and neck lesions using diffusion-weighted MR imaging with SPLICE. *Eur J Radiol* 2009;69(2):260–8. <https://doi.org/10.1016/j.ejrad.2007.10.008>.
- [26] Beetz I, Schilstra C, Burlage FR, Koken PW, Doornaert P, Bijl HP, et al. Development of NTCP models for head and neck cancer patients treated with three-dimensional conformal radiotherapy for xerostomia and sticky saliva: The role of dosimetric and clinical factors. *Radiother Oncol* 2012;105(1):86–93. <https://doi.org/10.1016/j.radonc.2011.05.010>.
- [27] Christianen MEMC, Schilstra C, Beetz I, Muijs CT, Chouvalova O, Burlage FR, et al. Predictive modelling for swallowing dysfunction after primary (chemo)radiation: Results of a prospective observational study. *Radiother Oncol* 2012;105(1):107–14. <https://doi.org/10.1016/j.radonc.2011.08.009>.
- [28] Van den Bosch L, van der Schaaf A, van der Laan HP, Hoebbers FJP, Wijers OB, van den Hoek JGM, et al. Comprehensive toxicity risk profiling in radiation therapy for head and neck cancer: A new concept for individually optimised treatment. *Radiother Oncol* 2021;157:147–54. <https://doi.org/10.1016/j.radonc.2021.01.024>.
- [29] Valstar MH, de Bakker BS, Steenbakkens RJHM, de Jong KH, Smit LA, Klein Nulent TJW, et al. The tubarial glands: A potential new organ at risk for radiotherapy. *Radiother Oncol* 2021;154:292–1228. <https://doi.org/10.1016/j.radonc.2020.09.034>.
- [30] Osorio EV, Abravan A, Green A, van Herk M, Lee LW, Ganderton D, et al. Dysphagia at 1 year is associated with mean dose to the inferior section of the brain stem. *Int J Radiat Oncol Biol Phys* 2023;00:1–11. <https://doi.org/10.1016/j.ijrobp.2023.06.004>.
- [31] Bruvo M, Mahmood F. Apparent diffusion coefficient measurement of the parotid gland parenchyma. *Quant Imaging Med Surg* 2021;11(8):3812–29. <https://doi.org/10.21037/qims-20-1178>.
- [32] Bos C, Duijndam A, Senegas J. Reference phantom validation of T2-mapping: maximum likelihood estimation of T2 from magnitude phased-array multi-echo data. In Proceedings of the 17th Annual Meeting of ISMRM, Honolulu, Hawaii, USA, 2009. p 4526.
- [33] Milford D, Rosbach N, Bendszus M, Heiland S. Mono-exponential fitting in T2-relaxometry: relevance of offset and first echo. *PLoS One* 2015;10:1–13. <https://doi.org/10.1371/journal.pone.0145255>.
- [34] Frans W, Frantzen-Stenekker M, Osinga J, van der Heide UA, Navest R. OC-0091 head coil head support: in support of the head and the head coil. *Radiother Oncol* 2021;161(1):66–7. [https://doi.org/10.1016/S0167-8140\(21\)06785-2](https://doi.org/10.1016/S0167-8140(21)06785-2).
- [35] van Houdt PJ, Kallehauge JF, Tanderup K, Nout R, Zaletel J, Tadic T, et al. Phantom-based quality assurance for multicenter quantitative MRI in locally advanced cervical cancer. *Radiother Oncol* 2020;153:114–21. <https://doi.org/10.1016/j.radonc.2020.09.013>.
- [36] Hu HH, Yokoo T, Bashir MR, Sirlin CB, Hernando D, Malyarenko D, et al. Linearity and bias of proton density fat fraction as a quantitative imaging biomarker: a multicenter, multiplatform. *Multivendor Phantom Study Radiology* 2021;298(3):640–51. <https://doi.org/10.1148/radiol.2021202912>.
- [37] Schakel T, Hoogduin JM, Terhaard CHJ, Philippens MEP. Technical Note: Diffusion-weighted MRI with minimal distortion in head-and-neck radiotherapy using a turbo spin echo acquisition method. *Med Phys* 2017;44(8):4188–93. <https://doi.org/10.1002/mp.12363>.
- [38] Kise Y, Chikui T, Yamashita Y, Kobayashi K, Yoshiura K. Clinical usefulness of the mDIXON quant method for estimation of the salivary gland fat fraction: comparison with MR spectroscopy. *Br J Radiol* 2017;90(1077):20160704. <https://doi.org/10.1259/bjr.20160704>.
- [39] Su GY, Wang CB, Hu H, Liu J, Ding HY, Xu XQ, et al. Effect of laterality, gender, age and body mass index on the fat fraction of salivary glands in healthy volunteers: assessed using iterative decomposition of water and fat with echo asymmetry and least-squares estimation method. *Dentomaxillofac Radiol* 2019;48(3):20180263. <https://doi.org/10.1259/dmfr.20180263>.
- [40] Chu C, Feng Q, Zhang H, Zhao S, Chen W, He J, et al. Evaluation of salivary gland fat fraction values in patients with primary Sjögren's syndrome by mDIXON quant imaging: Initial findings. *Eur J Radiol* 2020;123:108776. <https://doi.org/10.1016/j.ejrad.2019.108776>.
- [41] Scott J, Flower EA, Burns J. A quantitative study of histological changes in the human parotid gland occurring with adult age. *J Oral Pathol* 1987;16:505–10. <https://doi.org/10.1111/j.1600-0714.1987.tb00681.x>.
- [42] Dayan D, Vered M, Paz T, Buchner A. Aging of human palatal salivary glands: a histomorphometric study. *Exp Gerontol* 2000;35:85–93. [https://doi.org/10.1016/S0531-5565\(99\)00079-0](https://doi.org/10.1016/S0531-5565(99)00079-0).