Published in final edited form as:

Int J Radiat Oncol Biol Phys. 2018 November 15; 102(4): 1330–1338. doi:10.1016/j.ijrobp.2018.05.054.

Image-based data mining to probe dosimetric correlates of radiation-induced trismus

W Beasley^{1,*}, M Thor², A McWilliam³, A Green⁴, R Mackay³, N Slevin⁵, C Olsson⁶, N Pettersson⁶, C Finizia⁷, C Estilo⁸, NY Lee⁹, JO Deasy², and M van Herk⁴

¹Division of Cancer Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, United Kingdom; Christie Medical Physics and Engineering, Christie NHS Foundation Trust, Manchester, United Kingdom

²Department of Medical Physics, Memorial Sloan Kettering Cancer Center, New York, New York

³Division of Cancer Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, United Kingdom; Christie Medical Physics and Engineering, Christie NHS Foundation Trust, Manchester, United Kingdom

⁴Division of Cancer Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, United Kingdom

⁵Department of Clinical Oncology, Christie NHS Foundation Trust, Manchester, United Kingdom

⁶Department of Radiation Physics, Institute of Clinical Sciences, Sahlgrenska Academy at the University of Gothenburg, Gothenburg, Sweden

⁷Department of Otorhinolaryngology, Institute of Clinical Sciences, Sahlgrenska Academy at the University of Gothenburg, Gothenburg, Sweden

⁸Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, New York

⁹Department of Radiation Oncology, Memorial Sloan Kettering Cancer Center, New York, New York

Abstract

Purpose: To identify imaged regions in which dose is associated with radiation-induced trismus after head and neck cancer radiotherapy (HNRT) using a novel image-based data mining (IBDM) framework.

Materials and methods: A cohort of 86 HNRT patients were analysed for region identification. Trismus was characterised as a continuous variable by the maximum incisor-to-incisor opening distance (MID) at 6 months post-RT. Patient anatomies and dose distributions were spatially normalised to a common frame of reference using deformable image registration. IBDM was used to identify clusters of voxels associated with MID ($p < 0.05$ based on permutation testing). The result was externally tested on a cohort of 35 HNC patients. Internally, we also performed a dose-volume histogram (DVH)-based analysis by comparing the magnitude of the correlation between

* william.beasley@christie.nhs.uk.

MID and the mean dose (D_{mean}) for the IBDM-identified cluster in comparison with five delineated masticatory structures.

Results: A single cluster was identified with the IBDM approach ($p < 0.01$), partially overlapping with the ipsilateral masseter. The DVH-based analysis confirmed that the IBDM cluster had the strongest association with MID, followed by the ipsilateral masseter and the ipsilateral medial pterygoid (Spearman's rank correlation coefficients: $R_s = -0.36, -0.35, -0.32$; $p = 0.001, 0.001, 0.002$). External validation confirmed an association between D_{mean} to the IBDM cluster and MID ($R_s = -0.45$; $p = 0.007$).

Conclusions: IBDM bypasses the common assumption that dose patterns within structures are unimportant. Our novel IBDM approach for continuous outcome variables successfully identified a cluster of voxels that are highly associated with trismus, overlapping partially with the ipsilateral masseter. Tests on an external validation cohort showed an even stronger correlation with trismus. These results support use of the region in HNRT treatment planning to potentially reduce trismus.

1. Introduction

The aim of radiotherapy (RT) is to deliver a high dose to the tumour whilst minimising the dose to nearby sensitive structures. This balance is optimised using assumed dose-response relationships describing the probability of response and toxicity. Dose-response relationships are typically identified from dose-volume histogram (DVH)-based analyses. However, such techniques collapse 3D dose distributions into a single or few parameters, and do not account for potential heterogeneities in radiosensitivity within organs. In addition, DVH-based analyses require prior hypotheses regarding the structures responsible for toxicity, and subsequent delineation of these structures.

Image-based data mining (IBDM) is an alternative technique for identifying dose-response relationships. Instead of assessing dose-response relationships from pre-hypothesised critical structures, IBDM operates on each voxel of the dose distribution, providing a spatial map of dose-response relationships¹. Recently, IBDM has been applied to identify regions related to treatment failure in prostate cancer RT¹, as well as for various toxicities after RT for prostate cancer²⁻⁷, Hodgkin's lymphoma⁸, head and neck cancer (HNC)⁹, and overall survival in lung cancer¹⁰. All of these studies have investigated binary representations of the outcome (*e.g.* recurrence or not, dead or alive). However, many clinical outcomes are described by continuous variables, and should ideally be modelled as such to maintain the highest level of detail.

Difficulty in mouth opening, *i.e.* trismus, is a common radiation-induced toxicity that can be characterised by the maximum incisor-to-incisor distance (MID). Trismus is observed in between 5% and 45% of all HNC patients treated with RT alone and up to 50% for patients treated with RT and surgery^{11,12}. Risk factors include receiving surgery, advanced age, tumour location and dose to the masticatory structures¹³⁻¹⁶. The masticatory muscles include the masseter, lateral and medial pterygoid, temporomandibular joint and temporalis, and dose-response relationships have been reported for the bilateral lateral pterygoid¹³, ipsilateral masseter¹⁶⁻¹⁸ and ipsilateral medial pterygoid¹⁶. The aim of the present study was to introduce a novel IBDM approach for continuous outcome variables and use it to identify

an anatomic cluster of voxels involved in radiation-induced trismus, characterised by MID. This methodology was applied to a HNC cohort and validated internally using a DVH-based approach, as well as externally in an independent patient cohort.

2. Method

2.1. Patient data

One hundred and nineteen patients treated with RT for HNC at XXX in an ethically approved prospective study¹⁸ were available for analysis. Patients received either 3D conformal RT or fixed-gantry intensity modulated radiotherapy (IMRT), delivering 65–73Gy to the primary tumour over 33–38 fractions using either a simultaneous integrated boost or a two-phase approach. Patients were immobilised with a head mask for imaging and treatment, and treatment plans were created with the Eclipse treatment planning system (Varian Medical Systems, Palo Alto, CA, USA). CT slice thicknesses were either 3mm or 5mm. To account for fractionation effects, dose distributions were converted into biologically effective doses (BEDs) using an α/β ratio of 3Gy¹⁹. The BED at voxel i is given by $BED_i = nd(1 + d_i/(\alpha/\beta))$, where n is the number of fractions and d_i is the dose per fraction at voxel i . For patients treated in two phases, the BED distributions of individual phases were summed. For each patient the MID at baseline and at six months post-RT were available. Patients were not included in the analysis if they received an intervention before the six month MID measurement (23 patients), or if they had an MID <35mm at baseline, as this could indicate tumour-related trismus (5 patients). In addition, dicom dose and/or images were not available for 5 further patients. A total of 86 patients were therefore available for analysis (Table 1). The masseter, medial pterygoid, lateral pterygoid, temporalis and temporomandibular joint (TMJ) had been previously systematically delineated for each patient according to written instructions and anatomical extensions¹⁸.

2.2. Image-based data mining

2.2.1. Image pre-processing—Prior to IBDM (Figure 1), various pre-processing steps were performed: The ipsilateral side of each patient was determined by measuring the mean dose (D_{mean}) to all outlined mastication structures and calculating the difference between the left and right-sided structures. If the difference in D_{mean} was >5Gy, the side with the higher D_{mean} was classified as ipsilateral; otherwise patients were classified as bilateral. To ensure that all ipsilateral structures were on the same side, the CTs and BED-corrected doses for patients with the ipsilateral side on the right hand side were mirrored through the sagittal plane by inverting the lateral coordinates. For bilateral patients, both sides were considered ipsilateral and so the sagittal reflection was also performed; these patients were therefore included twice, both with and without this sagittal reflection.

2.2.2. Deformable image registration—After the initial pre-processing steps, air cavities were masked from the CT images by setting the minimum value for voxels within the patient to 0 HU. Patients were then spatially normalised into a common frame of reference, which was chosen manually by visual inspection, using rigid registration within an automatically generated clip box covering the muscles of mastication and the mandible, followed by deformable image registration (DIR). The DIR used a fast free-form

deformation method based on cubic B-splines²⁰, with the images being resampled to the voxel dimensions of the reference patient. DIR was implemented in NiftyReg, which is an open-source DIR platform (<http://cmictig.cs.ucl.ac.uk/wiki/index.php/NiftyReg>).

The DIR uncertainty was assessed by calculating the standard deviation (SD) in the centre of mass (COM_{SD}) of the delineated muscles of mastication. The mean COM_{SD} across all muscles was calculated in the superior-inferior (SI), anterior-posterior (AP) and left-right (LR) directions. To account for DIR uncertainty, these COM_{SD} were used as the SD for a 3D Gaussian kernel to blur the dose distributions of each patient.

2.2.3. Voxel-wise correlation map—A voxel-wise correlation map was created from the BED distributions of the spatially normalised patients using an in-house toolkit. As illustrated in Figure 2, the correlation map was calculated from measuring the Spearman's rank correlation coefficient (R_s) between the dose and MID at each voxel. The R_s map was masked from approximately 3cm below the mandible to the base of the skull (see Figure 3(a)) to limit noise in superior (brain) regions and relatively large registration uncertainties in the inferior (neck) regions.

Correction for multiple comparisons posed by the large number of voxels was performed using permutation testing with the maximum suprathreshold cluster size ($STCS_{max}$) as the test statistic²¹. This method has not previously been applied to IBDM in RT for continuous outcome variables, and is therefore briefly described here. The R_s map was first dichotomised based on a manually-defined threshold (75% of the maximum R_s within the image), and the volume of the largest correlation cluster in the image was recorded (this is $STCS_{max}$). The MID values were permuted 1000 times and the R_s map recalculated each time; $STCS_{max}$ was recorded for each permutation and combined into a permutation histogram. Clusters in the original R_s map were considered statistically significant if they had a volume greater than the 95th percentile of the permutation histogram (for $\alpha=0.05$). Permutation testing with $STCS_{max}$ therefore enables spatial localisation of clusters exhibiting a statistically significant dose-response relationship. Since the test was performed on a single variable, it accounts for multiple testing.

2.3. Validation

2.3.1. Internal—In addition to our IBDM approach, a DVH-based analysis was performed within the training cohort for the previously segmented masticatory structures, as well as for the IBDM-identified cluster. The R_s between D_{mean} to each of these structures and MID was calculated, and statistical significance ($p < 0.05$) was assessed from a permutation test.

The influence of previously reported risk factors for trismus was also investigated in a multivariable analysis (Table 4(supplementary)): age, baseline MID (continuous), smoking (current/former vs never-smoker), tumour site (categorical), as well as chemotherapy (no vs concurrent), gender, and tumour laterality (Left/right/bilateral). First, each potential risk factor, F , was investigated using a univariate linear regression analysis ($MID = \beta_0 + \beta_1 F$). If presenting with $p < 0.05$, that candidate risk factor was included in a multivariable linear regression analysis ($MID = \beta_0 + \beta_1 F_1 + \beta_2 F_2 + \dots + \beta_n F_n$). Risk factors to be included in the final

model were chosen using forward selection. For this, the baseline model used D_{mean} to the IBDM-identified cluster, and each candidate risk factor was added individually. The difference between the model with and without the candidate risk factor was assessed using a likelihood ratio test, and the candidate predictor was considered a final risk factor if this difference was statistically significant ($p < 0.05$). All statistical analyses were performed in R 3.2.4²².

2.3.2. External—An external validation was also performed within a cohort of 35 HNC patients treated in an IRB-approved study at XXX (Table 1). Patients had been treated using fixed-gantry IMRT to 50–66Gy in 30–33 fractions using either a simultaneous integrated boost or a two-phase approach. Similar to the training cohort, patients had been immobilised with a head mask both for imaging and treatment, and treatment plans were generated with the Eclipse treatment planning system (Varian Medical Systems, Palo Alto, CA, USA). CT slice thicknesses were either 2.5mm or 3mm. For this cohort, MID at baseline and at a median of five (range: 3–8) months post-RT were available. Dose distributions were converted into BED as previously described, and CT and BED images were mirrored in the sagittal axis according to the dose distribution relative to the masticatory structures. Each patient was spatially normalised to the reference patient, and D_{mean} to the voxel cluster identified from IBDM in the training cohort was recorded.

External validation for D_{mean} to the IBDM-identified cluster was performed using the linear regression coefficients obtained from within the training cohort to predict MID in the external validation cohort. The performance of the model was assessed using R_s between the predicted and actual MID. The multivariable model, including the risk factors identified in the training cohort, was assessed using the same approach.

3. Results

3.1. DIR accuracy

The COM_{SD} , averaged over all muscles of mastication, was 4.0, 0.34 and 0.30 mm in the LR, AP and SI directions, respectively. These COM_{SD} were applied to the Gaussian kernel used to blur the BED distributions of each patient prior to IBDM.

3.2. Image-based data mining

The left- and right-hand side was deemed the ipsilateral side for 42% and 24% of the patients, respectively; the images were reflected through the sagittal axis for the right-sided patients. The remaining 33% of the patients had bilateral disease; these patients were included both with and without mirroring the images through the sagittal plane. This resulted in a total of 115 images on which IBDM was performed.

The voxel-based R_s map (Figure 3(a)) showed large regions of negative correlations, suggesting that a higher dose was associated with lower MID (minimum R_s in image = -0.41). The resulting IBDM cluster, identified with a threshold of -0.30 for $STCS_{\text{max}}$, is illustrated in Figure 3(b). This cluster partially overlapped with the ipsilateral masseter, demonstrating a possible dose-response relationship for this structure.

3.3. Validation

3.3.1. Internal—The internal DVH-based validation indicated a statistically significant and negative correlation between D_{mean} to the IBDM cluster and MID ($R_s = -0.36$; $p = 0.001$; Table 2). In addition, D_{mean} to the ipsilateral masseter correlated with MID ($R_s = -0.35$; $p = 0.001$). This indicates that the IBDM cluster, which overlapped with the ipsilateral masseter, correctly identified this structure. The D_{mean} to the ipsilateral medial pterygoid and ipsilateral lateral pterygoid were also correlated with MID, although the correlation was weaker for these structures.

Amongst the investigated potential risk factors, D_{mean} to the IBDM-identified cluster, baseline MID, and smoking status were univariate candidate predictors ($R_s = 0.37, 0.52, 0.29$; $p < 0.001, < 0.001, 0.013$). Systematic forward variable selection using these predictors resulted in all of these variables being included in the final multivariable model (Table 4 (supplementary)). The linear regression equation for the final multivariable model was: $\text{MID}(\text{mm}) = 21.31 - 0.31D_{\text{meanIBDM}} + 0.51\text{MID}_{\text{baseline}} + 4.19\text{Smoking}$ ($R_s = 0.60$).

3.3.2. External—Despite some differences between the internal and external validation cohorts (Table 1), the correlation between D_{mean} for the IBDM cluster and MID in the validation cohort was of similar magnitude as in the training cohort ($R_s = -0.45$; $p = 0.008$; Table 3), confirming that the dose to the identified IBDM cluster was also significantly associated with MID in these patients.

The univariate correlation between MID in the external cohort and MID predicted from D_{mean} to the IBDM-identified cluster using the training linear regression coefficients was of similar magnitude as in the training cohort ($R_s = 0.45$, $p = 0.007$; Table 3). The R_s for the multivariable model on the external cohort was $R_s = 0.76$ ($p < 0.001$) (Table 3). A calibration plot for the multivariable model is provided in supplementary Figure 4; the coefficient of determination was 0.39 and the root mean squared error was 7.1 mm.

4. Discussion

This work has demonstrated a novel framework with which to perform image-based data mining (IBDM) with continuous outcome data, which was applied to radiation-induced trismus following HNC RT. To the best of our knowledge, this is the first demonstration of IBDM for continuous outcome data. Our approach successfully identified an anatomic cluster of voxels demonstrating an association with trismus at six months post-RT, and both internal and external validation supported the importance of this region. Furthermore, the identified cluster overlapped with the ipsilateral masseter, emphasising the importance of this structure for the development of radiation-induced trismus.

Several different techniques have been suggested for improving upon traditional DVH-based dose-response modelling so as to incorporate spatial information. These include calculating DVH parameters for a sub-volume of an organ²³, quantifying the distance between high dose regions²⁴, and parameterising the shape of 2D rectum dose-surface maps^{25–27}. Whilst these techniques improve upon simplistic whole-organ DVH models, they are used for hollow cylindrical structures such as the rectum, and are not suited to 3D localisation of

dose-response relationships. Incorporation of 3D dose information within dose-response models can be performed with 3D moment invariants. This has been used for the parotid glands (xerostomia)²⁸, oral mucosa (mucositis)^{29,30} and pharyngeal mucosa (dysphagia)³¹. This technique parameterises the centre of mass, skewness and spread of the dose distribution within an organ, which enables toxicities in both parallel and serial structures to be investigated. By including these parameters within the modelling, 3D moment invariants account for regional variations in dose within a structure. However, 3D moment invariants require prior delineations of structures with a pre-hypothesised dose-response relationship; IBDM is an alternative technique that is not biased by prior segmentations.

IBDM is an emerging technique for spatially localising dose-response relationships in RT. In contrast to DVH-based analyses, in which prior segmentation of structures with potential dose-response relationships is required, no assumptions are made on the size and shape of critical structures. This means that identified dose-response relationships are not limited to the delineated structures, and can be found anywhere in which the local dose is correlated to the outcome. IBDM is hypothesis-generating and has been applied to several tumour sites including the prostate, head and neck and lung^{3–10,32}. It has recently been used successfully by Monti *et al.* (2017) to identify the cricopharyngeus muscle and cervical oesophagus as being related to acute dysphagia in HNC RT. Similarly, McWilliam *et al.* (2017) used IBDM to identify the dose to the base of the heart as an important structure for overall survival in lung RT.

Trismus is a common side-effect following RT for HNC³³, and previous DVH-based studies have reported dose-response relationships with the mean dose for primarily the ipsilateral masseter and to a lesser extent the ipsilateral medial pterygoid¹⁶. With our IBDM analysis we identified a single cluster of voxels that overlapped with the ipsilateral masseter, and internal DVH-based validation confirmed a dose-response relationship both with this cluster and the ipsilateral masseter. The doses to neighbouring structures are correlated with each other and so it is difficult to determine whether the dose to a single structure is more important, or whether the interplay between doses is important. However, the correlation between D_{mean} and MID was confirmed for the IBDM-identified cluster on external validation.

Furthermore, within multivariable linear regression analysis on the training cohort, D_{mean} to the IBDM-identified cluster was found to be predictive of MID, along with baseline MID and smoking status. Reduced mouth opening before treatment was more likely to result in a smaller MID six months post-RT, and never-smokers were more likely to have a larger MID at six months post-RT than current or former smokers. This model, fitted on the training cohort, was then shown to predict MID in the validation cohort. Previous studies have identified baseline MID and smoking status as predictors for trismus¹⁸, in agreement with our results. Other previously-reported risk factors such as treatment site, age and chemotherapy were not found to be related to MID in our analysis. Surgery, which has also been identified as a risk factor for trismus³⁴, was not evaluable as none of the patients in the training cohort had undergone surgery. Nevertheless, D_{mean} to the IBDM-identified cluster was found to be a predictor for trismus in both patient cohorts even after correcting for several confounding clinical factors.

The multivariable model was developed using forward variable selection, which in some circumstances can lead to overfitting. However, overfitting was unlikely in this case due to the small number of variables used in the model. Nevertheless, model stability was investigated, as recommended by Heinze et al. (2018)³⁵ using the R package *mplog*³⁶. For models with three variables, our model was identified as being the most stable (see Table 5 (supplementary) and Figure 6 (supplementary)). Furthermore, the final model was tested in an independent external validation cohort, which is one of the crucial steps of exploring model performance, as further explained in the TRIPOD statement³⁷.

There were differences between the training and validation cohort (Table 1), and the possibility of a selection bias cannot be ruled out. Differences included the baseline MID (two patients had MID<35mm in the validation cohort), treatment site, technique, surgery, chemotherapy and designated laterality of disease. These differences were likely caused by the relatively small size of the validation cohort, and the different populations from which patients originated. Although these differences pose a risk of selection bias, the model was able to predict MID this independent cohort.

In the present study trismus was characterised by MID, which is an objective measure of mouth opening. Other subjective measures of trismus are also commonly performed, including the Common Terminology Criteria for Adverse Events (CTCAE), which grades trismus according to the severity of impaired mouth opening. However, MID provides an objective measure of trismus and is well suited to the IBDM methodology introduced here. Furthermore, MID has been shown to correlate with a patient's perceived experience of trismus¹⁷.

One of the challenges with IBDM is the multiple comparison problem due to the considerable number of voxels within dose distributions being investigated. To overcome this issue, permutation testing has been successfully employed^{21,32}. For this we selected the size of the largest cluster in the image after dichotomising the R_s map with a pre-defined threshold ($STCS_{max}$) as the test statistic to identify statistically significant voxels. The exact choice of threshold or the use of threshold-free techniques⁸ is likely to have a negligible influence on our results, as although the threshold choice could affect the size of the cluster, it would not affect its position. Therefore, the ipsilateral masseter would still have been identified as the critical structure. We acknowledge that our identified structure covered only part of the ipsilateral masseter, but a further radiosensitivity analysis within this structure was beyond the scope of this study. The exact location of the identified region is dependent on the variation of dose throughout the image (see supplementary figures 7 and 8), although this was mitigated by the variety of treatment sites and modalities in the patient cohort.

Another challenge with any IBDM approach is accurate spatial normalisation using DIR. This is important to ensure that the dose at each voxel represents the same anatomic point in each patient. In the present study we addressed this potential challenge thoroughly by assessing the uncertainty in the DIR by measuring the SD in the COM of the masticatory structures, and by accounting for this uncertainty by blurring the dose distributions according to the measured SD.

Prior to IBDM, the dose distributions were flipped according to the side of the patient receiving the highest dose. This is analogous to DVH-based analyses that consider ipsilateral and contralateral structures separately, and was performed here due to recent literature suggesting that the ipsilateral structures are more important for radiation-induced trismus^{16,18}. Indeed, IBDM analysis without flipping (results not presented) identified a similar but smaller region on the ipsilateral (left side for the majority of patients) side of the patient. Furthermore, patients with bilateral dose distributions were included twice, with and without flipping the dose distributions. This ensured that no assumptions were made about the most important side of the patient.

The results of this study have indicated that the ipsilateral masseter has a dose-response relationship with radiation-induced trismus. Remarkably, prediction on the external validation dataset was even stronger than on the training dataset. This is therefore evidence that reducing the dose to this region during RT is likely to decrease the severity and rate of trismus. Despite this, uncertainty remains as to the precise dose thresholds that should be used for this anatomic region, as well as the extent to which this structure could be spared during RT.

5. Conclusion

A better understanding of the relationship between trismus and the spatial dose distributions could be used to guide treatment planning on an individual basis. This is usually done by assuming that dose distribution details with a segmented structure are unimportant. To the best of our knowledge, this study is the first clinical application of IBDM with a continuous outcome variable. The IBDM approach was applied to radiation-induced trismus in HNC RT, and identified a cluster of voxels, only partially overlapping with the ipsilateral masseter, that was associated with trismus. The cluster was validated internally using a DVH-based approach, as well as externally in another patient cohort. Multivariable linear regression in both cohorts confirmed that the dose to the IBDM-identified cluster was predictive of trismus, along with baseline MID and smoking status. These results demonstrate that IBDM with continuous outcome variables enables outcome associations to be spatially localised, thereby improving our knowledge of toxicity etiology in radiotherapy.²²

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Witte MG et al. Relating dose outside the prostate with freedom from failure in the Dutch trial 68 Gy vs. 78 Gy. *Int. J. Radiat. Oncol. Biol. Phys* 77, 131–8 (2010). [PubMed: 20004526]
2. Heemsbergen WD et al. Urinary obstruction in prostate cancer patients from the dutch trial (68 Gy vs. 78 Gy): relationships with local dose, acute effects, and baseline characteristics. *Int. J. Radiat. Oncol. Biol. Phys* 78, 19–25 (2010). [PubMed: 20056354]
3. Acosta O et al. Voxel-based population analysis for correlating local dose and rectal toxicity in prostate cancer radiotherapy. *Phys. Med. Biol* 58, 2581–95 (2013). [PubMed: 23528429]

4. Wortel RC et al. Prostate cancer radiotherapy Dose-surface maps identifying local dose-effects for acute gastrointestinal toxicity after radiotherapy for prostate cancer. *Radiother. Oncol* 117, 515–20 (2015). [PubMed: 26522060]
5. Improta I et al. Bladder spatial-dose descriptors correlate with acute urinary toxicity after radiation therapy for prostate cancer. *Phys. Medica* 32, 1681–1689 (2016).
6. Palorini F et al. First application of a pixel-wise analysis on bladder dose-surface maps in prostate cancer radiotherapy. *Radiother. Oncol* 119, 123–8 (2016). [PubMed: 26993415]
7. Yahya N et al. Modeling Urinary Dysfunction After External Beam Radiation Therapy of the Prostate Using Bladder Dose-Surface Maps: Evidence of Spatially Variable Response of the Bladder Surface. *Int. J. Radiat. Oncol. Biol. Phys* 97, 420–426 (2017). [PubMed: 28068247]
8. Palma G et al. A voxel-based approach to explore local dose differences associated with radiation-induced lung damage. *Int. J. Radiat. Oncol. Biol. Phys* 96, 127–33 (2016). [PubMed: 27511851]
9. Monti S et al. Voxel-based analysis unveils regional dose differences associated with radiation-induced morbidity in head and neck cancer patients. *Sci. Rep* 7, 7220 (2017). [PubMed: 28775281]
10. McWilliam A et al. Radiation dose to heart base linked with poorer survival in lung cancer patients. *Eur. J. Cancer* 85, 106–113 (2017). [PubMed: 28898766]
11. Louise Kent M et al. Radiation-Induced trismus in head and neck cancer patients. *Support. Care Cancer* 16, 305–309 (2008). [PubMed: 17965892]
12. Wetzels J-WGH, Merkx MAW, de Haan AFJ, Koole R & Speksnijder CM Maximum mouth opening and trismus in 143 patients treated for oral cancer: A 1-year prospective study. *Head Neck* 36, 1754–1762 (2014). [PubMed: 24478217]
13. Hsieh L-C et al. Predicting the severity and prognosis of trismus after intensity-modulated radiation therapy for oral cancer patients by magnetic resonance imaging. *PLoS One* 9, e92561 (2014). [PubMed: 24658376]
14. Lindblom U et al. Radiation-induced trismus in the ARTSCAN head and neck trial. *Acta Oncol. (Madr)* 53, 620–7 (2014).
15. Pauli N, Fagerberg-Mohlin B, Andréll P & Finizia C Exercise intervention for the treatment of trismus in head and neck cancer. *Acta Oncol. (Madr)* 53, 502–509 (2014).
16. Rao SD et al. Dose-volume factors correlating with trismus following chemoradiation for head and neck cancer. *Acta Oncol. (Madr)* 55, 99–104 (2016).
17. van der Molen L et al. 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* 106, 364–9 (2013). [PubMed: 23540551]
18. Pauli N et al. Risk structures for radiation-induced trismus in head and neck cancer. *Acta Oncol. (Madr)* 55, 788–792 (2016).
19. Kehwar TS Analytical approach to estimate normal tissue complication probability using best fit of normal tissue tolerance doses into the NTCP equation of the linear quadratic model. *J. Cancer Res. Ther* 1, 168–79 (2005). [PubMed: 17998649]
20. Modat M et al. Fast free-form deformation using graphics processing units. *Comput. Methods Programs Biomed* 98, 278–84 (2010). [PubMed: 19818524]
21. Nichols TE & Holmes AP Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum. Brain Mapp* 15, 1–25 (2002). [PubMed: 11747097]
22. R Core Team. R: A Language and Environment for Statistical Computing R Foundation for Statistical Computing (2016).
23. Dréan G et al. Identification of a rectal subregion highly predictive of rectal bleeding in prostate cancer IMRTRectal subregion involved in bleeding. *Radiother. Oncol* 119, 388–397 (2016). [PubMed: 27173457]
24. Munbodh R & Jackson A Quantifying cell migration distance as a contributing factor to the development of rectal toxicity after prostate radiotherapy. *Med. Phys* 41, (2014).
25. Buettner F et al. Assessing correlations between the spatial distribution of dose to the rectal wall and late rectal toxicity after prostate radiotherapy. *IFMBE Proc.* 25, 908–911 (2009).

26. Buettner F, Gulliford SL, Webb S & Partridge M Modeling late rectal toxicities based on a parameterized representation of the 3D dose distribution. *Phys. Med. Biol* 56, 2103–2118 (2011). [PubMed: 21386140]
27. Moulton CR et al. Spatial features of dose-surface maps from deformably-registered plans correlate with late gastrointestinal complications. *Phys. Med. Biol* 62, 4118–4139 (2017). [PubMed: 28445167]
28. Buettner F et al. Novel approaches to improve the therapeutic index of head and neck radiotherapy: An analysis of data from the PARSPORT randomised phase III trial. *Radiother. Oncol* 103, 82–87 (2012). [PubMed: 22444242]
29. Dean JA et al. Normal tissue complication probability (NTCP) modelling using spatial dose metrics and machine learning methods for severe acute oral mucositis resulting from head and neck radiotherapy. *Radiother. Oncol* 120, 21–27 (2016). [PubMed: 27240717]
30. Dean JA et al. Normal Tissue Complication Probability (NTCP) Modelling of Severe Acute Mucositis using a Novel Oral Mucosal Surface Organ at Risk. *Clin. Oncol* 29, 263–273 (2017).
31. Dean J et al. Incorporating spatial dose metrics in machine learning-based normal tissue complication probability (NTCP) models of severe acute dysphagia resulting from head and neck radiotherapy. *Clin. Transl. Radiat. Oncol* 8, 27–39 (2018). [PubMed: 29399642]
32. Chen C, Witte M, Heemsbergen W & van Herk M Multiple comparisons permutation test for image based data mining in radiotherapy. *Radiat. Oncol* 8, 293 (2013). [PubMed: 24365155]
33. Dijkstra PU, Kalk WWI & Roodenburg JLN Trismus in head and neck oncology: a systematic review. *Oral Oncol.* 40, 879–89 (2004). [PubMed: 15380165]
34. Agarwal P, Shiva Kumar HR & Rai KK Trismus in oral cancer patients undergoing surgery and radiotherapy. *J. Oral Biol. Craniofacial Res* 6, S9–S13 (2016).
35. Heinze G, Wallisch C & Dunkler D Variable selection - A review and recommendations for the practicing statistician. *Biometrical J* 1–19 (2018). doi:10.1002/bimj.201700067
36. Tarr G, Müller S & Welsh AH mplot: An R Package for Graphical Model Stability and Variable Selection Procedures. *J. Stat. Softw* 83, (2018).
37. Collins GS, Reitsma JB, Altman DG & Moons KGM Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD): The TRIPOD Statement. *Eur. Urol* 67, 1142–1151 (2015). [PubMed: 25572824]

Figure 1: Methodology for image-based data mining. Patients are spatially normalised into a common frame of reference using deformable image registration (DIR), and the resulting transformation is applied to the dose distributions. A voxel-wise correlation map is then created to spatially localise associations with an outcome variable.

Figure 2:

An R_s map was produced by correlating the dose at each voxel in the spatially normalised BED distributions with the outcome data. This allows the location of any dose-response relationships to be determined.

Figure 3:

(a) Correlation maps between mouth opening distance and voxel dose, for a representative axial (top) and sagittal (bottom) slice, with the colour representing the Spearman rank correlation coefficient (R_s). The ipsilateral side is on the right hand side of the images as viewed, and due to the method of data analysis the correlations on the left hand side of the images are not valid. (b) The cluster identified as exhibiting a dose-response relationship with MID (red shaded region) shown on a representative axial (top) and sagittal (bottom) slice. The masseter is shown in red, the medial pterygoid in blue, the lateral pterygoid in yellow, and the temporalis in magenta.

Table 1:

Patient characteristics in the training and validation cohorts. Values are numbers of patients unless otherwise specified. The difference between cohorts was assessed using an independent samples t-test for continuous variables and the chi-squared test for categorical variables. *Surgery=neck resection prior to RT.

Patient characteristic	Training (N=86)	External validation (N=35)	p-value
Mean age (range) / years	60 (29–85)	56 (37–73)	0.073
Median MID at baseline (range) / mm	50 (35–68)	46 (21–68)	<0.001
Median MID at 6 months (range) / mm	40 (18–65)	45 (24–58)	0.014
Gender			0.111
Male	63 (73%)	31 (89%)	
Female	23 (27%)	4 (11%)	
Site			0.015
Oropharynx	41 (48%)	24 (68%)	
Oral cavity	11 (13%)	0 (0%)	
Nasopharynx	13 (15%)	8 (23%)	
Other	21 (24%)	3 (9%)	
Technique			
IMRT	42 (49%)	35 (100%)	<0.001
3D conformal	44 (51%)	0 (0%)	
Surgery*			<0.001
Yes	0 (0%)	17 (49%)	
No	86 (100%)	18 (51%)	
Chemotherapy			0.001
Yes (concurrent)	54 (60%)	33 (94%)	
No	32 (40%)	2 (6%)	
Smoking			0.870
Current	20 (23%)	4 (11%)	
Former	24 (28%)	14 (40%)	
Never	38 (44%)	17 (49%)	
Unknown	4 (5%)	0 (0%)	
Laterality			0.014
Left	36 (42%)	8 (23%)	
Right	21 (24%)	18 (51%)	
Bilateral	29 (34%)	9 (26%)	

Table 2:

Spearman rank correlation coefficient (R_s) and associated p-value between the mean dose and MID six months post-radiotherapy in the training cohort.

Structure	R_s	p
IBDM cluster	-0.36	0.001
Masseter	-0.35	0.001
Medial pterygoid	-0.32	0.002
Lateral pterygoid	-0.28	0.001
Temporalis	-0.16	0.08
TMJ	-0.24	0.007

Table 3:

Spearman rank correlation coefficient between MID predicted from fitting on the training cohort and MID in the validation cohort.

Variable	R_s	p-value
D_{mean} to IBDM cluster (univariate)	0.45	0.007
Baseline MID (univariate)	0.69	<0.001
Smoking (univariate)	0.05	0.77
Full model (multivariable)	0.76	<0.001