

SCIENTIFIC DATA

OPEN Data Descriptor: Imaging and clinical data archive for head and neck squamous cell carcinoma patients treated with radiotherapy

Received: 4 January 2018

Accepted: 5 June 2018

Published: xx xx 2018

Aaron J. **Grossberg**^{1,2,†}, Abdallah S. R. **Mohamed**^{1,3,†}, Hesham **El Halawani**^{1,†}, William C. **Bennett**^{4,‡}, Kirk E. **Smith**^{4,‡}, Tracy S. **Nolan**^{4,‡}, Bowman **Williams**^{1,†}, Sasikarn **Chamchod**^{1,5,†}, Jolien **Heukelom**^{1,6,†}, Michael E. **Kantor**^{1,†}, Theodora **Browne**^{1,†}, Katherine A. **Hutcheson**^{7,†}, G. **Brandon Gunn**^{1,†}, Adam S. **Garden**^{1,†}, William H. **Morrison**^{1,†}, Steven J. **Frank**^{1,†}, David I. **Rosenthal**^{1,†}, John B. **Freyermann**^{8,‡} & Clifton D. **Fuller**^{1,†}

Cross sectional imaging is essential for the patient-specific planning and delivery of radiotherapy, a primary determinant of head and neck cancer outcomes. Due to challenges ensuring data quality and patient de-identification, publicly available datasets including diagnostic and radiation treatment planning imaging are scarce. In this data descriptor, we detail the collection and processing of computed tomography based imaging in 215 patients with head and neck squamous cell carcinoma that were treated with radiotherapy. Using cross sectional imaging, we calculated total body skeletal muscle and adipose content before and after treatment. We detail techniques for validating the high quality of these data and describe the processes of data de-identification and transfer. All imaging data are subject- and date-matched to clinical data from each patient, including demographics, risk factors, grade, stage, recurrence, and survival. These data are a valuable resource for studying the association between patient-specific anatomic and metabolic features, treatment planning, and oncologic outcomes, and the first that allows for the integration of body composition as a risk factor or study outcome.

Design Type(s)	subject-based data analysis objective • dataset evaluation objective • data annotation objective • image creation and editing objective
Measurement Type(s)	tumor measurement • body composition measurement
Technology Type(s)	Imaging Technique • Diagnostic Imaging
Factor Type(s)	tumor grading • Stage • site
Sample Characteristic(s)	Homo sapiens

¹Department of Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA. ²Department of Radiation Medicine, Oregon Health and Science University, Portland, Oregon 97238, USA.

³Department of Clinical Oncology and Nuclear Medicine, Faculty of Medicine, University of Alexandria, Alexandria, Egypt. ⁴Department of Biomedical Informatics, University of Arkansas for Medical Sciences, Little Rock, Arkansas 72205, USA. ⁵Radiation Oncology Unit, Chulabhorn Hospital, Bangkok 10210, Thailand.

⁶Department of Radiation Oncology, Netherlands Cancer Institute, Amsterdam, The Netherlands. ⁷Department of Head and Neck Surgery, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA.

⁸Leidos Biomedical Research, Inc. Frederick National Laboratory for Cancer Research, Frederick, Maryland 20892, USA. [†]The MD Anderson Head and Neck Quantitative Imaging Working Group. [‡]In concert with The Cancer Imaging Archive. Correspondence and requests for materials should be addressed to A.J.G.

(email: grossber@ohsu.edu) or to C.D.F. (email: cdfuller@mdanderson.org).

Q3
Q4