

NSCLC-Radiomics-Genomics

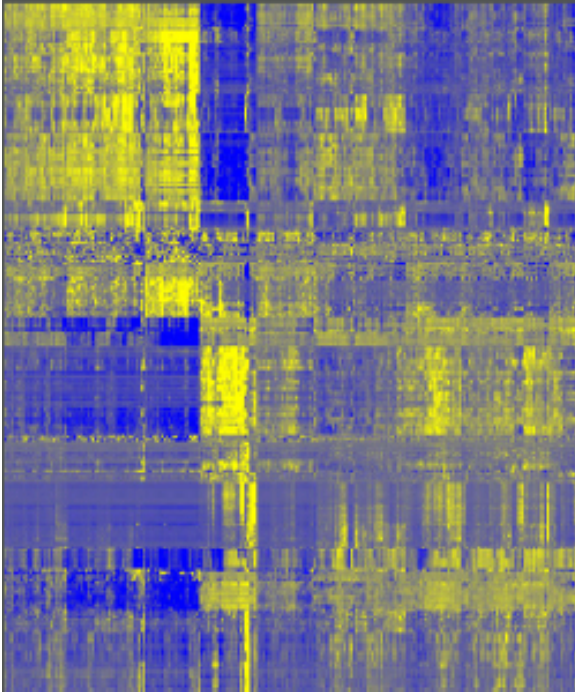
Summary

This collection contains images from 89 non-small cell lung cancer (NSCLC) patients that were treated with surgery. For these patients pretreatment CT scans, gene expression, and clinical data are available. This dataset refers to the Lung3 dataset of the [study published in Nature Communications](#).

In short, this publication applies a radiomic approach to computed tomography data of 1,019 patients with lung or head-and-neck cancer. Radiomics refers to the comprehensive quantification of tumour phenotypes by applying a large number of quantitative image features. In present analysis 440 features quantifying tumour image intensity, shape and texture, were extracted. We found that a large number of radiomic features have prognostic power in independent data sets, many of which were not identified as significant before. Radiogenomics analysis revealed that a prognostic radiomic signature, capturing intra-tumour heterogeneity, was associated with underlying gene-expression patterns. These data suggest that radiomics identifies a general prognostic phenotype existing in both lung and head-and-neck cancer. This may have a clinical impact as imaging is routinely used in clinical practice, providing an unprecedented opportunity to improve decision-support in cancer treatment at low cost.

The dataset described here (Lung3) was used to investigate the association of radiomic imaging features with gene-expression profiles. The Lung2 dataset used for training the radiomic biomarker and consisting of 422 NSCLC CT scans with outcome data can be found here: [NSCLC-Radiomics](#).

For scientific inquiries about this dataset, please contact Dr. Hugo Aerts of the Dana-Farber Cancer Institute / Harvard Medical School (hugo_aerts@dfci.harvard.edu).



Data Access

Data Access

Click the **Download** button to save a ".tcia" manifest file to your computer, which you must open with the [NBIA Data Retriever](#). Click the **Search** button to open our Data Portal, where you can browse the data collection and/or download a subset of its contents.

Data Type	Download all or Query/Filter
Images (DICOM, 6.6GB)	<input type="button" value="Download"/> <input type="button" value="Search"/>
Lung3 clinical (CSV)	<input type="button" value="Download"/>
Gene Expression (web)	<input type="button" value="Search"/>

Click the Versions tab for more info about data releases.

Detailed Description

Detailed Description

Collection Statistics	
Modalities	CT
Number of Patients	89
Number of Studies	89
Number of Series	89
Number of Images	13,482

Image Size (GB)	6.6
-----------------	-----

Gene-expression Data

Corresponding microarray data acquired for the imaging samples are available at National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (Link to GEO: <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE58661>). The patient names used to identify the cases on GEO are identical to those used in the DICOM files on TCIA and in the clinical data spreadsheet.

Clinical Data

Corresponding clinical data can be found here: [Lung3.metadata.xls](#).

Please note that survival time is measured in days from start of treatment. DICOM patients names are identical in TCIA and clinical data file.

Citations & Data Usage Policy

Citations & Data Usage Policy

This collection **may not be used for commercial purposes**. This collection is freely available to browse, download, and use **for scientific and educational purposes** as outlined in the [Creative Commons Attribution 3.0 Unported License](#).

See TCIA's [Data Usage Policies and Restrictions](#) for additional details. Questions may be directed to help@cancerimagingarchive.net.

Please be sure to include the following citations in your work if you use this data set:

i Data Citation

Aerts, Hugo J. W. L., Rios Velazquez, Emmanuel, Leijenaar, Ralph T. H., Parmar, Chintan, Grossmann, Patrick, Carvalho, Sara, ... Lambin, Philippe. (2015). Data From NSCLC-Radiomics-Genomics. The Cancer Imaging Archive. <http://doi.org/10.7937/K9/TCIA.2015.L4FRET6Z>

i Publication Citation

Aerts, H. J. W. L., Velazquez, E. R., Leijenaar, R. T. H., Parmar, C., Grossmann, P., Cavalho, S., ... Lambin, P. (2014, June 3). Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nature Communications. Nature Publishing Group. <http://doi.org/10.1038/ncomms5006> (link)

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. **The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository**, Journal of Digital Imaging, Volume 26, Number 6, December, 2013, pp 1045-1057. ([paper](#))

Other Publications Using This Data

TCIA maintains [a list of publications](#) that leverage our data. At this time we are not aware of any additional publications based on this data. If you have a publication you'd like to add, please [contact the TCIA Helpdesk](#).

Versions

Version 1 (Current): Updated 2014/07/02

Data Type	Download all or Query/Filter
Images (DICOM, 6.6GB)	<div style="display: flex; gap: 10px;">  Download  Search </div> <p>(Requires the NBIA Data Retriever.)</p>
Lung3 clinical (CSV)	<div style="display: flex; gap: 10px;">  Download </div>
Gene Expression (web)	<div style="display: flex; gap: 10px;">  Search </div>