HHS Public Access

Author manuscript

*Phys Imaging Radiat Oncol*. Author manuscript; available in PMC 2018 January 01.

Published in final edited form as:

*Phys Imaging Radiat Oncol*. 2017 January ; 1: 34–40. doi:10.1016/j.phro.2017.02.006.

# A Magnetic Resonance Imaging-based approach to quantify radiation-induced normal tissue injuries applied to trismus in head and neck cancer

**Maria Thor**[1], **Neelam Tyagi**[1], **Vaios Hatzoglou**[2], **Aditya Apte**[1], **Ziad Saleh**[1], **Nadeem Riaz**[3], **Nancy Y Lee**[3], and **Joseph O Deasy, PhD**[1]

[1]Dept. of Medical Physics, Memorial Sloan Kettering Cancer Center, New York, US

[2]Dept. of Radiology, Memorial Sloan Kettering Cancer Center, New York, US

[3]Dept. of Radiation Oncology, Memorial Sloan Kettering Cancer Center, New York, US

## Abstract

**Background and Purpose—**In this study we investigated the ability of textures from T1-weighted MRI scans post-contrast ($T1w_{post}$) to identify the critical muscle(s) for radiation-induced trismus.

**Materials and Methods—**The study included ten cases (Trismus: ≥Grade 1), and ten age-sex-tumor-location-and-stage-matched controls treated with intensity-modulated radiotherapy to 70Gy@2.12Gy in 2005–2009. Trismus status and $T1w_{Post}$ were conducted within one year post-radiotherapy. For the masseter, lateral and medial pterygoids, and temporalis (*M/LP/MP/T*), 24 textures were extracted (Grey Level Co-Occurrence (*GLCM), Histogram*, and *Shape*). Univariate logistic regression with Bootstrapping (1000 populations) was applied to compare the muscle mean dose (Dmean) and textures between cases and controls (ipsilateral muscles); candidate predictors were suggested by an average p≤0.20 across all Bootstrap populations.

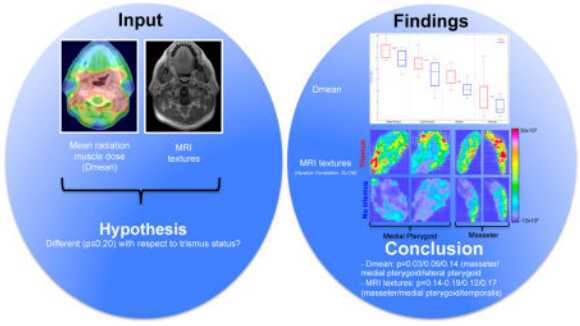
**Results—**Dmean to *M*/*LP*/*MP* (p=0.03/0.14/0.09), one *MP/T* (p=0.12/0.17), and three *M* (p=0.14–0.19) textures were candidate predictors. Three of these textures were *GLCM*- and two *Histogram* textures with the former being generally higher and the latter lower for cases compared to controls. The Dmean to *M* and MP, and *Haralick Correlation* (*GLCM*) of *MP* presented with the best discriminative ability (area under the receiver-operating characteristic curve: 0.85, 0.77, and 0.78), and the correlation between Dmean and this texture was weak (Spearman's rank correlation coefficient: 0.26–0.27).

**Conclusions—**Our exploratory study points towards an interplay between the dose to the masseter, and the medial pterygoid together with the local relationship between the mean MRI intensity relative to its variance of the medial pterygoid for radiation-induced trismus. This opens

Corresponding author: Maria Thor, Ph.D., Dept. of Medical Physics, Memorial Sloan Kettering Cancer Center 485 Lexington Avenue, New York, NY 10012 Tel: (+001) 646 888 8013, thorm@mskcc.org.

up for exploration of this interplay within the radiation-induced trismus etiology in the larger multi-institutional setting.

## TOC Image




### Keywords

head and neck cancer; radiotherapy; trismus; quantitative image; texture; radiomics

## Introduction

Primary or adjuvant radiotherapy (RT) is the standard of care for individuals diagnosed with head and neck cancer (HNC) [1]. Owing to a high five-year relative survival using state-of-the-art RT planning and delivery techniques such as Intensity-modulated RT (IMRT), and the consequently sizeable number of HNC survivors, avoidance of radiation-induced oral complications is increasing in importance [1–3].

Mouth-opening limitation *i.e.* trismus is observed in up to 50% of all HNC survivors, and progresses most dramatically between three and twelve months post-RT manifesting as eating and speech difficulties, as well as impaired oral hygiene, altered facial expression, and/or pain [1, 4–7]. The structures and related RT dose thresholds involved in the trismus etiology are unsettled, but it has been proposed that trismus following RT primarily results from fibrosis of the masticatory muscles [1, 4]. Furthermore, five recent studies have pointed out that trismus is in particular explained by the dose to tumor-ipsilateral masticatory muscles [5, 8–11]. By means of magnetic resonance imaging (MRI) both Bhatia *et al* [12], and Hsieh *et al* [13] qualitatively identified abnormalities in multiple masticatory muscles that were associated with trismus. However, quantitative image measures may provide additional information, and identified metrics are likely to be more reproducible compared to qualitatively identified metrics [14, 15]. Thus, quantitative MRI metrics may add to the understanding of RT-induced trismus both by providing a measurable endpoint, as well as elucidating the trismus etiology.

Bearing in mind a multi musculature- and pathophysiological process behind RT-induced trismus, we in this exploratory study anticipate that quantitative metrics derived from MRI scans after gadolinium administration in any of the four major masticatory muscles following RT is associated with clinically observed radiation-induced trismus. The goal was to identify the muscles being critical for trismus, and the related intra-muscular intensity

patterns. We illustrated this concept within a case control study including ten cases and ten controls.

## Materials and Methods

### Study design and follow-up

This retrospective and Institutional Review Board approved case-control study included subjects previously treated with primary RT for squamous cell carcinoma in the head and neck region at the Memorial Sloan Kettering Cancer Center in 2005–2009 [9]. An overview of the characteristics for the included subjects is given in Table 1. Treatment was planned and delivered using IMRT with a 7-field arrangement and a two-phase approach with a total dose of 70 Gy in 33 fractions prescribed to the primary tumor. Trismus was assessed post-RT using the CTCAE.v.4 scoring system [16], and was in this study defined as having at least "*Decreased range of motion without impaired eating*" (≥Grade 1).

To qualify for inclusion, trismus had to be assessed within the first year post-RT along with simultaneously acquired (±two months) T1-weighted MRI scans post-contrast injection ($T1w_{post}$). Subjects were excluded if trismus was present before RT. In total, ten cases fulfilled the inclusion criteria as identified from the research database. Each case was matched with a control based on age, gender, primary tumor location, and stage. As illustrated in Table 1, the 20 included subjects presented primarily with locally advanced oropharyngeal tumors (base of tongue/tonsil: n=10/8), were mainly men (n=16) with an average age of 56 years (standard deviation, SD: ±7 years); all in accordance with the typical demographics for squamous cell carcinoma of the head and neck [2]. The median follow-up time on trismus and the median time to MRI acquisition were both three months (range: 3–12, 3–10 months, respectively). Age, follow-up time, time to MRI acquisition, and tumor stage were not statistically significantly different with respect to trismus status (Wilcoxon signed-rank test: p=0.56–1.00). Follow-up time, as well, was not statistically significantly different from the time to MRI acquisition (p=0.35).

### Comparison of MRI intensity for masticatory muscles

All scans were obtained on a 1.5T GE scanner. For each subject, a two-dimensional (2D) axial spin echo T1-weighted MRI sequence was used (Echo/Repetition time: 550/12 ms; matrix size: 256×256; slice thickness: 4 mm), both before and after gadolinium contrast administration. For the purpose of this study, we focused on the uptake of gadolinium and, therefore, the $T1w_{Post}$ scans. The level of gadolinium was on average 7 ml (SD: ±1ml), and was based on each patient's weight. There was no statistical significant difference with respect to the amount of gadolinium between cases and controls (Wilcoxon signed-rank test: p=1.00; Table 1). All four masticatory muscles – masseter, lateral and medial pterygoid, and temporalis – were segmented under supervision of a board certified radiologist (V.H), and were segmented separately on the left and the right side. For consistency, the cranial boarder of the temporalis was truncated to the first slice below the orbit since the field-of-view did not cover the entire muscle in some subjects; all other muscles were defined in the slices of their visibility (3–11 slices). To assess the mean RT dose (Dmean) to all muscles these were segmented in the $T1w_{Post}$ scans and propagated onto the Computed Tomography (CT) scan

used for RT planning following affine and deformable image registration (Eclipse SmartAdapt, Varian Medical Systems v.13, Palo Alto, CA, US), with the latter utilizing an intensity-based *modified* Demon's algorithm [17], followed by a dose recalculation using our clinically applied RT treatment planning settings (Eclipse AAA v.13.0.31, Varian Medical Systems, Palo Alto, CA, US). An overview of the path of analysis applied is given in Figure S1.

The MRI intensity was quantified by a total of 24 textures: three shape-based (*Shape*), twelve first order histogram-based (*Histogram*), and nine second order Haralick grey level co-occurrence-based (*GLCM*) textures [18–23]. The representation of these textures was based on their implementation in CERR [24]; further details are available here: https://github.com/adityaapte/CERR. All textures were in addition studied as the ratio relative to a subject-specific and non-irradiated structure (a portion of the posterior, right temporal lobe; extension: 2–3 slices) to enable comparison of these textures across patients [25]. No apparent abnormalities were observed within this non-irradiated structure as confirmed by qualitatively evaluating the simultaneously acquired T2-weighted MRI scans. The *GLCM* textures were calculated within patches in 2D (in-plane; patch size: 3×3 pixels) using 16 grey levels, and the median value over all patches represented the MRI intensity for these textures. An overview of all 24 textures and their definitions is given in the Supplementary material.

Previous dose-response-focused studies have implicated that predominantly tumor-ipsilateral representations of the masticatory muscles are important to explain the RT-induced trismus etiology [5, 8–10]. We, therefore, considered ipsilateral representations as the most critical determinant for the development of trismus. More specifically, we compared Dmean and textures with respect to trismus status for ipsilateral representations ($Dmean_{ipsi-contra}$>5 Gy for the corresponding muscle). In case of non-distinct ipsilaterality, bilateral representations were used where the median Dmean and texture of paired muscles was considered. The Dmean and textures were compared using a univariate logistic regression approach and, given the fairly small number of included patients, we applied an internal robustness approach using Bootstrapping with replacement (1000 sample populations). Performance was evaluated by the area under the receiver-operating characteristic curve (AUC), and Dmean and textures presenting with an average p-value ≤0.20 across all Bootstrap samples were considered candidate predictors. The reported AUCs and p-values are given as the average±SD across all Bootstrap populations. Correlation was evaluated between candidate predictors with high correlation defined by a Spearman's rank correlation coefficient (|Rs|) ≥0.70.

## Results

### Candidate trismus predictors

The Dmean to the masseter, as well as the lateral and the medial pterygoid was a candidate predictor (p=0.03, 0.14, and 0.09; AUC=0.85, 0.76, and 0.77), and was significantly higher for cases compared to controls. The corresponding difference in Dmean between cases and controls ranged between 0.50–26 Gy, 0.80–19 Gy, and 1.0–21.0 Gy for the masseter, lateral and medial pterygoid, respectively (Figure 1; Table 2). The correlation between Dmean to

the lateral pterygoid and either that to the masseter, or the medial pterygoid was, however, strong (|Rs|=0.78, or 0.79), whilst that between the masseter and the medial pterygoid was somewhat more modest (|Rs|=0.66). For these reasons, Dmean to the lateral pterygoid was not considered a final candidate predictor.

Among the textures, five candidate predictors were initially identified (population median (range): AUC=0.71 (0.68–0.78); p=0.17 (0.12–0.19); Figure 2; Table 2). These textures included *Haralick correlation (GLCM)* of the masseter and of the medial pterygoid, *Root Mean Square (Histogram)*, and *Sum Average (GLCM)* of the masseter, as well as *Skewness (Histogram)* of the temporalis. The three *GLCM textures* were on average higher for cases compared to controls, whilst the two *Histogram* textures were lower for cases compared to controls. A similar pattern of *GLCM* textures being higher and *Histogram* textures being lower for cases was also observed for textures not suggested as candidate predictors (Table S1). In general, the correlation between textures of the same domain for the corresponding muscle was high (Figure S2). However, amongst the initial candidate predictors, the texture-to-texture correlation was weak (population median (range) |Rs|=0.20 (0.10–0.91)), and only *Sum Average* (*GLCM*) of the masseter was highly correlated with *Haralick Correlation* (*GLCM*) of the same muscle (|Rs|=0.91), and since the latter texture presented with a somewhat better discriminative ability compared to the former (AUC=0.72 vs. 0.70; p=0.14 vs. 0.19;), *Sum Average (GLCM)* was not considered a final candidate predictor. The best discriminative ability among the four final candidate texture predictors was observed for *Haralick Correlation (GLCM)* of the medial pterygoid (AUC=0.78; p=0.12). The distribution of this texture is illustrated both within the medial pterygoid and the masseter for cases and controls in Figure 3. The median correlation between any of the four final candidate texture predictors and Dmean to either the masseter or the medial pterygoid was weak (|Rs|=0.18), and a modest correlation was observed only between Dmean to the masseter and *Root Mean Square (Histogram)* of the masseter (|Rs|=0.55).

## Discussion

This study has focused on identifying the muscles being critical for radiation-induced trismus and the related intra-muscular intensity patterns using T1-weighted MRI scans acquired after gadolinium administration ($T1w_{post}$). This was demonstrated within a case control design including a total of 20 subjects, and should as such primarily be regarded exploratory. With this in mind, the findings point primarily towards an involvement in the radiation-induced trismus etiology of the dose to the masseter, together with a trend of the dose to the medial pterygoid and the local dispersion of the mean relative the variance of the MRI intensity within the medial pterygoid as described by the *Haralick (GLCM)* texture.

Trismus following RT is primarily a consequence of perpetual masticatory muscle(s) contraction due to RT-induced fibrosis [1, 4]. Fibrosis manifests initially in a similar way as wound healing, but progressive accumulations lead to fibrotic changes *i.e.* decreased tissue elasticity, and may cause distortion of tissues, reduction in joint movement, distal lymphedema, neuropathy, and/or pain [26]. In an ideal scenario, fibrosis should be detected prior to the manifestation of trismus since once trismus is established the fibrosis causing the restriction is difficult to manage [1, 26]. Interestingly, in two previous studies muscle

abnormalities post-RT for HNC were qualitatively identified from conventional MRI sequences [12, 13]. In the study by Bhatia *et al* [12] abnormalities of the muscles involved in jaw mobility (high intensity within T2-weighted scans, or intensity/striking enhancement on $T1w_{post}$ scans) were observed in 19 of their 35 investigated subjects that all experienced some degree of trismus (dental gap ≤25mm), and were primarily identified in the lateral and the medial pterygoid (N=16, 11) followed by the masseter and the temporalis (N=5, 4). In the study by Hsieh *et al* [13], muscle abnormalities of various degree (Grade 1: high intensity within T2-weighted scans, Grade 2: Grade 1, and muscle atrophy/obvious enhancement within $T1w_{post}$ scans) were identified in all 22 subjects, and the resulting qualitative MRI-based signal abnormality score (SA) was moderately correlated with the clinical staging of trismus (CTCAE.v.3; Rs=0.52). In addition, Dmean to all masticatory muscles combined, and Dmean to the lateral pterygoid was significantly higher for subjects with a worse SA compared to those with a more favorable SA (p=0.05, 0.04). Added together, these two studies implicate a multi-musculature involvement in the RT-induced trismus etiology with an emphasis on the lateral and the medial pterygoid but it should be pointed out that the MRI abnormalities were qualitatively identified.

Using the same type of MRI sequence as in the aforementioned two studies [12, 13], but representing instead the MRI intensity using a more quantitative approach, we also identified a trend of a multi-musculature involvement in the development of trismus but between textures of the masseter, the medial pterygoid, and the temporalis with one to two identified textures per muscle being candidates for predicting trismus. In particular, *Haralick Correlation (GLCM)* of the medial pterygoid presented with the best discriminative ability (AUC=0.78; p=0.12; averaged across all Bootstrap samples), and showed a tendency of being higher for cases compared to controls (population median: −1.3 vs. −2.6). The same texture but of the masseter was the second best candidate predictor (population median: −1.5 vs. −3.0, AUC=0.72; p=0.14). This texture describes the ratio between the mean and the variance of the intensity within local patches [18–23], and is one of two out of the 24 investigated textures that explicitly addresses this relationship, and considering that we monitored the uptake of gadolinium this result is in line with the qualitatively identified increased MRI intensity in the $T1w_{post}$ scans as observed in [12], and [13]. Our residual two final candidate textures *i.e. Root Mean Square* of the masseter, and *Skewness* of the temporalis from the histogram domain were instead lower for cases compared to controls (population median: −0.1 vs. 0.0, and 1.1 vs. 1.3), but differences were less pronounced compared to the two *GLCM* textures. A potential explanation to this could be that *Histogram* textures in general involve a lower degree of spatial image complexity compared to *GLCM* textures, which may further support the increased ability of textures from the latter domain to capture significantly different patterns. The inability of the three investigated *Shape* textures to explain differences with respect to trismus status and muscle laterality could be an indication that fibrosis rather than atrophy is the key RT-induced muscle injury. However, since this study was based on in total 20 subjects, we encourage efforts within the larger, multi-institutional setting as pointed out in *e.g.* [27] to explore the robustness of these findings to explain radiation-induced trismus. Also, this study was based on contrast-enhanced MRI scans acquired after RT. The amount of contrast was not statistically significantly different with respect to trismus status and ranged between four and

nine milliliters across our cohort, and thus together with the normalization to a non-irradiated scan-specific structure supports the quantitative nature of the investigated textures. Including only MRI scans post-RT was justified as a surrogate for radiation-induced trismus given that no subject presented with trismus before RT, but an improved study design to discriminate the precise radiation-induced textural patterns would involve scans acquired both before and after RT.

The mean dose (Dmean) to all muscles except for the temporalis were candidate predictors, but only the mean dose to the masseter was strictly significantly higher for cases compared to controls (p=0.03; AUC=0.85) whilst a similar trend was observed for the medial pterygoid (p=0.09; AUC=0.77). The Dmean to the lateral pterygoid was highly correlated with that of both the masseter and the medial pterygoid, and presented with a lower p-value (p=0.14; AUC=0.76). Previously, Dmean has been the most widely used dose representation for masticatory muscles for the dose-response modeling of trismus [5, 8–11], and has been shown to describe its development equally well as that of other dose metrics [5, 9]. Following IMRT for HNC, Rao *et al* [9] found that Dmean to the ipsilateral medial pterygoid, or to the ipsilateral masseter best described trismus (CTCAE v.4: ≥Grade 1; median follow-up time: 33 months; AUC: 0.71, 0.73) on univariate analysis. Significant Dmean associations with a lower discriminative ability were observed for the ipsilateral and contralateral lateral pterygoid, the ipsilateral temporalis, and the contralateral masseter (AUC range: 0.60–0.65). However, on multivariate analysis only Dmean of the medial pterygoid remained a significant predictor for trismus (p=0.03). Van der Molen *et al* [5] also observed that Dmean to the masseter, and Dmean to the pterygoid muscles were significant predictors for trismus (maximum interincisor moth opening≤35mm; follow-up: ten weeks post-IMRT), but only Dmean of the masseter could be confirmed as an independent predictor using a questionnaire-based definition of trismus. The disagreement between the two assessments could to some degree be explained by the use of bilateral representations, and in addition by the combination of the lateral and medial pterygoid into one structure. Across the investigated muscles, the overall highest Dmean in this study, and in [9] and [10], was that of the ipsilateral medial pterygoid, which was followed by Dmean to the ipisalteral lateral pterygoid here and in [9]. The fibers of the four muscles investigated have a pinnate formation in which the muscle fascicles are attached obliquely. In contrast to the medial pterygoid and the masseter, which are built up by fibers oriented in multiple angles, the lateral pterygoid and the temporalis are formed by fibers oriented in one angle [28]. Therefore, the more complex architecture of the medial pterygoid and the masseter compared to the lateral pterygoid and temporalis could implicate a more complicated repair process of RT-induced fibrosis in the two former muscles. This together with the highest Dmean levels as observed for the medial pterygoid could explain that our overall best discriminative ability was observed for MRI textures of the medial pterygoid. It should also be pointed out that the correlation between any of our four final candidate texture predictors and Dmean to either the masseter or the medial pterygoid was in general weak (median |Rs|: 0.18). Hence, even if a strictly significant association between trismus and any of the textures was not identified, the weak correlations between textures and dose motivate added information from the MRI intensity over using dose only.

The subjects in our cohort included mainly men with primary tumors of the oropharynx (base of tongue, and tonsil), and an age range of 47–71 years. This cohort reflects the current trend of HNC since the estimated rates of new HNC cases in 2016 is up to five times higher among men compared to women, and the highest incidence of new HNC cases cover the investigated age range [2]. The exact applicability of our results in cohorts including subjects presenting with diverging characteristics, however, remains unclear. Our trismus definition was a combination of eating- and jaw mobility limitations since seven of the ten cases presented with Grade 1, one with Grade 2, and two with Grade 3, which corresponded to "*Increased range of motion without impaired eating*", "*Decreased range of motion requiring small bites/soft foods/purees*", and "*Decreased range of motion with inability to adequately aliment/hydrate orally*", respectively. Even though the RT-induced trismus etiology is likely multifactorial, and by such involves in addition to eating and jaw mobility limitations, also muscular tension, and/or pain [1, 4–7], it has been demonstrated that the rate of eating and jaw mobility limitations following IMRT for HNC is higher than the other domains, and significantly higher up to six months after compared to before IMRT [7].

In conclusion, this study demonstrates an image- and dose based rather than the widely used dose-based only approach to assess dose-response relationships and more specifically by quantifying intensities from T1-weighted MRI scans after gadolinium administration as a means to measure radiation-induced normal tissue injury. Although limited to a cohort of 20 subjects, our findings pointed towards a multi-masticatory involvement for trismus following IMRT for HNC with an emphasis on the dose to the masseter and the medial pterygoid, as well the local relationship between the mean and the variance of the MRI intensity within the medial pterygoid. Thus, this study opens up for exploration of 'the masseter and medial pterygoid dose and MRI intensity interplay' within the radiation-induced trismus etiology in the larger multi-institutional setting.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## References

1. Epstein JB, Thariat J, Bensadoun RJ, Barasch A, Murphy BA, Kolnick L, Popplewell L, Maghami E. Oral complications of cancer and cancer therapy: from cancer to treatment to survivorship. CA Cancer J Clin. 2012; 62:400–22. [PubMed: 22972543]
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. CA Cancer J Clin. 2016; 66:7–30. [PubMed: 26742998]
3. Nutting CM, Morden JP, Harrington KJ, Urbano TG, Bhide SA, Clark C, Milles EA, Miah AB, Newbold K, Tanay M, Adab F, Jefferies SJ, Scrase C, Yap BK, A'Hern RP, Sydenham MA, Emson N, Hall E, PARSPORT trail management group. Parotid-sparing intensity modulated versus conventional radiotherapy in head and neck cancer (PARSPORT): a phase 3 mutlicentre randomized controlled trial. Lancet Oncol. 2011; 12:127–36. [PubMed: 21236730]
4. Rapidis AD, Dijkstra PU, Roodenburg JL, Rodrigo JP, Rinaldo A, Strojan P, Takes RP, Ferlito A. Trismus in patients with head and neck cancer: etiopathogenesis, diagnosis, and management. Clin Otolaryngol. 2015; 40:516–26. [PubMed: 26098612]
5. van der Molen L, Heemsbergen WD, de Jong R, van Rossum MA, Smeele LE, Rasch CR, Hilgers FJ. Dysphagia and trismus after concomitant chemo-intensity-modulated radiation therapy (chemo-

IMRT) in advanced head and neck cancer; dose-effect relationships for swallowing and mastication structures. Radiother Oncol. 2013; 106:364–9. [PubMed: 23540551]

6. Kamstra JI, Dijkstra PU, van Leeuwen M, Roodenburg JL, Langendijk JA. Mouth opening in patients irradiated for head and neck cancer: a prospective repeated measures study. Oral Oncol. 2015; 51:548–55. [PubMed: 25703798]
7. Pauli N, Johnson J, Finizia C, Andréll P. The incidence of trismus and long-term impact on health-related quality of life in patients with head and neck cancer. Acta Oncol. 2013; 52:1137–45. [PubMed: 23193958]
8. Lindblom U, Gärskog O, Kjellén E, Laurell G, Levring Jägahgen E, Wahlberg P, Zackrisson B, Nilsson P. Radiation-induced trismus in the ARTSCAN head and neck trial.
9. Rao SD, Saleh ZH, Setton J, Tam M, McBride SM, Riaz N, Deasy JO, Lee NY. Dose-volume factors correlation with trismus following chemoradiation for head and neck cancer. Acta Oncol. 2016; 55:99–104. [PubMed: 25920361]
10. Pauli N, Olsson C, Pettersson N, Johanson M, Haugen H, Wideräng U, Steineck G, Finizia C. Risk structures for radiation-induced trismus in head and neck cancer. Acta Oncol. 2016; 55:788–92. [PubMed: 26935147]
11. Teguh DN, Levendag PC, Voet P, van der Est H, Noever I, de Kruijf W, van Rooij P, Schmitz PI, Heijmen BJ. Trismus in patients with oropheryngeal cancer: relationship with dose in structures of mastication apparatus. Head Neck. 2008; 30:622–30. [PubMed: 18213726]
12. Bhatia KS, King AD, Paunipagar BK, Abrigo J, Vlantis AC, Leung SF, Ahuja AT. MRI findings in patients with severe trismus following radiotherapy for nasopharyngeal carcinoma. Eur Radiol. 2009; 19:2586–93. [PubMed: 19504110]
13. Hsieh LC, Chen JW, Wang LY, Tsang YM, Sheung PW, Liao LJ, Lo WC, Tseng CF, Kuo YS, Jhuang JY, Tien HJ, Juan HF, Hsieh CG. Predicting the severity and prognosis of trismus after intensity-modulated radiation therapy for oral caner patients using magnetic resonance imaging. PLoS One. 2014; 9:e92561. [PubMed: 24658376]
14. Lambin P, Rios-Velasquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, Zegers CM, Gillies R, Boellard R, Dekker A, Aerts HJ. Radiomics: extracting more information from medical images using advanced features analysis. Eur J Cancer. 2012; 48:441–6. [PubMed: 22257792]
15. Evans ES, Hahn CA, Kocak Z, Zhou SM, Marks LB. The role of functional imaging in the diagnosis and management of late normal tissue injury. Semin Radiat Oncol. 2007; 17:72–80. [PubMed: 17395037]
16. Common Terminology Criteria for Adverse Events (CTCAEv.4.0). http://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03_2010-06-14_QuickReference_5×7.pdf
17. Thirion JP. Image matching as a diffusion process: an analogy with Maxwell's demons. Med Image Anal. 1998; 2:243–60. [PubMed: 9873902]
18. Aerts HJ, Velasquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, Bussink J, Monshouwer R, Haibe-Kains B, Rietveld D, Hoebers F, Rietbergen MM, Leemans CR, Dekker A, Quackenbush J, Gillies RJ, Lambin P. Decoding tumor phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun. 2014; 5:4006. [PubMed: 24892406]
19. El Naqa I, Grigsby P, Apte A, Kidd E, Donnelly E, Khullar D, Chaudari S, Yang D, Schmitt M, Laforest R, Thorstad W, Deasy JO. Exploring feature-based approaches in PET images for predicting cancer treatment outcome. Pattern Recognit. 2009; 42:1162–71. [PubMed: 20161266]
20. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. IEEE Transactions on systems, man, and cybernetics. 1973; 3:610–21.
21. Haralick RM. Statistical and structural approaches to textures. Proceedings of the IEEE. 1979; 67:786–804.
22. Conners RW, Harlow CA. A theoretical comparison of texture algorithms. IEEE Trans Pattern Anal Mach Intell. 1980; 2:204–22. [PubMed: 21868894]
23. Conners RW, Trivedi MM, Harlow CA. Segmentation of a high-resolution urban scene using texture operators. Computer vision, graphics, an image processing. 1984; 25:273–310.
24. Deasy JO, Blanco AI, Clark VH. CERR: a computational environment for radiotherapy research. Med Phys. 2003; 30:979–85. [PubMed: 12773007]

25. Messer JA, Mohamed AS, Hutcheson KA, Ding Y, Lewin JS, Wang J, Lai SY, Frank SJ, Garden AS, Sandulache V, Eichelberger H, French CC, Colen RR, Phan J, Kalpathy-Cramer J, Hazle JD, Rosenthal DI, Gunn GB, Fuller CD. Magnetic resonance imaging of swallowing-related structures in nasopharyngeal carcinoma patients receiving IMRT: Longitudinal dose-response characterization of quantitative signal kinetics. Radiother Oncol. 2016; 118:15–22.

26. Moloney EC, Brunner M, Alexander AJ, Clark J. Quantifying fibrosis in head and neck cancer treatment: An overview. Head Neck. 2015; 37:1225–31. [PubMed: 24797251]

27. Deasy JO, Bentzen SM, Jackson A, Ten Haken RK, Yorke ED, Constine LS, Sharma A, Marks LB. Improving normal tissue complication probability models: The need to adopt a “Data-pooling” culture. Int J Radiat Oncol Biol Phys. 2010; 76:151–4.

28. Hannam AG, McMillan AS. Internal organization in the human jaw muscles. Crit Rev Oral Biol Med. 1994; 5:55–89. [PubMed: 7999950]

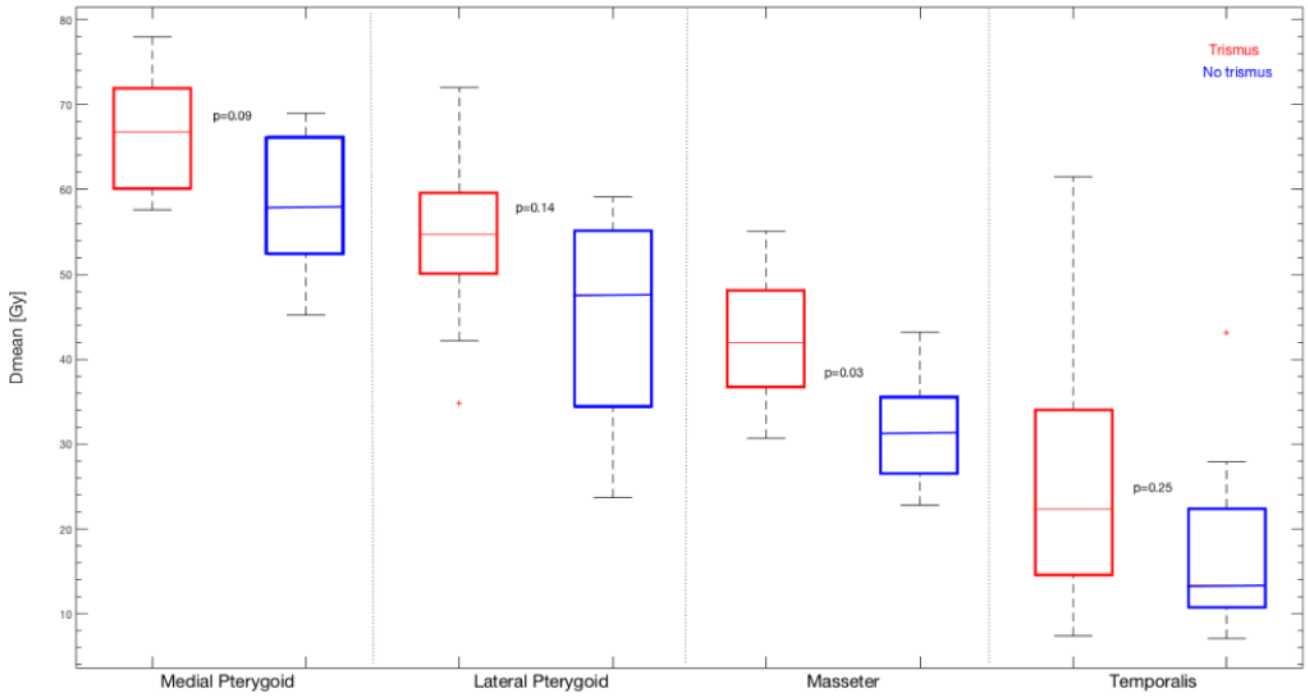



**Figure 1.**

The mean dose (Dmean) for the investigated muscles stratified with respect to trismus status (cases: red; controls: blue). Note: The p-values (averaged across the 1000 Bootstrap samples) have been inserted next to the compared distributions of Dmean.

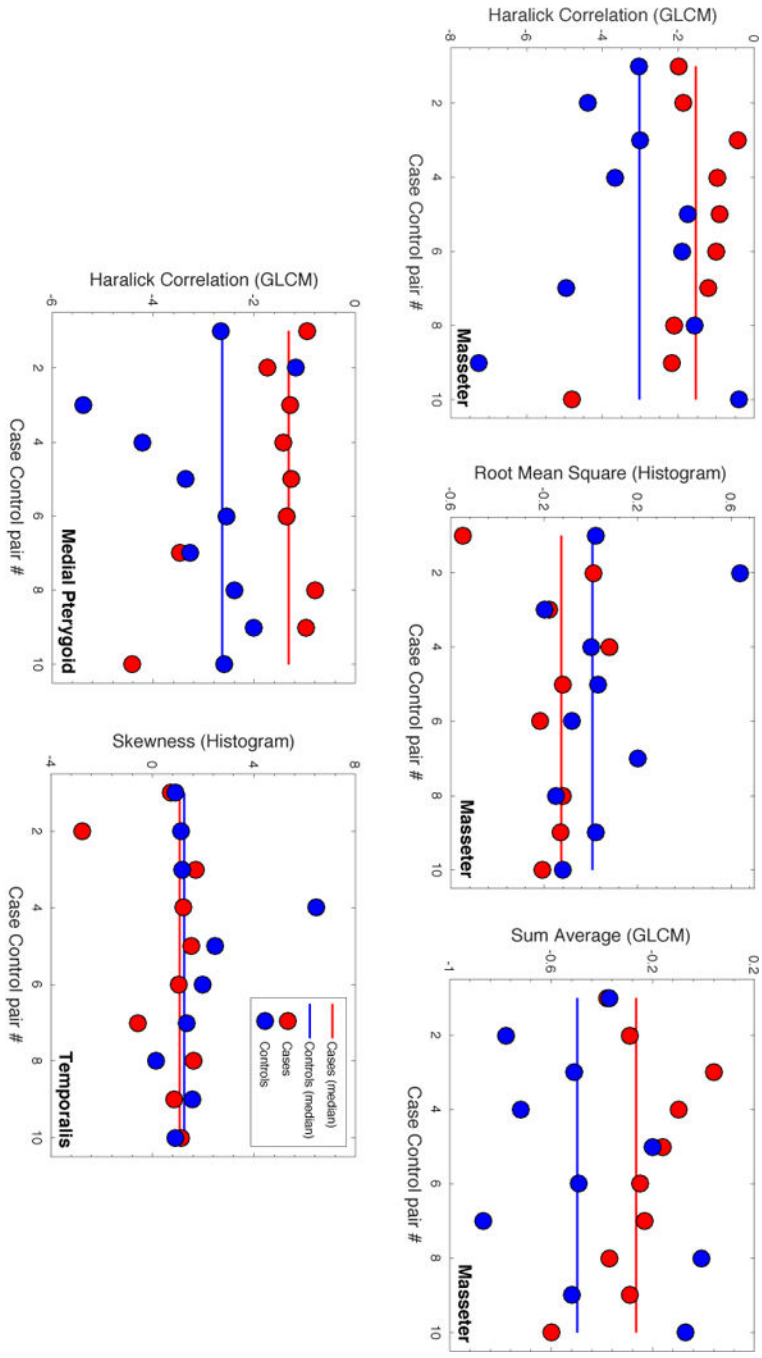



**Figure 2.**

Comparison of the five candidate trismus textures between cases (red dots) and controls (blue dots). Note: Population median values are given as solid lines (cases: red; controls: blue).

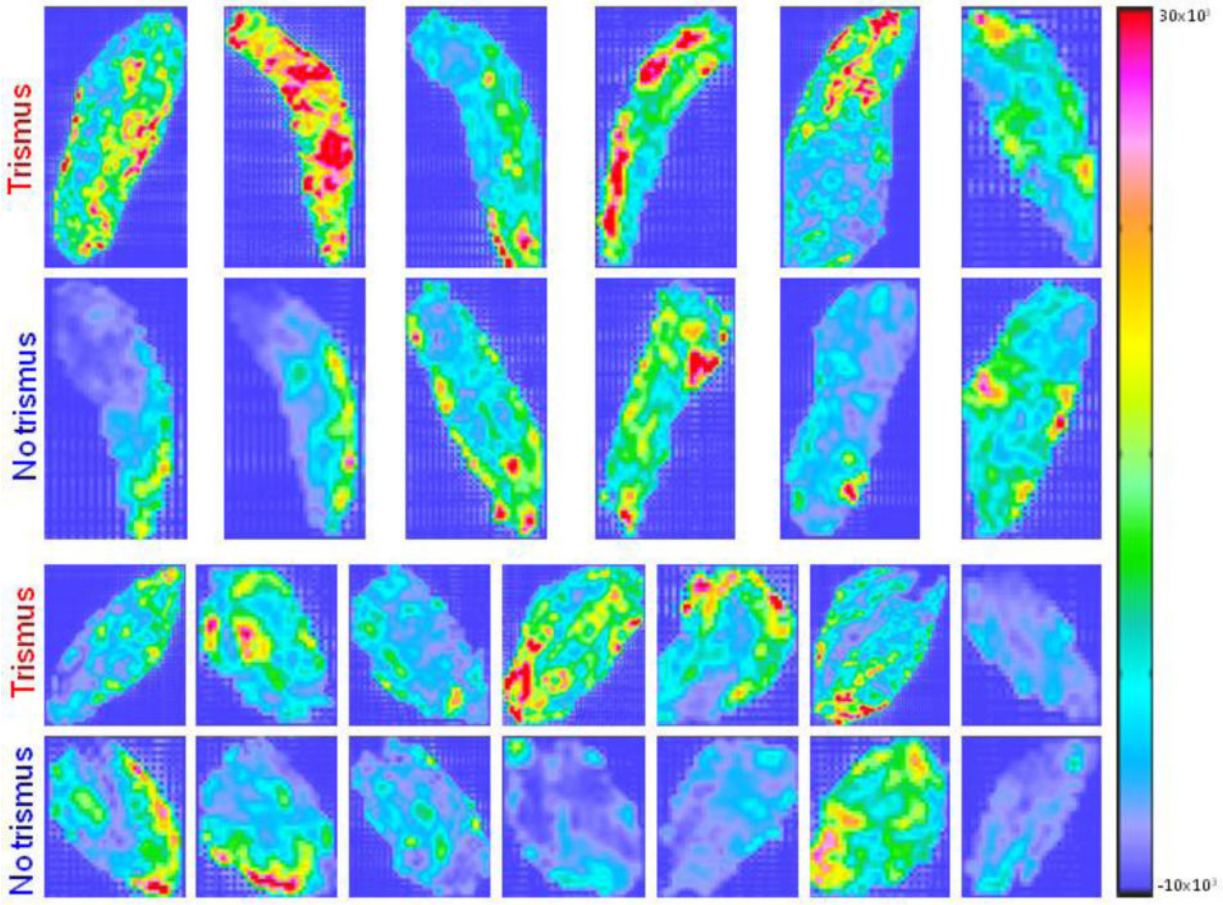



**Figure 3.**

The middle axial slice of the Haralick Correlation (GLCM) texture distribution within the ipsilateral masseter (two upper rows) and the ipsilateral medial pterygoid (two bottom rows) for the cases and controls with a distinct ipsilateral to contralateral separation i.e. $Dmean_{ipsicontra}$>5 Gy. Note: Each matched case-control pair is located in vertical direction.

**Table 1**

Comparison of characteristics between the ten included cases and controls (p-value: Wilcoxon signed rank-test).

| Case control pair | Age [y] | | Gadolinium level [ml] | | Follow-up (m) | | Sex | Time to MRI acqusition (m) | | Tumor site | Tumor stage | | Trismus grade | Weight [kg] | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *Cases/Controls* | *p* | *Cases/Controls* | *p* | *Cases/Controls* | *P* | *Cases+Controls* | *Cases/Controls* | *P* | *Cases+Controls* | *Cases/Controls* | *p* | *Cases* | *Cases/Controls* | *p* |
| 1 | 47/49 | 0.56 | 6/5 | 1.00 | 12/11 | 0.66 | F | 12/11 | 1.00 | Naso | II/IV | 1.00 | 2 | 66/57 | 0.73 |
| 2 | 53/51 | | 6/6 | | 7/6 | | M | 4/6 | | Oro (BOT) | IV/III | | 3 | 64/66 | |
| 3 | 55/58 | | 5/7 | | 3/3 | | M | 3/4 | | Oro (BOT) | IV/III | | 3 | 59/79 | |
| 4 | 51/53 | | 7/7 | | ¾ | | M | 4/3 | | Oro (BOT) | IV/III | | 1 | 75/75 | |
| 5 | 52/55 | | 8/8 | | 3/3 | | M | 3/3 | | Oro (BOT) | III/II | | 1 | 82/84 | |
| 6 | 65/60 | | 7/7 | | 4/3 | | M | 4/3 | | Oro (BOT) | IV/IV | | 1 | 75/79 | |
| 7 | 71/66 | | 6/7 | | 3/3 | | M | 3/3 | | Oro (Tonsil) | III/III | | 1 | 69/79 | |
| 8 | 47/50 | | 8/8 | | 3/3 | | M | 4/3 | | Oro (Tonsil) | IV/IV | | 1 | 86/84 | |
| 9 | 60/61 | | 9/8 | | 8/3 | | M | 5/3 | | Oro (Tonsil) | IV/IV | | 1 | 91/88 | |
| 10 | 54/59 | | 5/4 | | 7/10 | | F | 5/10 | | Oro (Tonsil) | III/IV | | 1 | 59/41 | |
| | **56+8/58+6** | | **7+1/7+1** | | **3(3–12)/3(3–11)** | | | **4(3–12)/3(3–11)** | | | **IV(II-IV)/IV(II-IV)** | | | **73+11/73+15** | |

Note: Numbers in bold (bottom line) denote the population average (± SD) age, gadolinium level, and weight, as well as the population median (range) tumor stage, follow-up time on trismus, and time to MRI acquisition. Abbreviations: BOT=Base of tongue; Naso=Nasopharyngeal; Oro=Oropharyngeal.

**Table 2**

Candidate trismus predictors from univariate logistic regression analyses stratified between cases and controls, the related discriminative ability (AUC) and p-values, as well as the logistic regression coefficients for the intercept (B0), and candidate predictor (B1).

| Muscle | Candidate predictor | Cases | Controls | AUC | P | B0/B1 |
|---|---|---|---|---|---|---|
| Lateral Pterygoid | Dmean [Gy]* | 55 (35, 72) | 48 (24, 59) | 0.76±0.11 | 0.14±0.19 | −4.3/0.09 |
| Masseter | Dmean [Gy] | 42 (31, 55) | 31 (23, 43) | 0.85±0.09 | 0.03±0.07 | −8.2/0.22 |
| | Haralick Correlation *GLCM* | −1.5 (−4.8, −0.4) | −3.0 (−7.3, −0.4) | 0.72 (0.12) | 0.14 (0.20) | 1.5/0.63 |
| | Root Mean Square *Histogram* | −0.1 (−0.6, 0.2) | 0.0 (−0.2, 0.6) | 0.71 (0.11) | 0.19 (0.23) | 0.22/−4.3 |
| | Sum Average *GLCM*** | −0.3 (−0.9, 0.0) | −0.5 (−0.8, 0.0) | 0.70 (0.12) | 0.19 (0.24) | 1.3/3.5 |
| Medial Pterygoid | Dmean [Gy] | 67 (58, 78) | 58 (45, 69) | 0.77±0.11 | 0.09±0.14 | −10/0.16 |
| | Haralick Correlation *GLCM* | −1.3 (−4.4, −0.8) | −2.6 (−5.4, −1.2) | 0.78 (0.12) | 0.12 (0.20) | 2.1/0.92 |
| Temporalis | Skewness *Histogram* | 1.1 (−2.8, 1.7) | 1.3 (0.2, 6.5) | 0.68 (0.12) | 0.17 (0.21) | 0.96/−0.82 |

Note: Dmean and textures are given as the median (range), and AUC and p-values as the mean±SD across all Bootstrap samples, while B0 and B1 are the regression coefficients without Bootstrapping applied. Not considered a final candidate predictor due to being highly correlated with *Dmean of the masseter, and the medial pterygoid (|Rs|=0.78, 0.79), and ** the Haralick Correlation GLCM of the masseter (|Rs|=0.91).