

# **HHS Public Access**

Author manuscript *Phys Med Biol.* Author manuscript; available in PMC 2020 December 18.

Published in final edited form as:

Phys Med Biol.; 64(17): 175018. doi:10.1088/1361-6560/ab2f13.

# Biomechanical modeling of neck flexion for deformable alignment of the salivary glands in head and neck cancer images

Molly M McCulloch<sup>1,2,8</sup>, Brian M Anderson<sup>1</sup>, Guillaume Cazoulat<sup>1</sup>, Christine B Peterson<sup>3</sup>, Abdallah S R Mohamed<sup>4</sup>, Stefania Volpe<sup>4,5</sup>, Hesham Elhalawani<sup>4</sup>, Houda Bahig<sup>4</sup>, Bastien Rigaud<sup>1</sup>, Jason B King<sup>1,6</sup>, Alexandra C Ford<sup>1</sup>, Clifton D Fuller<sup>4</sup>, Kristy K Brock<sup>1,7</sup>

<sup>1</sup>Department of Imaging Physics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, United States of America

<sup>2</sup>Department of Nuclear Engineering and Radiological Sciences, University of Michigan, Ann Arbor, MI 48109, United States of America

<sup>3</sup>Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, United States of America

<sup>4</sup>Department of Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, United States of America

<sup>5</sup>Radiation Oncology Department, European Institute of Oncology IRCSS, Milan, via Ripamonti 435, Milan, Italy

<sup>6</sup>Department of Biomedical Engineering, The University of Texas at Austin, Austin, TX 78705, United States of America

<sup>7</sup>Department of Radiation Physics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, United States of America

# Abstract

During head and neck (HN) cancer radiation therapy, analysis of the dose-response relationship for the parotid glands (PG) relies on the ability to accurately align soft tissue organs between longitudinal images. In order to isolate the response of the salivary glands to delivered dose, from deformation due to patient position, it is important to resolve the patient postural changes, mainly due to neck flexion. In this study we evaluate the use of a biomechanical model-based deformable image registration (DIR) algorithm to estimate the displacements and deformations of the salivary glands due to postural changes.

A total of 82 pairs of CT images of HN cancer patients with varying angles of neck flexion were retrospectively obtained. The pairs of CTs of each patient were aligned using bone-based rigid registration. The images were then deformed using biomechanical model-based DIR method that focused on the mandible, C1 vertebrae, C3 vertebrae, and external contour. For comparison, an intensity-based DIR was also performed. The accuracy of the biomechanical model-based DIR was assessed using Dice similarity coefficient (DSC) for all images and for the subset of images

<sup>&</sup>lt;sup>8</sup>Author to whom any correspondence should be addressed. mmmcculloch@mdanderson.org.

where the PGs had a volume change within 20%. The accuracy was compared to the intensitybased DIR.

The PG mean  $\pm$  STD DSC were  $0.63 \pm 0.18$ ,  $0.80 \pm 0.08$ , and  $0.82 \pm 0.15$  for the rigid registration, biomechanical model-based DIR, and intensity based DIR, respectively, for patients with a PG volume change up to 20%. For the entire cohort of patients, where the PG volume change was up to 57%, the PG mean  $\pm$  STD DSC were  $0.60 \pm 0.18$ ,  $0.78 \pm 0.09$ , and  $0.81 \pm 0.14$  for the rigid registration, biomechanical model-based DIR, and intensity based DIR, respectively. The difference in DSC of the intensity and biomechanical model-based DIR methods was not statistically significant when the volume change was less than 20% (two-sided paired *t*-test, *p* = 0.12). When all volume changes were considered, there was a significant difference between the two registration approaches, although the magnitude was small.

These results demonstrate that the proposed biomechanical model with boundary conditions on the bony anatomy can serve to describe the varying angles of neck flexion appearing in images during radiation treatment and to align the salivary glands for proper analysis of dose-response relationships. It also motivates the need for dose response modeling following neck flexion for cases where parotid gland response is noted.

#### Keywords

deformable image registration; head and neck cancer; volumetric response; image registration; biomechanical modeling

# 1. Introduction

During external beam radiation therapy (EBRT) of head and neck (HN) cancer, dose to the parotid glands (PG) can lead to toxicity (Deasy *et al* 2010). Studies have shown that limiting the mean dose to the salivary glands during HN radiation therapy leads to lower toxicity to the patient (Vainshtein *et al* 2015). However, preclinical studies show that considering specific subregions of the glands could improve dose response modeling (Konings *et al* 2005). Understanding the effect of the dose to the subregions of the glands over the course of radiotherapy is challenging due to the volumetric response combined with the sharp dose gradient within the glands. Determining the dose to subregions over the course of treatment requires spatial alignment of longitudinal images. Deformable registration of the gland is challenging due to the volumetric response, deformation, and the uniform contrast of the glands on CT images which remain the standard imaging modality in the radiation therapy workflow.

The majority of techniques proposed for DIR of longitudinal images of HN cancer patients use intensity-driven algorithms to deform one image so that the structure boundaries match those in the other image (Kierkels *et al* 2018, Loi *et al* 2018, Nix *et al* 2017, Liu *et al* 2018). Using these conventional DIR approaches, the estimated deformations inside PGs will only depend on the displacement vector field (DVF) regularization model of the DIR algorithm. However, the shrinkage scheme is likely more complex, correlating to the heterogeneity of the dose inside the organ and potentially the sensitivity of the local structure within the

gland. Biomechanical model-based DIR algorithms allows exploration of such complex deformations and volume change as the PGs respond to radiation therapy. Early work by Al-Mayah demonstrates that the use of dose-based boundary conditions in a biomechanical models can simulate the radiation dose response during HN radiation therapy (Al-Mayah *et al* 2015).

Based on the above study, biomechanical model-based DIR provides the potential to investigate the PGs response to radiation. However, in order to understand the PG response to radiation dose other sources of displacements and deformations of the glands should be resolved first. In particular, flexion of the neck or movement of the mandible are often observed between images acquired at different time points because of the difficulty in reproducing the patient position in the presence of weight loss or when acquisition is performed by different imaging devices, limiting the consistency of immobilization devices. In order to accurately model the volumetric changes of the PG subregions due to dose-response during radiation therapy, an initial alignment of the PGs is necessary to first resolve the changes due to neck flexion. A previous feasibility study (Al-Mayah *et al* 2010) investigated a biomechanical model of the HN with various boundary conditions. This study found that the highest accuracy, based on the DSC of the tumor and PGs, was found when placing boundary conditions on the vertebrae and mandible. However, the evaluation was tested on a small dataset of four pairs of images and requires further validation.

The goal of this study is to perform a comprehensive investigation and validation of the use of biomechanical model-based registration to resolve the misalignment of the salivary glands only due to those postural changes and compare to an intensity-based DIR method for a large cohort of patients demonstrating a range of levels of volumetric response of the parotid gland. The use of a commercially available biomechanical model of the patient anatomy with boundary conditions on relevant bones and external contours is proposed. For evaluation of the model, a cohort of 82 patients was retrospectively evaluated and for each patient a pair of CT scans presenting noticeable differences in the angle of neck flexion was selected. The overlap between mapped deformed gland volumes and original gland volumes were measured after solving the neck flexion using the biomechanical model-based DIR algorithm. In order to assess the accuracy of the biomechanical approach to resolve patient positioning, the performances of the biomechanical approach were compared to the performances of the rigid and intensity-based DIR methods for PGs with volume change within 20%, to evaluate the accuracy of alignment when little response is noted, as well as the whole cohort of patients, with volume change up to 57%, to evaluate the need of further dose-based boundary conditions on the glands when response is noted.

# 2. Methods

# 2.1. Patient data

A retrospective IRB approved evaluation was performed on 164 PG from 82 oropharynx cancer patients who underwent EBRT. For each patient, two non-contrast enhanced CTs showing neck flexion acquired on different dates were selected. 14 out of the 82 patients included in the study had bite blocks; however removing them from the cohort and analyzing them alone did not change the statistics of the study. All patients were treated at the

University of Texas MD Anderson Cancer Center. The patients in this study were not immobilized in the images used, as they were diagnostic scans, and thus, there was no need for masks as there is during radiation therapy. This caused the need to resolve neck flexion. Each patient was treated with intensity-modulated radiation therapy (IMRT) following previously described protocols (Gunn *et al* 2016). CT scans exhibiting differences of neck flexion angles were selected for each patient among the diagnostics images that were part of the treatment management. The normal tissue structures were previously auto-segmented (Admire ABAS, Elekta, Stockholm, Sweden) on all images. For this study, all images and contours were imported into a TPS with biomechanical model-based deformable registration capabilities (RayStation v6.99, RaySearch Laboratories, Stockholm, Sweden).

### 2.2. Image registration

The following registration and evaluation steps were performed automatically for all 82 patients using a Python-based script in the TPS.

**2.2.1.** Rigid registration—For each patient, rigid registration was performed between the two CTs using the option available in the TPS of focusing on the bony structures. This registration resulted in a compromise between the alignments of the different bony structures, mainly of the skull due to its relatively large volume. This registration step served as the initial registration for the DIR that followed. Examples of the rigid registrations are illustrated in figure 1. The bottom row of figure 1 shows the overlay of the images after a global rigid registration based on all bony anatomy. For all three alignments, the global rigid registration resulted in a perfect alignment of the skull. However, in the left column, both the vertebra and mandible were still misaligned. This can likely be explained by the bite block in the secondary image. For the case in the middle column, only the vertebra were still misaligned. Vertebra misalignment in the cases in the left and middle column can partially be explained by differences in the positioning of the patient in the CT scan. For the case in the third column, only the mandible was still misaligned. This can likely be explained by the patient's mouth is slightly open. It is expected that a biomechanical model driven by boundary conditions on the vertebrae and mandible would allow a better global alignment, particularly for the PG.

**2.2.2. Deformable image registration**—Following the initial rigid registration, both intensity and biomechanical-based DIRs were performed between the CTs of each patient. The intensity-based DIR algorithm used in this study was the ANAtomicaly CONstrained Deformation Algorithm (ANACONDA), commercially available in the TPS used for this study (RayStation v6.99, RaySearch Laboratories, Stockholm, Sweden). For the ANACONDA method, a smooth deformation vector field (DVF) is optimized using the quasi-Newton algorithm. Similarity between the images is determined by correlation coefficients, which guide the DVF. Regularization is attained with minimization of the weighted Dirichlet energy for coordinate functions of the DVF, and involves first resolving the DVF smoothness and invertibility, and then penalizing large deviations in the regions of interest (ROI) (Weistrand and Svensson 2015).

The biomechanical model-based DIR algorithm used for this study was the commercial implementation of Morfeus (Brock *et al* 2005) in the TPS. Briefly, Morfeus creates tetrahedral meshes from the contours of the body and organs included in the model and assigns elastic properties to each of them. For each organ, a surface projection method between the organ surface on the reference and secondary images determines the displacement of the surface nodes of the tetrahedral meshes. Those displacements are used as boundary conditions in the model to solve the displacement of all the internal mesh nodes in a finite-element analysis. In the proposed model, boundary conditions were applied on the mandible, C1 vertebrae, and C3 vertebrae, as well as the patient external contour. These boundary conditions were chosen for their proximity to the PG, and their capacity to describe most possible postural changes.

#### 2.3. Registration evaluation

Based on the rigid registration and two DIR methods (intensity and biomechanical), the left and right PG segmentations were propagated from the CT0 (the earlier dated CT) to CT1 (the later dated CT). The performance of each method to accurately propagate the PG structures were reported using the following metric:

**2.3.1.** Dice similarity coefficient (DSC)—The DSC, which calculates the overlap between two ROIs as defined by equation (1), was used to assess the accuracy of the DIR based on the overlap between deformed and original contours.

$$DSC = 2 \frac{|A \cap B|}{|A| + |B|}.$$
(1)

Where *A* is the primary ROI, and *B* is the secondary ROI. Based on this equation, a DSC of 1 indicates complete overlap, and a DSC of 0 indicates no overlap. The DSC was calculated in the TPS for the left and right PG, individually, based on the rigid registration, intensity-based DIR method, and the biomechanical model-based DIR method. For each case, the DSC was calculated between the deformed ROI mapped on the second CT and the original ROI on that CT. The mean DSC for PG contouring variability (based on multiple observers) is 0.76 (Nelms *et al* 2012).

**2.3.2.** Volume change—The volume of the left and right PG from each CT were calculated in the TPS in order to understand the analysis of the PG DSC. The DSC from the intensity-based DIR and the biomechanical model-based DIR were evaluated in cohorts based on absolute percent volume change of the PG. A two-tailed paired Student's *t*-test was used to assess the statistical significance between the PG DSC of each DIR method for cases with volume change within 20% and for the entire cohort, as 20% shrinkage of a sphere will yield a DSC of 0.89 with the original sphere. The DSC as a function of volume change is important as the intensity based registration should account for volume change, whereas the biomechanical based registration does not have boundary conditions on the PGs and therefore will not account for volume change (e.g. when the 2 images represent both neck flexion and PG volume chance, the biomechanical model, by definition, will only account for the neck flexion and the volume difference should be noted by a DSC less than 1).

DSC for cases with less than 20% volume change using biomechanical-based DIR will evaluate the accuracy and robustness of the algorithm to resolve the deformation and positional changes due to neck flexion. The DSC for cases with more than 20% volume change will evaluate the accuracy and robustness of the algorithm to partially resolve complete deformation, but highlight the potential need for further boundary constraints to resolve the volumetric response. The DSC for the intensity-based DIR will indicate the accuracy and robustness of the algorithm to resolve this complex deformation.

# 3. Results

Table 1 shows the mean, standard deviation (STD), maximum, and minimum DSC of the PG over all patients with volume change within 20% (N= 63). The intensity-based DIR method resulted in a slightly higher mean DSC (0.02) and slightly higher standard deviation (0.07) at  $(0.82 \pm 0.15)$  than the biomechanical model-based DIR method  $(0.80 \pm 0.08)$ , and both DIR methods resulted in a higher mean DSC and lower standard deviation than rigid registration  $(0.63 \pm 0.18)$ . No significant differences were observed (mean difference = 0.017, 95%) confidence interval = -0.004 to 0.038, p = 0.12) between the DIR methods and both methods had a DSC greater than the DSC reported for inter-observer variation (0.76). This result was partially driven by a limited number of cases with volume change < 20% where the intensity-based method performed very poorly. The minimum PG DSC was 0.17 for the intensity-based and 0.53 for the biomechanical model-based DIR methods, indicating the potential for more 'catastrophic' registration errors with the intensity-based algorithm. The minimum PG DSC for intensity was so low potentially due to artifact and a large angle of neck flexion between the two images. The volume change of the PG with the minimum DSC for the biomechanical model-based DIR method was 30%, potentially causing the low minimum DSC of 0.53. When the two observations with strongly outlying differences (where the intensity-based approach had unusually poor performance) were omitted, a significant difference in performance was observed, but the magnitude of this difference was small (mean difference = 0.029, 95% confidence interval = 0.016 to 0.042, p < 0.0001). The maximum PG DSC were 0.96 and 0.93 for intensity-based and biomechanical model-based, respectively.

Figure 2 is a histogram depicting the number of PG with DSC greater than values ranging from 0 to 1. All PG (n = 164) resulted in DSC greater than 0.1 for both methods, and 0 PG resulted in DSC greater than 0.95. Fifty (30%) PGs had a DSC that exceeded 0.9 for the intensity-based method, while only 11 (7%) PG exceeded 0.9 for the biomechanical model-based method. However, all of the biomechanical model-based registrations had a DSC of 0.5 or greater, whereas 157 (96%) of the intensity-based registrations had a DSC of 0.5 or greater.

Table 2 shows the mean, standard deviation (STD), maximum, and minimum DSC of the PG over all patients. The intensity-based DIR method resulted in a slightly higher mean DSC average (0.03) and slightly higher standard deviation (0.05) at (0.81  $\pm$  0.14) than the biomechanical model-based DIR method (0.78  $\pm$  0.09), and both DIR methods resulted in a higher mean DSC and lower standard deviation than rigid registration (0.60  $\pm$  0.18). Similar

to the subset evaluated above, both DIR algorithms had an average DSC greater than the inter-observer variability (0.76).

Analyzing all cases together, including those where the volume change exceeded 20%, the differences in DSC between the intensity-based DIR and biomechanical model-based DIR becomes statistically significant (mean difference = 0.030, 95% confidence interval = 0.015to 0.044, p < 0.0001). The significant difference between the PG DSC in cases with a 20% volume change demonstrates the need for additional boundary conditions to describe this volumetric response. Figure 3 depicts the results of the biomechanical model-based DIR method for a patient presenting large neck flexion. While the anterior part of the parotid glands appears well aligned, misalignment of the posterior part of the glands remained for the biomechanical model-based method, caused by the change in the PG volume of 30%. For this case, the intensity-based method matched all the boundaries with a DSC of 0.88, where the biomechanical model-based method, which successfully aligned the bony anatomy (indicating that Morfeus resolved the neck flexion), results in a DSC of 0.79. The larger STD and lower minimum DSC of the intensity-based registration demonstrates the potential advantage of the biomechanical model-based approach when potentially combined with a dose-based boundary condition on the PGs. Figure 4 illustrates the range of DSC for each method as a function of volume change.

# 4. Discussion

In this study, we retrospectively evaluated the registration of 82 repeat CT scans for oropharyngeal cancer using the DSC between the deformed and original delineations of the PGs. For each set of CT scans, DIR was performed using intensity-based registration for full image DIR and a biomechanical model-based methods focused only on resolving the neck flexion, and therefore PGs with minimal volume change were evaluated in an isolated cohort.

There was no statistically significant difference (p = 0.12) observed between the DIR methods for cases with a PG volume change within 20%: this result reflects the combination of DSC values which were generally slightly better using the intensity-based approach but were quite poor for a couple of exceptional cases. Both DIR methods had a median DSC of greater than the inter-observer variation. Additionally, the clinical significance of a 0.02 decrease in mean DSC is arguably minimal. The standard deviation for the intensity-based method was larger than for the biomechanical model-based method, and the minimum DSC was lower due to the outlying points mentioned above, indicating more potential failures in the results of the intensity-based method, which was further demonstrated by 3 PGs that had a DSC of less than 0.5 following intensity-based DIR.

This data demonstrates that modeling neck flexion alone using a biomechanical model-based registration algorithm aligns the PGs as accurately as an intensity-based registration and within the expected contour variation. Across the full data set, including observations with volume changes exceeding 20%, the results of each method were statistically significant, as expected given the lack of boundary conditions driving the alignment of the PGs in the biomechanical model. Future work will include applying dose-based boundary conditions to

the biomechanical model-based algorithm (Al-Mayah *et al* 2015) in order to build a comprehensive model that involves accurate neck flexion as well as a deformation model function of the dose distribution in the PG.

Evaluation of the DSC trends showed that there are 10 outliers (i.e. points > 3 \* STD) in the total 164 PGs for the intensity-based method and 0 for the biomechanical model-based method. This indicates more failures in the intensity-based method. Table 2 shows the minimum DSC over all patients for the intensity-based method was 0.17, while the minimum DSC for biomechanical model-based method was 0.53. There were 10 outlier cases for the intensity-based method so for 6% of the parotid glands analyzed. Of these 10 outlier cases, six parotid glands belonged to the same three patients. The other four parotid glands belonged to four different patients. One patient had major changes in the angle of neck flexion as well as metal artifact. The parotid glands of this patient were those with the lowest DSC from intensity-based (0.17 and 0.23). Four patients had major changes in neck flexion and metal artifact are likely reasons for the biomechanical model-based DIR outperforming the intensity-based DIR. This data suggests that for challenging cases where intensity-based methods fail, even a simple biomechanical model-based method may give reasonable results.

# 5. Conclusion

This retrospective study demonstrated that PG deformation due to neck flexion can be modeled using the biomechanical model-based method with indistinguishable results from the intensity-based method (p = 0.12 for volume change up to 20%). The minimum DSC based on the biomechanical-based method is 0.53, while the minimum DSC based on the intensity-based method is 0.17. This combined with the many outliers from the intensitybased method shows that while the intensity-based method can slightly outperform the biomechanical model-based method on average, there can be substantial failures using the intensity-based method, which is not observed with the biomechanical model-based method. The biomechanical model can be an initial step in a comprehensive model to describe the anatomical, patient positioning, and volumetric changes to the salivary glands during HN radiation therapy, and eventually aid in the development of toxicity models for this region.

# Acknowledgments

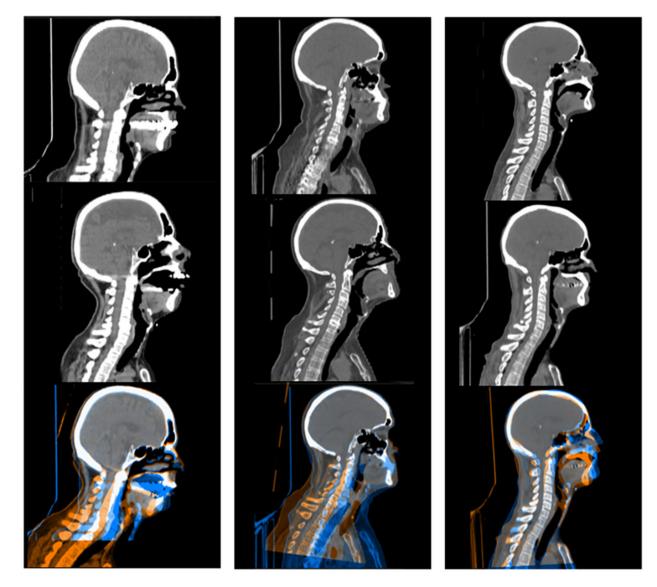
This work was funded in part by the Helen Black Image Guided Cancer Therapy Fund and a grant from RaySearch Laboratories. The work was also supported by the Image Guided Cancer Therapy Research Program at The University of Texas MD Anderson Cancer Center. Dr Brock has a licensing agreement with RaySearch Laboratories. Dr Peterson was partially supported by NIH/NCI CCSG Grant P30CA016672 (Biostatistics shared resource). This research is supported by the Andrew Sabin Family Foundation; Dr. Fuller is a Sabin Family Foundation Fellow. Dr. Fuller receives funding and salary support from the National Institutes of Health (NIH), including: the National Institute for Dental and Craniofacial Research Award (1R01DE025248/R56DE025248); a National Science Foundation (NSF), Division of Mathematical Sciences, Joint NIH/NSF Initiative on Quantitative Approaches to Biomedical Big Data (QuBBD) Grant (NSF 1557679); a National Institute of Biomedical Imaging and Bioengineering (NIBIB) Research Education Programs for Residents and Clinical Fellows Grant (R25EB025787-01); the NIH Big Data to Knowledge (BD2K) Program of the National Cancer Institute (NCI) Early Stage Development of Technologies in Biomedical Computing, Informatics, and Big Data Science Award (1R01CA218148); an NIH/NCI Cancer Center Support Grant (CCSG) Pilot Research Program Award from the UT MD Anderson CCSG Radiation Oncology and Cancer Imaging Program (P30CA016672) and an NIH/NCI Head

and Neck Specialized Programs of Research Excellence (SPORE) Developmental Research Program Award (P50 CA097007. Dr. Fuller has received direct industry grant support and travel funding from Elekta AB. Mr King was supported by an NIH T32 training grant, grant #T32EB007507.

# References

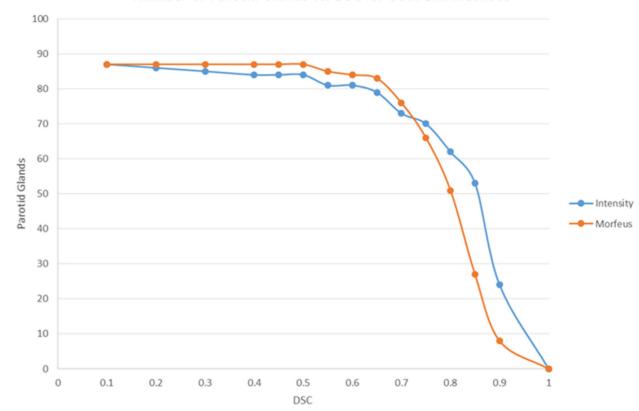
- Al-Mayah A, Moseley J, Hunter S and Brock K 2015 Radiation dose response simulation for biomechanical-based deformable image registration of head and neck cancer treatment Phys. Med. Biol 60 8481 [PubMed: 26485227]
- Al-Mayah A, Moseley J, Hunter S, Velec M, Chau L, Breen S and Brock K 2010 Biomechanical-based image registration for head and neck radiation treatment Phys. Med. Biol 55 6491 [PubMed: 20959687]
- Brock K, Sharpe M, Dawson L, Kim S and Jaffray D 2005 Accuracy of finite element model-based multi-organ deformable image registration Med. Phys 32 1647–59 [PubMed: 16013724]
- Deasy JO, Moiseenko V, Marks L, Chao KC, Nam J and Eisbruch A 2010 Radiotherapy dose-volume effects on salivary gland function Int. J. Radiat. Oncol. Biol. Phys 76 S58-63 [PubMed: 20171519]
- Gunn GB, Blanchard P, Garden AS, Zhu XR, Fuller CD, Mohamed AS, Morrison WH, Phan J, Beadle BM and Skinner HD 2016 Clinical outcomes and patterns of disease recurrence after intensity modulated proton therapy for oropharyngeal squamous carcinoma Int. J. Radiat. Oncol. Biol. Phys 95 360–7 [PubMed: 27084653]
- Kierkels RG, den Otter LA, Korevaar EW, Langendijk JA, van der Schaaf A, Knopf A-C and Sijtsema NM 2018 An automated, quantitative, and case-specific evaluation of deformable image registration in computed tomography images Phys. Med. Biol 63 045026 [PubMed: 29182154]
- Konings AW, Cotteleer F, Faber H, van Luijk P, Meertens H and Coppes RP 2005 Volume effects and region-dependent radiosensitivity of the PG Int. J. Radiat. Oncol. Biol. Phys 62 1090–5 [PubMed: 15990013]
- Liu C, Kim J, Kumarasiri A, Mayyas E, Brown SL, Wen N, Siddiqui F and Chetty IJ 2018 An automated dose tracking system for adaptive radiation therapy Comput. Methods Prog. Biomed 154 1–8
- Loi G, Fusella M, Lanzi E, Cagni E, Garibaldi C, Iacoviello G, Lucio F, Menghi E, Miceli R and Orlandini LC 2018 Performance of commercially available deformable image registration platforms for contour propagation using patient-based computational phantoms: a multi-institutional study Med. Phys 45 748–57 [PubMed: 29266262]
- Nelms BE, Tomé WA, Robinson G and Wheeler J 2012 Variations in the contouring of organs at risk: test case from a patient with oropharyngeal cancer Int. J. Radiat. Oncol. Biol. Phys 82 368–78 [PubMed: 21123004]
- Nix MG, Prestwich RJ and Speight R 2017 Automated, reference-free local error assessment of multimodal deformable image registration for radiotherapy in the head and neck Radiother. Oncol 125 478–84 [PubMed: 29100697]
- Vainshtein JM, Moon DH, Feng FY, Chepeha DB, Eisbruch A and Stenmark MH 2015 Long-term quality of life after swallowing and salivary-sparing chemo–intensity modulated radiation therapy in survivors of human papillomavirus–related oropharyngeal cancer Int. J. Radiat. Oncol. Biol. Phys 91 925–33 [PubMed: 25832685]
- Weistrand O and Svensson S 2015 The ANACONDA algorithm for deformable image registration in radiotherapy Med. Phys 42 40–53 [PubMed: 25563246]

McCulloch et al.



#### Figure 1.

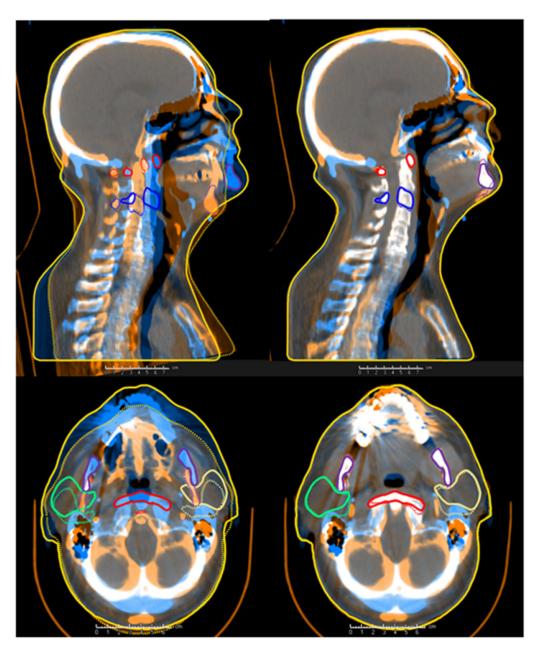
Examples of rigid registration performance for three cases. The top row shows the primary images, the middle row shows the secondary images, and the bottom row shows the rigid registration results (with focus on skeletal anatomy) between the primary and secondary images.



# Number of Parotid Glands vs. DSC for both DIR methods

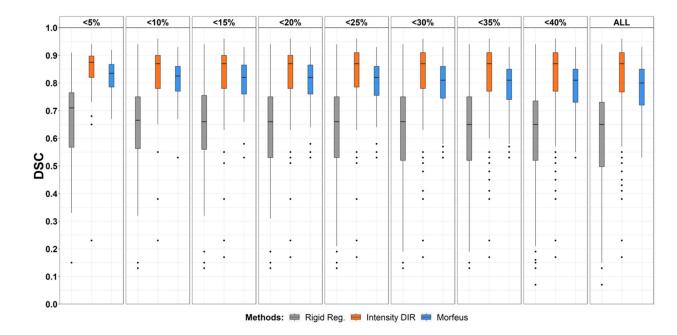
# Figure 2.

Histogram depicting the number of PG with DSC greater than the corresponding value, ranging from 0 to 1. The results of the intensity-based DIR method is depicted in blue and the results of Morfeus (the biomechanical model-based method) are depicted in orange, for patients with < 20% volume change.



#### Figure 3.

Top: Sagittal view of a HN patients with sub substantial neck flexion and parotid gland volumetric response between each image. The rigid alignment is the shown on the left and the alignment based on the Morfeus DIR, with boundary conditions on the C1, C3, mandible, and external body is shown on the right. Morfeus successfully aligns the mandible in this case. Bottom: Axial view of the same HN patient showing the results of the PG alignment. The rigid alignment is shown on the left and the alignment based on the Morfeus DIR, with boundary conditions on the C1, C3, mandible, and external body is shown on the left and the alignment based on the Morfeus DIR, with boundary conditions on the C1, C3, mandible, and external body is shown on the right.



# Figure 4.

Boxplots of the DSC obtained for rigid registration (gray), intensity-based DIR (orange), and biomechanical model-based DIR (blue) for different volume changes of the DSC obtained for the propagated parotid gland contours based on each DIR method. PG 5 represents volume change of 5%, and the same for other values. Outliers (i.e. points > 3\*STD) are shown in the bubbles below each box plot. Median DSCs are represented by the horizontal lines, the lower whiskers represent the minimum, the upper whiskers represent the maximum, the base of the box represents the lower quartile, and the ceiling of each box represents the upper quartile. Intensity-based and biomechanical model-based results are shown in pairs for each volume change threshold.

#### Table 1.

Registration method comparison according to PG DSC over all patients with volume change within 20% (N= 63). For each method, the mean, STD, max, and min DSC values of the PG are reported. The percentage of patients with a DSC superior to 0.75 are also reported.

DIR method	Mean	STD	Min	Median	Max	% of PG with DSC > 0.75
Rigid	0.63	0.18	0.07	0.66	0.94	20%
Intensity-based	0.82	0.15	0.17	0.87	0.96	80%
Biomechanical model-based	0.80	0.08	0.53	0.82	0.93	76%

#### Table 2.

Registration method comparison according to PG DSC for all patients. For each method, the mean, STD, max, and min DSC values of the PG are reported. The percentage of patients with a DSC superior to 0.75 are also reported.

DIR method	Mean	STD	Max	Min	% of PG with DSC > 0.75
Rigid	0.60	0.18	0.94	0.07	21%
Intensity-based	0.81	0.14	0.96	0.17	77%
Biomechanical model-based	0.78	0.09	0.93	0.53	67%