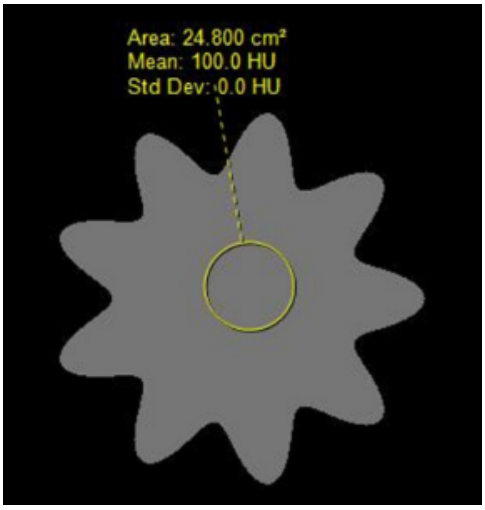
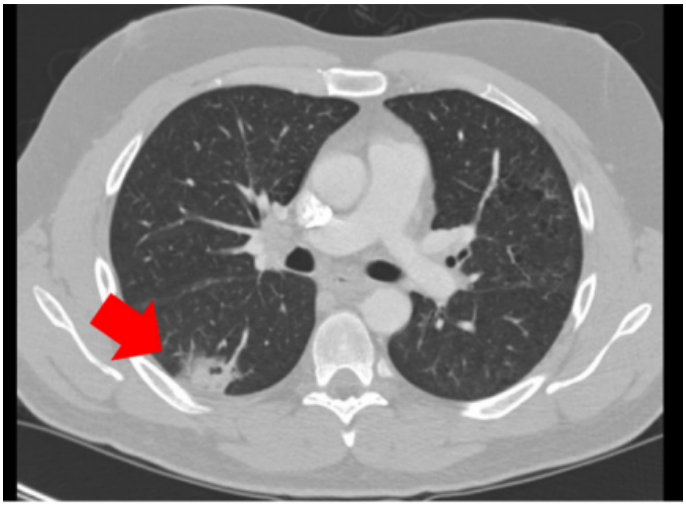
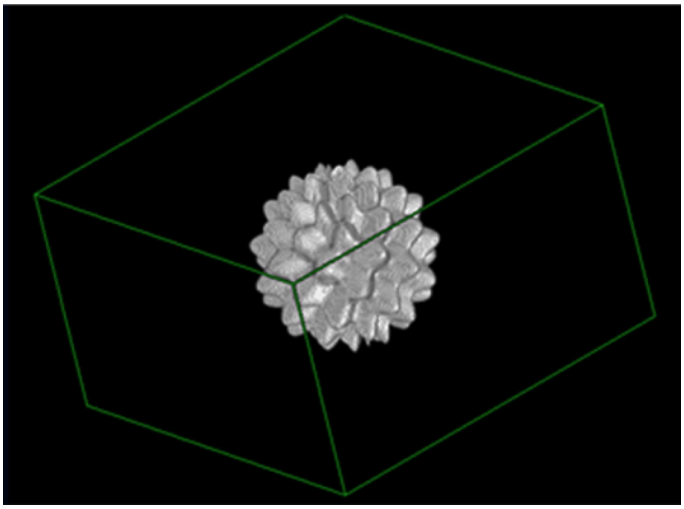
Standardization in Quantitative Imaging: A Multi-center Comparison of Radiomic Feature Values (Radiomic-Feature-Standards)

Summary

This dataset was used by the NCI's Quantitative Imaging Network (QIN) PET-CT Subgroup for their project titled: **Multi-center Comparison of Radiomic Features from Different Software Packages on Digital Reference Objects and Patient Datasets**. The purpose of this project was to assess the agreement among radiomic features when computed by several groups by using different software packages under very tightly controlled conditions, which included common image data sets and standardized feature definitions.







The image datasets (and Volumes of Interest – VOIs) provided here are the same ones used in that project and reported in the publication listed below (ISS N 2379-1381 https://doi.org/10.18383/j.tom.2019.00031). In addition, we have provided detailed information about the software packages used (Table 1 in that publication) as well as the individual feature value results for each image dataset and each software package that was used to create the summary tables (Tables 2, 3 and 4) in that publication.

For that project, nine common quantitative imaging features were selected for comparison including features that describe morphology, intensity, shape, and texture and that are described in detail in the International Biomarker Standardisation Initiative (IBSI, https://arxiv.org/abs/1612.07003 and publication (Zwanenburg A. Vallières M, et al, **The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping**. Radiology. 2020 May;295(2):328-338. doi: https://doi.org/10.1148/radiol.2020191145).

Acknowledgements

The authors gratefully acknowledge the following sources of support:

- The National Cancer Institute Quantitative Network (QIN)

Data Access

Data Access

Data Type	Download all or Query/Filter	License
Segmentation (NIfTI, zip, 4 MB)	*Download*	CC BY 3.0
Feature Variability Software Package details (xlsx, 13 kb)	*Download*	CC BY 3.0
DRO Results (xlsx, 31 kb)	*Download*	CC BY 3.0
Patient Dataset Results (xlsx, 400 kb)	*Download*	CC BY 3.0
Harmonized GLCM Entropy Results (xlsx, 17 kb)	*Download*	CC BY 3.0

Collections Used in this Third Party Analysis

Below is a list of the Collections used in these analyses:

Source Data Type	Download	License
Corresponding Original CT images from LIDC-IDRI and DRO-Toolkit (DICOM, 2.0 GB)	*Download* *Search* (Requires NBIA Data Retriever.)	CC BY 3.0
Corresponding second-generation SEG images from QIN-LungCT-Seg (DICOM, 123 MB)	*Download* *Search* (Requires NBIA Data Retriever.)	CC BY 3.0

- LIDC-IDRI
- Stanford DRO Toolkit: Digital Reference Objects for Standardization of Radiomic Features (DRO Toolkit)
- QIN-LungCT-Seg

Detailed Description

Detailed Description

DICOM Image Statistics	
Modalities	CT, SEG
Number of Patients	13
Number of Studies	13
Number of Series	26
Number of Images	3,867
Images Size (GB)	2 GB

There are three datasets provided – two image datasets and one dataset consisting of four excel spreadsheets containing feature values.

1. The first image dataset is a set of three Digital Reference Objects (DROs) used in the project, which are: (a) a sphere with uniform intensity, (b) a sphere with intensity variation (c) a nonspherical (but mathematically defined) object with uniform intensity. These DROs were created by the team at Stanford University and are described in (Jaggi A, Mattonen SA, McNitt-Gray M, Napel S. **Stanford DRO Toolkit: digital reference objects for standardization of radiomic features.** Tomography. 2019;6:–.) and are a subset of the DROs described in Stanford DRO Toolkit: Digital Reference Objects for Standardization of Radiomic Features. Each DRO is represented in both DICOM and NIfTI format and the VOI was provided in each format as well (DICOM Segmentation Object (DSO) as well as NIfTI segmentation boundary).
2. The second image dataset is the set of 10 patient CT scans, originating from the **LIDC-IDRI** dataset, that were used in the **QIN multi-site collection of Lung CT** data with Nodule Segmentations project (https://doi.org/10.7937/K9/TCIA.2015.1BUVFJR7). In that QIN study, a single lesion from each case was identified for analysis and then nine VOIs were generated using three repeat runs of three segmentation algorithms (one from each of three academic institutions) on each lesion. To eliminate one source of variability in our project, only one of the VOIs previously created for each lesion was identified and all sites used that same VOI definition. The specific VOI chosen for each lesion was the first run of the first algorithm (algorithm 1, run 1). DICOM images were provided for each dataset and the VOI was provided in both DICOM Segmentation Object (DSO) and NIfTI segmentation formats.

3. The third dataset is a collection of four excel spreadsheets, each of which contains detailed information corresponding to each of the four tables in the publication. For example, the raw feature values and the summary tables for Tables 2,3 and 4 reported in the publication cited (https://doi.org/10.18383/j.tom.2019.00031). These tables are:

    Software Package details : This table contains detailed information about the software packages used in the study (and listed in Table 1 in the publication) including version number and any parameters specified in the calculation of the features reported.

    DRO results : This contains the original feature values obtained for each software package for each DRO as well as the table summarizing results across software packages (Table 2 in the publication) .

    Patient Dataset results: This contains the original feature values for each software package for each patient dataset (1 lesion per case) as well as the table summarizing results across software packages and patient datasets (Table 3 in the publication).

    Harmonized GLCM Entropy Results : This contains the values for the "Harmonized" GLCM Entropy feature for each patient dataset and each software package as well as the summary across software packages (Table 4 in the publication).

Patient IDs for the 3 DROs from (https://doi.org/10.7937/t062-8262)
Phantom-100.0-1.0-1.0-1.0-9.0-0.0-100.0-10.0-0.0-0.0
Phantom-100.0-1.0-1.0-1.0-9.0-0.0-100.0-10.0-50.0-0.0
Phantom-100.0-1.0-1.0-1.0-9.0-0.2-100.0-10.0-0.0-0.0

Patient IDs for the 10 LIDC-IDRI subjects (https://doi.org/10.7937/K9/TCIA.2015.LO9QL9SX)
LIDC-IDRI-0314
LIDC-IDRI-0325
LIDC-IDRI-0580
LIDC-IDRI-0766
LIDC-IDRI-0771
LIDC-IDRI-0811
LIDC-IDRI-0905
LIDC-IDRI-0963
LIDC-IDRI-0965
LIDC-IDRI-1012

Additional options for download:

DRO Data (3 subjects)	Download all or Query/Filter
Image Data (DICOM, 452.0 MB) CT only	*Download* *Search* (Requires NBIA Data Retriever.)
Segmentation Data - DSO (DICOM, 29.0 MB)	*Download* *Search* (Requires NBIA Data Retriever.)
Segmentation Data - (NIfTI, zip, 926 KB)	*Download*

Patient Datasets (10 subjects)	Download all or Query/Filter
Image Data (DICOM, 1.0 GB) CT only	*Download* *Search* (Requires NBIA Data Retriever.)
Segmentation Data - (DICOM, 94 MB)	*Download* *Search* (Requires NBIA Data Retriever.)
Segmentation Data - (NIfTI, zip, 21.0 KB)	*Download*

Citations & Data Usage Policy

Citations & Data Usage Policy

Users must abide by the TCIA Data Usage Policy and Restrictions. Attribution should include references to the following citations:

Data Citation

McNitt-Gray, M.*, Napel, S.*, Jaggi, A., Mattonen, S.A., Hadjiiski, L., Muzi, M., Goldgof, D., Balagurunathan, Y., Pierce, L.A., Kinahan, P.E., Jones, E.F., Nguyen, A., Virkud, A., Chan, H-P., Emaminejad, N., Wahi-Anwar, M., Daly, M., Abdalah, M., Yang, H., Lu, L., Lv, W., Rahmim, A., Gastounioti, A., Pati, S., Bakas, S., Kontos, D., Zhao, B., Kalpathy-Cramer, J., Farahani, K. (2020). *Data from the* **Standardization in**

Quantitative Imaging: A Multi-center Comparison of Radiomic Feature Values [Data set]. The Cancer Imaging Archive. DOI: https://doi.org/10.7937/tcia.2020.9era-gg29.

Publication Citation

McNitt-Gray, M., Napel, S., Jaggi, A., Mattonen, S.A., Hadjiiski, L., Muzi, M., Goldgof, D., Balagurunathan, Y., Pierce, L.A., Kinahan, P.E., Jones, E.F., Nguyen, A., Virkud, A., Chan, H-P., Emaminejad, N., Wahi-Anwar, M., Daly, M., Abdalah, M., Yang, H., Lu, L., Lv, W., Rahmim, A., Gastounioti, A., Pati, S., Bakas, S., Kontos, D., Zhao, B., Kalpathy-Cramer, J., Farahani, K. (2020). **Standardization in Quantitative Imaging: A Multi-center Comparison of Radiomic Feature Values,** Tomography. https://doi.org/10.18383/j.tom.2019.00031.

TCIA Citation

Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., Koppel, P., Moore, S., Phillips, S., Maffitt, D., Pringle, M., Tarbox, L., & Prior, F. (2013). **The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository**. Journal of Digital Imaging, 26(6), 1045–1057. https://doi.org/10.1007/s10278-013-9622-7

Acknowledgement - Grant support

- David Geffen School of Medicine at UCLA - U01CA181156

Acknowledgement - Grant support

- Stanford University School of Medicine – U01CA187947 and U24CA180927

Acknowledgement - Grant support

- University of Michigan - U01CA232931

Acknowledgement - Grant support

- University of Washington – R50CA211270, U01CA148131

Acknowledgement - Grant support

- University of South Florida - U24CA180927, U01CA200464

Acknowledgement - Grant support

- Moffitt Cancer Center – U01CA143062, U01CA200464, P30CA076292

Acknowledgement - Grant support

- UC San Francisco - U01CA225427

Acknowledgement - Grant support

- BC Cancer Research Centre - NSERC Discovery Grant: RGPIN-2019-06467

Acknowledgement - Grant support

- Columbia University- U01CA225431

Acknowledgement - Grant support

- Center for Biomedical Image Computing and Analytics at the University of Pennsylvania - U24CA189523, R01NS042645

Acknowledgement - Grant support

- Massachusetts General Hospital- U01CA154601, U24CA180927

In addition to the dataset citation above, please be sure to cite the following if you utilize these data in your research:

Analysis Citation

Kalpathy-Cramer, J., Napel, S., Goldgof, D., Zhao, B. (2015). **Multi-site collection of Lung CT data with Nodule Segmentations**. The Cancer Imaging Archive. https://doi.org/10.7937/K9/TCIA.2015.1BUVFJR7

Other Publications Using This Data

TCIA maintains a list of publications which leverage TCIA data. If you have a manuscript you'd like to add please contact TCIA's Helpdesk.

Versions

Version 1 (Current): Updated 2020/06/09

Data Type	Download all or Query/Filter
Corresponding Original CT images from LIDC-IDRI and DRO-Toolkit (DICOM, 2.0 GB)	*Download* *Search* (Requires NBIA Data Retriever.)
Corresponding second-generation SEG images from QIN-LungCT-Seg (DICOM, 123 MB)	*Download* *Search* (Requires NBIA Data Retriever.)
Segmentation (NIfTI, zip, 4 MB)	*Download*
Feature Variability Software Package details (xlsx)	*Download*
DRO Results (xlsx)	*Download*
Patient Dataset Results (xlsx)	*Download*
Harmonized GLCM Entropy Results (xlsx)	*Download*