Journal of Visualized Experiments

Predicting treatment response to image-guided therapies using machine learning: An example for trans-arterial treatment of hepatocellular carcinoma

--Manuscript Draft--

Article Type:	Invited Methods Article - JoVE Produced Video
Manuscript Number:	JoVE58382R1
Full Title:	Predicting treatment response to image-guided therapies using machine learning: An example for trans-arterial treatment of hepatocellular carcinoma
Keywords:	Machine learning, artificial intelligence, interventional radiology, hepatocellular carcinoma, trans-arterial chemoembolization, supervised machine learning, predictive modeling, predicting outcomes, pre-procedure planning
Corresponding Author:	Julius Chapiro XX
Corresponding Author's Institution:	
Corresponding Author E-Mail:	julius.chapiro@yale.edu
Order of Authors:	Aaron Abajian
	Nikitha Murali
	Lynn Jeanette Savic
	Fabian Max Laage-Gaupp
	Nariman Nezami
	James S. Duncan
	Todd Schlachter
	MingDe Lin
	Jean-François Geschwind
	Julius Chapiro
Additional Information:	
Question	**Response**
Please indicate whether this article will be Standard Access or Open Access.	Standard Access (US$2,400)
Please indicate the **city, state/province, and country** where this article will be **filmed**. Please do not use abbreviations.	300 Cedar Street, Yale School of Medicine, New Haven, CT, 06520

TITLE:

Predicting Treatment Response to Image-Guided Therapies Using Machine Learning: An Example for Trans-Arterial Treatment of Hepatocellular Carcinoma

AUTHORS & AFFILIATIONS:

Aaron Abajian, MD, Nikitha Murali, BA, Lynn Jeanette Savic, MD, Fabian Max Laage-Gaupp, MD, Nariman Nezami, MD, James S. Duncan, PhD, Todd Schlachter, MD, MingDe Lin, PhD, Jean-François Geschwind, MD, and Julius Chapiro, MD

A.A., N.M., L.J.S., F. M.L.-G., N.N., T.S., J.C.
Department of Radiology and Biomedical Imaging Yale School of Medicine
333 Cedar Street, New Haven, CT

L.J.S.
Department of Diagnostic and Interventional Radiology
Universitätsmedizin Charité Berlin
Berlin, Germany

J.S.D.
Department of Biomedical Engineering
Yale School of Engineering and Applied Science
New Haven, Connecticut

M.L.
Philips Research North America
Cambridge, Massachusetts

J.F.G.
Prescience Labs
Westport, Connecticut

aaron.abajian@yale.edu
nikhitha.murali@yale.edu
lynn.savic@yale.edu
fabian.laage-gaupp@yale.edu
nariman.nezami@yale.edu
james.duncan@yale.edu
todd.schlachter@yale.edu
mingde.lin@yale.edu
jeff.geschwind@gmail.com
julius.chapiro@yale.edu

Corresponding Author:
Aaron Abajian

aaron.abajian@yale.edu

KEYWORDS:
Machine learning, artificial intelligence, interventional radiology, hepatocellular carcinoma, trans-arterial chemoembolization, supervised machine learning, predictive modeling, predicting outcomes, pre-procedure planning

SHORT ABSTRACT:
Intra-arterial therapies are the standard of care for patients with hepatocellular carcinoma who cannot undergo surgical resection. A method for predicting response to these therapies is proposed. The technique uses pre-procedural clinical, demographic, and imaging information to train machine learning models capable of predicting response prior to treatment.

LONG ABSTRACT:
Intra-arterial therapies are the standard of care for patients with hepatocellular carcinoma who cannot undergo surgical resection. The objective of this study was to develop a method to predict response to intra-arterial treatment prior to intervention.

The method provides a general framework for predicting outcomes prior to intra-arterial therapy. It involves pooling clinical, demographic and imaging data across a cohort of patients and using these data to train a machine learning model. The trained model is applied to new patients in order to predict their likelihood of response to intra-arterial therapy.

The method entails the acquisition and parsing of clinical, demographic and imaging data from *N* patients who have already undergone trans-arterial therapies. These data are parsed into discrete features (age, sex, cirrhosis, degree of tumor enhancement, *etc.*) and binarized into true/false values (*e.g.,* age over 60, male gender, tumor enhancement beyond a set threshold, *etc.*). Low-variance features and features with low univariate associations with the outcome are removed. Each treated patient is labeled according to whether they responded or did not respond to treatment. Each training patient is thus represented by a set of binary features and an outcome label. Machine learning models are trained using $N - 1$ patients with testing on the left-out patient. This process is repeated for each of the *N* patients. The *N* models are averaged to arrive at a final model.

The technique is extensible and enables inclusion of additional features in the future. It is also a generalizable process that may be applied to clinical research questions outside of interventional radiology. The main limitation is the need to derive features manually from each patient. A popular modern form of machine learning called deep learning does not suffer from this limitation, but requires larger datasets.

INTRODUCTION:
Patients with hepatocellular carcinoma who are not surgical candidates are offered intra-arterial therapies[1,2,3]. There is no single metric that determines whether a patient will respond to an intra-arterial therapy before the treatment is administered. The objective of this study

was to demonstrate a method that predicts treatment response by applying methods from machine learning. Such models provide guidance to practitioners and patients when choosing whether to proceed with a treatment.

The protocol entails a reproducible process for training and updating a model starting from primary patient data (clinical notes, demographics, laboratory data, and imaging). The data is initially parsed for specific features, with each patient represented by a set of binary features and a binary outcome target label. The outcome label is determined using an established imaging-based response criterion for hepatocellular therapy[4,5,6,7]. The features and target labels are passed to machine learning software that learns the mapping between features and outcomes under a specific learning model (logistic regression or random forest)[8,9,10]. Similar techniques have been applied in radiology and other areas of cancer research for diagnosis and treatment prediction[11,12,13].

The method adapts techniques from computer science to the field of interventional radiology. Traditional significance studies in interventional radiology, and medicine in general, rely upon mono- or oligo- feature analyses. For example, the Model for End-Stage Liver Disease incorporates five clinical metrics to assess the extent of liver disease. The benefit of the proposed method is the ability to add features liberally; twenty-five features are considered in the example analysis. Additional features may be added as desired.

The technique may be applied to other radiographic interventions where pre- and post-intervention imaging data are available. For example, outcomes following percutaneous treatments could be predicted in a similar manner. The main limitation of the study is the need to manual curate features for inclusion in the model. Data curation and feature extraction is time-consuming for the practitioner and may impede clinical adoption of such machine learning models.

PROTOCOL:

1. Workstation Setup for Machine Learning

1.1. Use a system with the following:

Intel Core 2 Duo or higher CPU at 2.0 GHz
4 GB or more system memory
POSIX-compliant operating system (Linux or Mac OS) or Microsoft Windows 7
User permissions for executing programs and saving files

1.2. Install the following tools:

Anaconda Python3: https://www.anaconda.com/download
DICOM to NIfTI converter (dcm2niix) - https://github.com/rordenlab/dcm2niix
Sublime Text Editor: https://www.sublimetext.com/
itk-SNAP (optional): http://www.itksnap.org/

1.2.1. Install Anaconda Python3, dcm2nii and Sublime Text, visit their respective websites for operating system specific installation steps.

1.2.2. Create and activate an Anaconda environment.

```
conda create --name mlenv
conda activate mlenv
```

1.2.3. Install Anaconda packages for machine learning.

```
conda install numpy scipy scikit-learn nltk nibabel
```

Note: The nltk package is useful for parsing plaintext clinical notes, while the nibabel package provides useful functions for medical image manipulation. itk-SNAP may be installed for segmenting organs and tumors from medical images. It is useful for constraining features to specific regions.

2. Feature Extraction from Plaintext Clinical Notes and Structured Clinical Data

2.1. Create a parent directory for the project and create a folder for each patient within the parent folder. Directory structure should resemble:

```
Project/
Project/Patient_1/
Project/Patient_2/
Project/Patient_3/
...
```

2.2. Obtain plaintext clinical notes from the electronic medical record (EMR). Retrieve notes manually through the EMR or by means of the hospital information technology (IT) office through a data-dump. Store each patient's notes in their respective folders.

```
Project/Patient_1/History_and_Physical.txt
Project/Patient_1/Procedure_Note.txt
```

2.1.1. Decided which clinical features to include in the model. Parse the plaintext clinic notes for these features. The Python Natural Language Toolkit (nltk) library provides useful commands for splitting documents into sentences. Each sentence may be searched for appropriate terms such as jaundice. Store each patient's features in a file with one feature per line.

```
Project/Patient_1/Features.txt:
age 67
sex male
albumin 3.1
cirrhotic no
hepatitis_c no
...
```

2.1.2. For non-binary features, take the median value of each feature across all patients. Binarize each feature as a true(1) or false(0) value based on the median value.
Project/Patient_1/Binary_Features.txt:
age_over_60 0
male_sex 1
albumin_less_than_3.5 1
presence_of_cirrhosis 0
hepatitis_c 0
...

3. Feature Extraction from Medical Images

Note: See Step 3 Supplementary Materials for Code Examples.

3.1. Download pre- and post- therapy magnetic resonance DICOM images from the hospital PACS. Store images in the corresponding patient folders.
Project/
Project/Patient_1/Pre_TACE_MRI_Pre-Contrast.dcm
Project/Patient_1/Pre_TACE_MRI_Arterial.dcm
Project/Patient_1/Post_TACE_MRI_Pre-Constrast.dcm
Project/Patient_1/Post_TACE_MRI_Arterial.dcm

3.2. Convert DICOM images into NIfTI format using the dcm2niix program. The following commands converts all .dcm images in specified folder. Repeat for all patients.
dcm2niix Project/Patient_1/
dcm2niix Project/Patient_2/

3.3. Load each NIfTI file into Python.
```
import nibabel
image = nibabel.load('Project/Patient_1/Pre_TACE_MRI_Pre-Contrast.dcm')
```

3.3.1. Canonicalize the orientation of each image. This ensures that the x, y, and z axes are identical, irrespective of the machine used to acquire the images.
```
cImage = nibabel.as_closest_canonical(image)
```

3.4. Use itk-SNAP (or an equivalent software package) to segment binary liver and tumor masks for each image.
Project/Patient_1/Pre_TACE_MRI_Pre-Contrast_Liver_Mask.bin
Project/Patient_1/Pre_TACE_MRI_Pre-Contrast_Tumor_Mask.bin

3.5. Read the liver and tumor masks into Python. The code below demonstrates how to correct orientation issues in order to orient the masks along the same canonical axes as the MR images.
```
import numpy as np
```

```
with open(liver_mask_file, 'rb') as f:
 liver_mask = f.read()
 liver_mask = np.fromstring(liver_mask, dtype='uint8')
 liver_mask = np.reshape(liver_mask, diff.shape, order='F')
 liver_mask = liver_mask[:,::-1,:]
 liver_mask[liver_mask > 0] = 1
```

3.5.1. Use the liver and tumor masks to isolate voxels containing liver and tumor.

```
liver = np.copy(cImage)
liver[liver_mask <= 0] = 0
```

3.6. Calculate mean liver enhancement feature.

```
mean_liver_enhancement = mean(liver)
```

3.6.1. Calculate liver volume feature.

```
pixdim = cImage.header['pixdim']
units = pre.header['xyzt_units']
dx, dy, dz = pre_pixdim[1:4]
liver_volume = length(liver) * dx * dx * dz
```

3.6.2. (Optional) Calculate additional features as desired.

3.7. Update patient-specific features file with the image features.

Project/Patient_1/Features.txt:

```
age 67
sex male
albumin 3.1
cirrhotic no
hepatitis_c no
pre_tace_mean_liver_enhancement 78
pre_tace_liver_volume 10000
```

3.7.1. Calculate median values for each imaging feature and binarize as in Step 2.2.2.

Project/Patient_1/Binary_Features.txt:

```
age_over_60 0
male_sex 1
albumin_less_than_3.5 1
presence_of_cirrhosis 0
hepatitis_c 0
pre_tace_mean_liver_enhancement 1
pre_tace_liver_volume 0
```

4. **Feature Aggregation and Reduction**

Note: See Step 4 Supplementary Materials for Code Examples.

4.1. Combine the Binary_Features.txt files for each patient into a spreadsheet with patients on the y-axis and features on the x-axis.

Patient	Age > 60	Male Sex	Albumin < 3.5	Presence of Cirrhosis	Hepatitis C Present	mean liver enhancement > 50	liver volume > 20000
1	0	1	1	0	0	1	0
2	1	1	1	0	0	0	0
3	0	1	1	0	1	0	0

4.1.1. Add qEASL outcome response labels as the final column.

Patient	Age > 60	Male Sex	Albumin < 3.5	Presence of Cirrhosis	Hepatitis C Present	mean liver enhancement > 50	liver volume > 20000	qEASL Responder
1	0	1	1	0	0	1	0	1
2	1	1	1	0	0	0	0	1
3	0	1	1	0	1	0	0	0

4.1.2. Export the spreadsheet as a tab-delimited file.

Project/ML_Matrix.tsv:

Patient	Age > 60	Male Sex	Albumin < 3.5	Presence of Cirrhosis	Hepatitis C Present	mean liver enhancement > 50	liver volume > 20000	qEASL Responder
1	0	1	1	0	0	1	0	1
2	1	1	1	0	0	0	0	1
3	0	1	1	0	1	0	0	0

4.2. Remove low-variance features from consideration.

```
import numpy as np
from sklearn.feature_selection import VarianceThreshold

# Read in the binary matrix.
features = []
labels = []
for i, L in enumerate(sys.stdin):
 if i == 0:
  continue
 n_fs_L = L.strip().split('\t')
 features.append([float(_) for _ in n_fs_L[1:-1]])
 labels.append(n_fs_L[-1])
X = np.array(features)
```

```
y = np.array(labels)

# Compute features appearing in at least 20% of both responders and non-respnders.
model = VarianceThreshold(threshold=0.8 * (1 - 0.8))
X_new = model.fit_transform(X, y)
```

The male sex, albumin < 3.5, presence of cirrhosis, and liver volume > 2000 features have been removed.

Patient	Age > 60	Hepatitis C Present	mean liver enhancement > 50	qEASL Responder
1	0	0	1	1
2	1	0	0	1
3	0	1	0	0

4.3. Remove features with low univariate-association with the outcome. Filter only those features that passed 4.2. Retain ceil($\log_2(N)$) features, where N is the number patients. Ceil($\text{Log}_2(3)$) = 2.

```
import math
from sklearn.feature_selection import SelectKBest
from sklearn.feature_selection import chi2

# Read in the binary matrix as in 4.2.1
...

# Compute top ceil(log2(N)) features by univariate association.
k = math.ceil(log2(length(y)))
model = SelectKBest(chi2, k=k)
X_new = model.fit_transform(X, y)
```

The male sex age > 60 feature has been removed from the remaining features from 4.2.1.

Patient	Hepatitis C Present	mean liver enhancement > 50	qEASL Responder
1	0	1	1
2	0	0	1
3	1	0	0

5. Model Training and Testing

See Step 5 Supplementary Materials for Code Examples

5.1. Train a logistic regression model using the binary features matrix from 4.3.

```
import math
from sklearn.linear_model import LogisticRegression

# Read in the binary matrix as in 4.2 and 4.3.
...

# For each patient, train a model on all other patients.
score = 0.0
models = []
for patient in len(X):
# Train model on all but one of the patients.
train_x = np.array([_ for i, _ in enumerate(X) if i != patient])
train_y = np.array([_ for i, _ in enumerate(y) if i != patient])
model = LogisticRegression(C=1e15)
model.fit(train_x, train_y)

# Test on the left-out patient.
y_prediction = model.predict(X[patient])
if y_prediction == y[patient]:
                    score += 1
        models.append(model)
```

5.2. Train a random forest model using the binary features matrix from 4.2.2. Steps are identical to 5.2.1, except the model instantiation should be updated as follows:

```
from sklearn.ensemble import RandomForestClassifier
...
model = RandomForestClassifier(n_estimators=100)
...
```

5.3. Print out score / len(X)for 5.1 and 5.2. This represents the average accuracy of all logistic regression models and all random forest models, respectively. All *N* models should be applied to new patients with the average classification taken as the prediction outcome.

REPRESENTATIVE RESULTS:

The proposed method was applied to 36 patients who had undergone trans-arterial therapies for hepatocellular carcinoma. Twenty-five features were identified and binarized using steps 1-5. Five features satisfied both the variance and univariate association filters (see steps 5.1 and 5.2) and were used for model training. Each patient was labeled as either a responder or non-responder under the qEASL response criteria. The features matrix was thus a 36 x 5 array while the target labels vector was 36 x 1.

Logistic regression and random forest classifiers were used for model fitting. Leave-one-out cross-validation was used to assess the performance of the resulting models. Two additional models were trained using just the top-two features (presence of cirrhosis and pre-TACE tumor signal intensity greater than 27.0). **Figure 1** illustrates the performance of the models as features were added. Both logistic regression and random forest models predicted trans-arterial chemoembolization treatment response with an overall accuracy of 78% (sensitivity 62.5%, specificity 82.1%, positive predictive value 50.0%, negative predictive value 88.5%).

FIGURE AND TABLE LEGENDS:

Figure 1: Performance of machine learning algorithms. (a,b) Logistic regression and (c,d) random forest classifier accuracies as features are added. Features were added in the following order: 1) ethiodized oil, 2) sorafenib, 3) cirrhosis, 4) pre–transarterial chemo- embolization relative tumor signal intensity >27.0, and 5) number of tumors >2. Figure used with permission in unmodified form from the *Journal of Vascular and Interventional Radiology*[14].

DISCUSSION:

Patients with hepatocellular carcinoma who are not candidates for surgical resection are offered intra-arterial therapies. Few methods exist to determine if a patient will respond *pre*-treatment. Post-treatment evaluation techniques rely upon changes in tumor size or tumor contrast uptake. These are called response criteria, with the most accurate being the Quantitative European Association for the Study of the Liver (qEASL) criterion. qEASL relies upon both volumetric and enhancement changes following therapy to predict a likelihood of response. Despite the strengths of qEASL, it is nonetheless a post-treatment evaluation criteria, and cannot aid in treatment planning.

There is a need to assess which patients are likely to respond to intra-arterial therapies before performing the intervention. The method demonstrated in this protocol incorporates clinical, laboratory and imaging features into a predictive model using techniques from the fields of computer science and statistics. A machine learning model is trained that maps these features from patients who have undergone intra-arterial therapies to their qEASL outcomes. The model may then be applied to new patients who will undergo treatment to predict their qEASL response using only their pre-treatment features.

Step 1 describes workstation setup for machine learning. It provides brief instructions on how to setup a workstation with required tooling. Steps 2-4 go into technical detail about how clinical and imaging data might be parsed to obtain features of interest. These steps are critical as selection of appropriate features will determine the effectiveness of the model. Certain substeps were chosen to ease feature extraction. For example, radiographic images are typically stored in DICOM format which is not ideal for image analysis. The National Institutes of Health (NIH) Neuroimaging Branch developed the Neuroimaging Informatics Technology Initiative (NIfTI) standard to facilitate image manipulation in research environments. Step 3.2 entails the conversion from DICOM to NIfTI format to ease the process of feature extraction.

Liver and tumor masks may be extracted from NIfTI format using a program such as itk-SNAP or similar segmentation software.

Each feature, whether acquired from clinical notes or imaging data, should be binarized as true-false values. For example, a continuous image enhancement gradient ranging from 0 to 10 may be binarized to a single feature representing enhancement greater than or less than 5. Alternatively, the feature could be split into multiple binary features: $x < 3$; $3 <= x < 7$; $7 <= x$. Machine learning models that operate on binary features are easier to train.

The end-result of Steps 2-4 is a binary matrix with patients on the y-axis, features on the x-axis, and a final column representing the outcome (responder or non-responder) as determined for that patient under the qEASL response criterion. Certain features may be under-represented or over-represented in an outcome population. For example, if all treatment responders were male, it may incorrectly be concluded that male gender implies response. One way to deal with this fallacy is to remove all features that are not found in both responders and non-responders above some threshold such as 20%.

Other features may have limited importance in determining the outcome under study. For example, eye color is likely irrelevant to outcomes following intra-arterial therapies. Such features will have a low univariate association with the outcome. Although no single feature is expected to have a significant ($p < 0.05$) association with the outcome, an effective strategy is to require that features have some minimal univariate correlation above a defined threshold.

Step 5 covers the process of training and applying a machine learning model. It is not strictly necessary to proceed in the fashion described, as long as the end-result is the same. The training process uses leave-one-out cross-validation whereby *N* models are trained for each of the *N* patients. The *N* resulting models generated from leave-one-out cross-validation may be averaged to produce a final model.

The technique is also applicable to other procedures done under imaging. Machine learning takes features from a set of training patients and weights them according to their relative contribution to a target outcome. The weighting process depends on the chosen model; logistic regression models calculate an exponentiated linear combination while random forest models employ a set of weighted decision trees. The target used in this protocol was response under the qEASL criterion. Other targets, such as years of disease-free survival or quality of life years, may be chosen depending on the desired outcome under question.

The main limitation of the protocol is the need to manually define and obtain the features under consideration. This entails a significant amount of manual work such as parsing clinical notes, segmenting tumor volumes and counting the number of lesions. Deep machine learning attempts to automatically derive features from the raw source data, but requires significantly more training data. This growing technology has been shown to be superior to supervised learning in a variety of contexts, and will likely be the next evolution of predictive models for patients undergoing image-guided procedures.

ACKNOWLEDGMENTS:

A.A. received funding support from the Office of Student Research, Yale School of Medicine.

L.J.S. receives grants from the National Institutes of Health (NIH/NCI R01CA206180), Leopoldina Postdoctoral Fellowship, and the Rolf W. Guenther Foundation of Radiological Sciences (Aachen, Germany).

J.C. receives grants from the National Institutes of Health (NIH/NCI R01CA206180), Philips Healthcare, and the German-Israeli Foundation for Scientific Research and Development (Jerusalem, Israel and Neuherberg, Germany); and scholarships from the Rolf W. Guenther Foundation of Radiological Sciences and the Charite Berlin Institute of Health Clinical Scientist Program (Berlin, Germany).

J.S.D. and M.L. receive grants from the National Institutes of Health (NIH/NCI R01CA206180) and Philips Healthcare (Best, The Netherlands).

J.F.G. receives grants from the National Institutes of Health (NIH/NCI R01CA206180), Philips Healthcare, BTG (London, United Kingdom), Boston Scientific (Marlborough, Massachusetts), and Guerbet Healthcare (Villepinte, France)

DISCLOSURES:

A.A. works as a software consult for Health Fidelity, Inc. that employs similar machine learning techniques on clinical notes for optimizing medical reimbursement.

J.F.G. receives personal fees from Guerbet Healthcare, BTG, Threshold Pharmaceuticals (San Francisco, California), Boston Scientific, and Terumo (Elkton, Maryland); and has a paid consultancy for Prescience Labs (Westport, Connecticut).

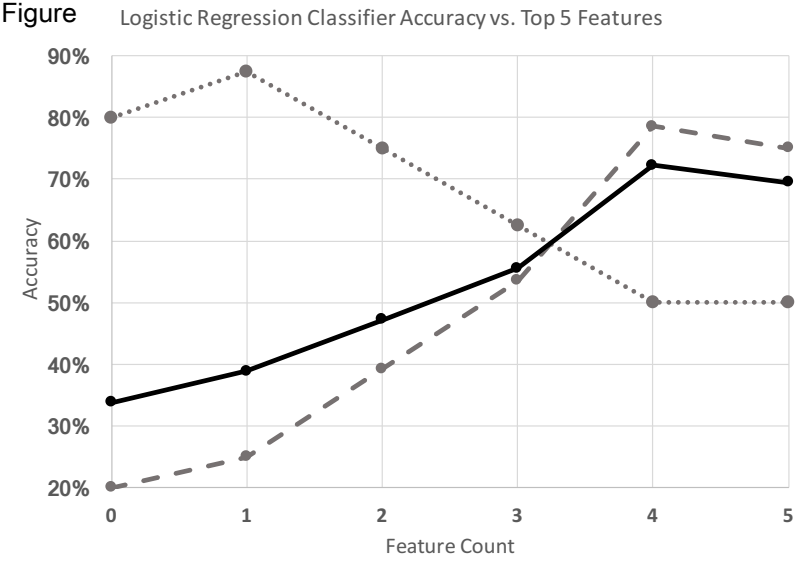
None of the other authors have identified a conflict of interest.

REFERENCES:

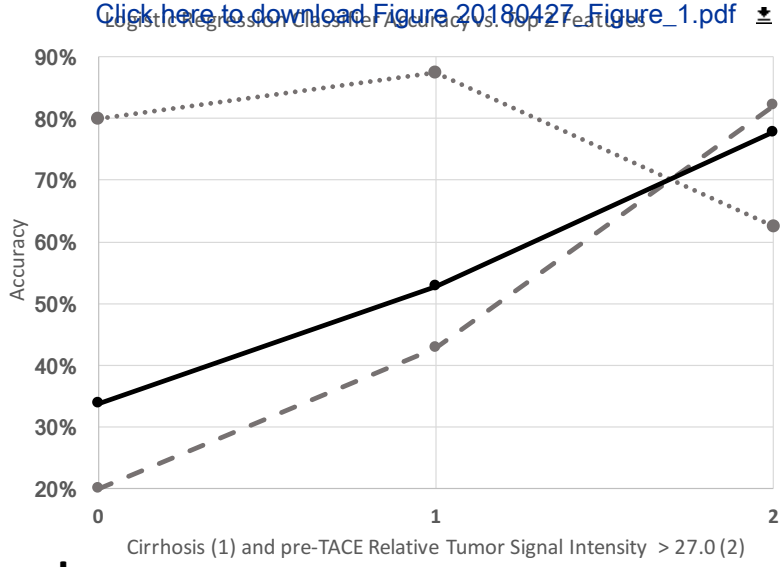
1. Benson, A., *et al.* NCCN clinical practice guidelines in oncology: hepatobiliary cancers. *J National Comprehensive Cancer Network* **7** (4), 350-391, (2009).
2. Siegel, R., Miller, K., Jemal, A. Cancer statistics, 2016. *CA Cancer J Clin*. **66** (1), 7-30, (2016).
3. Bruix, J, *et al.* Clinical management of hepatocellular carcinoma. Conclusions of the Barcelona-2000 European Association for the Study of the Liver conference. *Journal of Hepatology*. **35** (3), 421-430, (2001).
4. Eisenhauer E., *et al.* New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *European Journal of Cancer*. 45 (2), 228-247, (2009).
5. Gillmore, R., *et al.* EASL and mRECIST responses are independent prognostic factors for survival in hepatocellular cancer patients treated with transarterial embolization. *Journal of Hepatology*. **55** (6), 1309-1316, (2011).

6. Lin, M., *et al.* Quantitative and volumetric European Association for the Study of the Liver and Response Evaluation Criteria in Solid Tumors measurements: feasibility of a semiautomated software method to assess tumor response after transcatheter arterial chemoembolization. *Journal of Vascular and Interventional Radiology***23** (12), 1629-1637, (2012).
7. Tacher, V., *et al.* Comparison of Existing Response Criteria in Patients with Hepatocellular Carcinoma Treated with Transarterial Chemoembolization Using a 3D Quantitative Approach. *Radiology*. **278** (1), 275-284, (2016).
8. Pedregosa, F., *et al.* Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*. **12**, 2825-2830, (2011).
9. Bishop, C. Pattern recognition and machine learning. New York: Springer. 738 p.p. (2006).
10. Alpaydin, E. Introduction to machine learning. Third edition. ed. Cambridge, Massachusetts: The MIT Press. 613 p.p. (214).
11. Kim, S., Cho, K., Oh, S. Development of machine learning models for diagnosis of glaucoma. *PLoS One*. **12** (5), (2017).
12. Son, Y., Kim, H., Kim, E., Choi, S., Lee, S. Application of support vector machine for prediction of medication adherence in heart failure patients. *Healthcare Informatics Research*. **16** (4), 253-259, (2010).
13. Wang S., Summers, R. Machine learning and radiology. *Medical Image Analysis*. **16** (5), 933-951, (2012).
14. Abajian, A., *et al.* Predicting Treatment Response to Intra-arterial Therapies for Hepatocellular Carcinoma with the Use of Supervised Machine Learning—An Artificial Intelligence Concept. *Journal of Vascular and Interventional Radiology*. https://doi.org/10.1016/j.jvir.2018.01.769. (2018).

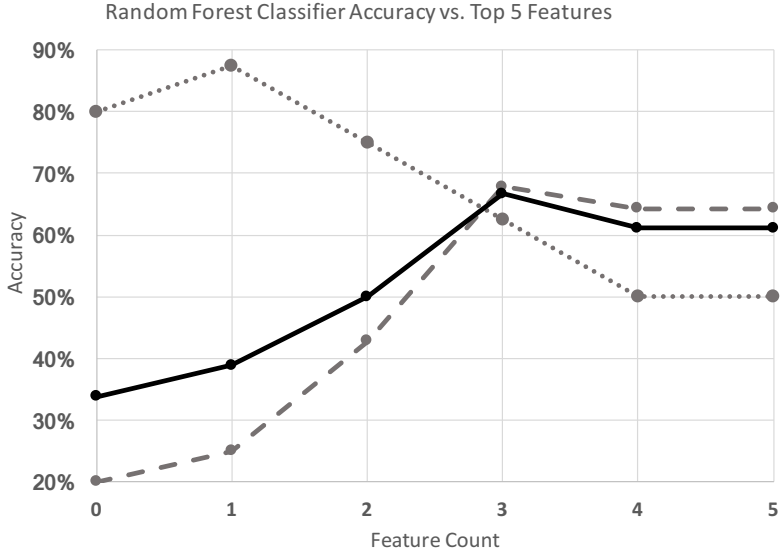
Figure



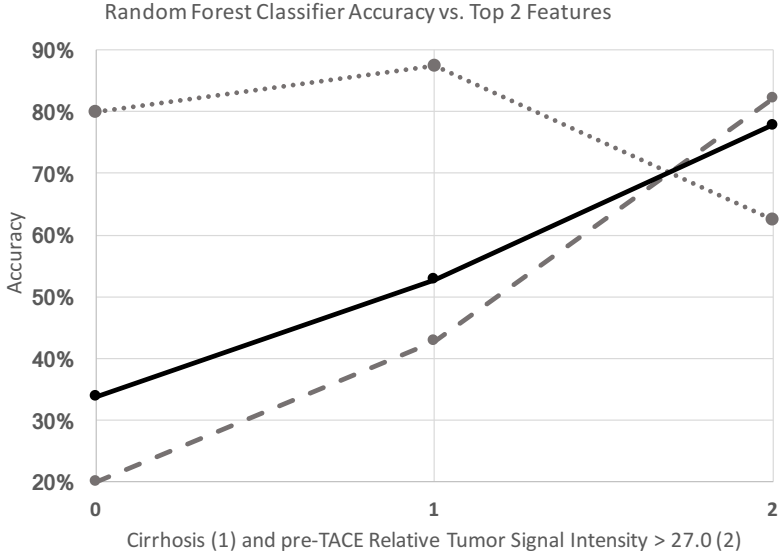
a



b



c



d

Responders Non-responders Overall

Materials Table

Click here to download Table of Materials
20180427_JoVE_Materials.xlsx

Name of Material/ Equipment	Company	Catalog Number
Computer workstation	N/A	N/A
Anaconda Python 3	Anaconda, Inc.	Version 3.6
DICOM to NIfTI	NeuroImaging Tools & Resources	Version 1.0 (4/4/2018 release)
Sublime Text Editor	Sublime HQ Pty Ltd	Version 3 (Build 3143)
Required Python Libraries	N/A	Version 3.2.25 (nltk) Version 0.19.1 (scikit-learn)
ITK-SNAP	N/A	Version 3.6.0

Comments/Description

Intel Core 2 Duo or higher CPU at 2.0 GHz; 4 GB or more system memory; POSIX-compliant operating system (Linux or Mac OS) or Microsoft Windows 7; User permissions for executing programs and saving files

Python 3 system and libraries packaged for scientists and researchers

Standalone program for converting DICOM imaging files to NIfTI format

Text-editor for writing Python code

Natural Language Toolkit (nltk)

Scikit-learn

Optional toolkit for performing segmentation of organ systems in medical images.

jove JOURNAL OF VISUALIZED EXPERIMENTS
1 Alewife Center #200
Cambridge, MA 02140
tel. 617.945.9051
www.jove.com

ARTICLE AND VIDEO LICENSE AGREEMENT

Title of Article:	Predicting treatment response to image-guided therapies using machine learning: An example for trans-arterial treatment of hepatocellular carcinoma
Author(s):	aaron abajian, nikhitha murali, lynn savic, fabian laage-gaupp, nariman nezami, james duncan, todd schlachter, mingde lin, jeff geschwind, julius chapiro

Item 1: The Author elects to have the Materials be made available (as described at http://www.jove.com/publish) via:

[X] Standard Access [ ] Open Access

Item 2: Please select one of the following items:

[X] The Author is **NOT** a United States government employee.

[ ] The Author is a United States government employee and the Materials were prepared in the course of his or her duties as a United States government employee.

[ ] The Author is a United States government employee but the Materials were NOT prepared in the course of his or her duties as a United States government employee.

ARTICLE AND VIDEO LICENSE AGREEMENT

1. **Defined Terms.** As used in this Article and Video License Agreement, the following terms shall have the following meanings: "**Agreement**" means this Article and Video License Agreement; "**Article**" means the article specified on the last page of this Agreement, including any associated materials such as texts, figures, tables, artwork, abstracts, or summaries contained therein; "**Author**" means the author who is a signatory to this Agreement; "**Collective Work**" means a work, such as a periodical issue, anthology or encyclopedia, in which the Materials in their entirety in unmodified form, along with a number of other contributions, constituting separate and independent works in themselves, are assembled into a collective whole; "**CRC License**" means the Creative Commons Attribution-Non Commercial-No Derivs 3.0 Unported Agreement, the terms and conditions of which can be found at: http://creativecommons.org/licenses/by-nc-nd/3.0/legalcode; "**Derivative Work**" means a work based upon the Materials or upon the Materials and other pre-existing works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which the Materials may be recast, transformed, or adapted; "**Institution**" means the institution, listed on the last page of this Agreement, by which the Author was employed at the time of the creation of the Materials; "**JoVE**" means MyJove Corporation, a Massachusetts corporation and the publisher of The Journal of Visualized Experiments; "**Materials**" means the Article and / or the Video; "**Parties**" means the Author and JoVE; "**Video**" means any video(s) made by the Author, alone or in conjunction with any other parties, or by JoVE or its affiliates or agents, individually or in collaboration with the Author or any other parties, incorporating all or any portion of the Article, and in which the Author may or may not appear.

2. **Background.** The Author, who is the author of the Article, in order to ensure the dissemination and protection of the Article, desires to have the JoVE publish the Article and create and transmit videos based on the Article. In furtherance of such goals, the Parties desire to memorialize in this Agreement the respective rights of each Party in and to the Article and the Video.

3. **Grant of Rights in Article.** In consideration of JoVE agreeing to publish the Article, the Author hereby grants to JoVE, subject to **Sections 4** and **7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Article in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Article into other languages, create adaptations, summaries or extracts of the Article or other Derivative Works (including, without limitation, the Video) or Collective Works based on all or any portion of the Article and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and(c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. If the "Open Access" box has been checked in **Item 1** above, JoVE and the Author hereby grant to the public all such rights in the Article as provided in, but subject to all limitations and requirements set forth in, the CRC License.

4. **Retention of Rights in Article.** Notwithstanding the exclusive license granted to JoVE in **Section 3** above, the Author shall, with respect to the Article, retain the non-exclusive right to use all or part of the Article for the non-commercial purpose of giving lectures, presentations or teaching classes, and to post a copy of the Article on the Institution's website or the Author's personal website, in each case provided that a link to the Article on the JoVE website is provided and notice of JoVE's copyright in the Article is included. All non-copyright intellectual property rights in and to the Article, such as patent rights, shall remain with the Author.

5. **Grant of Rights in Video – Standard Access.** This **Section 5** applies if the "Standard Access" box has been checked in **Item 1** above or if no box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby acknowledges and agrees that, Subject to **Section 7** below, JoVE is and shall be the sole and exclusive owner of all rights of any nature, including, without limitation, all copyrights, in and to the Video. To the extent that, by law, the Author is deemed, now or at any time in the future, to have any rights of any nature in or to the Video, the Author hereby disclaims all such rights and transfers all such rights to JoVE.

6. **Grant of Rights in Video – Open Access.** This **Section 6** applies only if the "Open Access" box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby grants to JoVE, subject to **Section 7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Video in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Video into other languages, create adaptations, summaries or extracts of the Video or other Derivative Works or Collective Works based on all or any portion of the Video and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and (c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. For any Video to which this **Section 6** is applicable, JoVE and the Author hereby grant to the public all such rights in the Video as provided in, but subject to all limitations and requirements set forth in, the CRC License.

7. **Government Employees.** If the Author is a United States government employee and the Article was prepared in the course of his or her duties as a United States government employee, as indicated in **Item 2** above, and any of the licenses or grants granted by the Author hereunder exceed the scope of the 17 U.S.C. 403, then the rights granted hereunder shall be limited to the maximum rights permitted under such statute. In such case, all provisions contained herein that are not in conflict with such statute shall remain in full force and effect, and all provisions contained herein that do so conflict shall be deemed to be amended so as to provide to JoVE the maximum rights permissible within such statute.

8. **Protection of the Work.** The Author(s) authorize JoVE to take steps in the Author(s) name and on their behalf if JoVE believes some third party could be infringing or might infringe the copyright of either the Author's Article and/or Video.

9. **Likeness, Privacy, Personality.** The Author hereby grants JoVE the right to use the Author's name, voice, likeness, picture, photograph, image, biography and performance in any way, commercial or otherwise, in connection with the Materials and the sale, promotion and distribution thereof. The Author hereby waives any and all rights he or she may have, relating to his or her appearance in the Video or otherwise relating to the Materials, under all applicable privacy, likeness, personality or similar laws.

10. **Author Warranties.** The Author represents and warrants that the Article is original, that it has not been published, that the copyright interest is owned by the Author (or, if more than one author is listed at the beginning of this Agreement, by such authors collectively) and has not been assigned, licensed, or otherwise transferred to any other party. The Author represents and warrants that the author(s) listed at the top of this Agreement are the only authors of the Materials. If more than one author is listed at the top of this Agreement and if any such author has not entered into a separate Article and Video License Agreement with JoVE relating to the Materials, the Author represents and warrants that the Author has been authorized by each of the other such authors to execute this Agreement on his or her behalf and to bind him or her with respect to the terms of this Agreement as if each of them had been a party hereto as an Author. The Author warrants that the use, reproduction, distribution, public or private performance or display, and/or modification of all or any portion of the Materials does not and will not violate, infringe and/or misappropriate the patent, trademark, intellectual property or other rights of any third party. The Author represents and warrants that it has and will continue to comply with all government, institutional and other regulations, including, without limitation all institutional, laboratory, hospital, ethical, human and animal treatment, privacy, and all other rules, regulations, laws, procedures or guidelines, applicable to the Materials, and that all research involving human and animal subjects has been approved by the Author's relevant institutional review board.

11. **JoVE Discretion.** If the Author requests the assistance of JoVE in producing the Video in the Author's facility, the Author shall ensure that the presence of JoVE employees, agents or independent contractors is in accordance with the relevant regulations of the Author's institution. If more than one author is listed at the beginning of this Agreement, JoVE may, in its sole

jove JOURNAL OF VISUALIZED EXPERIMENTS
1 Alewife Center #200
Cambridge, MA 02140
tel. 617.945.9051
www.jove.com

ARTICLE AND VIDEO LICENSE AGREEMENT

discretion, elect not take any action with respect to the Article until such time as it has received complete, executed Article and Video License Agreements from each such author. JoVE reserves the right, in its absolute and sole discretion and without giving any reason therefore, to accept or decline any work submitted to JoVE. JoVE and its employees, agents and independent contractors shall have full, unfettered access to the facilities of the Author or of the Author's institution as necessary to make the Video, whether actually published or not. JoVE has sole discretion as to the method of making and publishing the Materials, including, without limitation, to all decisions regarding editing, lighting, filming, timing of publication, if any, length, quality, content and the like.

12. **Indemnification.** The Author agrees to indemnify JoVE and/or its successors and assigns from and against any and all claims, costs, and expenses, including attorney's fees, arising out of any breach of any warranty or other representations contained herein. The Author further agrees to indemnify and hold harmless JoVE from and against any and all claims, costs, and expenses, including attorney's fees, resulting from the breach by the Author of any representation or warranty contained herein or from allegations or instances of violation of intellectual property rights, damage to the Author's or the Author's institution's facilities, fraud, libel, defamation, research, equipment, experiments, property damage, personal injury, violations of institutional, laboratory, hospital, ethical, human and animal treatment, privacy or other rules, regulations, laws, procedures or guidelines, liabilities and other losses or damages related in any way to the submission of work to JoVE, making of videos by JoVE, or publication in JoVE or elsewhere by JoVE. The Author shall be responsible for, and shall hold JoVE harmless from, damages caused by lack of sterilization, lack of cleanliness or by contamination due to the making of a video by JoVE its employees, agents or independent contractors. All sterilization, cleanliness or decontamination procedures shall be solely the responsibility of the Author and shall be undertaken at the Author's expense. All indemnifications provided herein shall include JoVE's attorney's fees and costs related to said losses or damages. Such indemnification and holding harmless shall include such losses or damages incurred by, or in connection with, acts or omissions of JoVE, its employees, agents or independent contractors.

13. **Fees.** To cover the cost incurred for publication, JoVE must receive payment before production and publication the Materials. Payment is due in 21 days of invoice. Should the Materials not be published due to an editorial or production decision, these funds will be returned to the Author. Withdrawal by the Author of any submitted Materials after final peer review approval will result in a US$1,200 fee to cover pre-production expenses incurred by JoVE. If payment is not received by the completion of filming, production and publication of the Materials will be suspended until payment is received.

14. **Transfer, Governing Law.** This Agreement may be assigned by JoVE and shall inure to the benefits of any of JoVE's successors and assignees. This Agreement shall be governed and construed by the internal laws of the Commonwealth of Massachusetts without giving effect to any conflict of law provision thereunder. This Agreement may be executed in counterparts, each of which shall be deemed an original, but all of which together shall be deemed to me one and the same agreement. A signed copy of this Agreement delivered by facsimile, e-mail or other means of electronic transmission shall be deemed to have the same legal effect as delivery of an original signed copy of this Agreement.

A signed copy of this document must be sent with all new submissions. Only one Agreement is required per submission.

CORRESPONDING AUTHOR

Name:	Julius Chapiro
Department:	Radiology and Biomedical Imaging
Institution:	Yale University
Title:	Research Scientist and Resident Physician
Signature:	
Date:	7/4/2018

Please submit a **signed** and **dated** copy of this license by one of the following three methods:

1. Upload an electronic version on the JoVE submission site
2. Fax the document to +1.866.381.2236
3. Mail the document to JoVE / Attn: JoVE Editorial / 1 Alewife Center #200 / Cambridge, MA 02140

Editorial comments:

Changes to be made by the Author(s):

1. Please take this opportunity to thoroughly proofread the manuscript to ensure that there are no spelling or grammar issues. The JoVE editor will not copy-edit your manuscript and any errors in the submitted revision may be present in the published version.
We have gone through and proofread the spelling and grammar.

2. Please verify the desired access type. The Author License Agreement states open access but Editorial Manager states standard access.
We would like standard access.

3. Please obtain explicit copyright permission to reuse any figures from a previous publication. Explicit permission can be expressed in the form of a letter from the editor or a link to the editorial policy that allows re-prints. Please upload this information as a .doc or .docx file to your Editorial Manager account. The Figure must be cited appropriately in the Figure Legend, i.e. "This figure has been modified from [citation]."
We have emailed JVIR and formally requested the right to use the figure, their response is as follows: "Please note that, as one of the Authors of this article, you retain the right to reuse portion or excerpts in a new work. You do not require formal permission to do so. For full details of your rights as a Journal Author, please visit: https://www.elsevier.com/about/our-business/policies/copyright"

4. Please avoid the use of lists in the Abstract. Please maintain the paragraph format.
Abstract has been rewritten to remove the list.

5. For in-text formatting, corresponding reference numbers should appear as numbered superscripts after the appropriate statement(s).
We have converted reference numbers to superscripts.

6. Please include Table 1 as an editable xls/xlsx file.
We have decided to remove Table 1 as it provides little standalone information. Rather, we have walked through the process of building this table in the protocol.

7. Please ensure that all text in the protocol section is written in the imperative tense as if telling someone how to do the technique (e.g., "Do this," "Ensure that," etc.). The actions should be described in the imperative tense in complete sentences wherever possible. Avoid usage of phrases such as "could be," "should be," and "would be" throughout the Protocol. Any text that cannot be written in the imperative tense may be added as a "Note." However, notes should be concise and used sparingly. Please include all safety procedures and use of hoods, etc.
We have revised the protocol to use the imperative mood. We have also reduced the use of text not in the imperative mode. Section 2 is almost entirely instructive now rather than descriptive.

8. The Protocol should contain only action items that direct the reader to do something. Please move the discussion about the protocol to the Discussion.
The commentary has been moved to the discussion section.

9. Please add more details to your protocol steps. Please ensure you answer the "how" question, i.e., how is the step performed? Alternatively, add references to published material specifying how to perform the protocol action.
We have updated the protocol to have step-by-step instructions regarding training and testing a model.

10. Please ensure that all terminal commands are explicitly written out.
We have written out terminal commands, provided code snippets, and provided an entire code file for generating features from images.

11. Steps 2-5 do not have enough details to replicate. We need explicit user input commands in order to film: File | Save | etc. Can a visual/screenshot be provided? A specific protocol (with specific values and numbers) in a specific example would help greatly.
We have updated these steps to have specific instructions.

12. Please highlight 2.75 pages or less of the Protocol (including headings and spacing) that identifies the essential steps of the protocol for the video, i.e., the steps that should be visualized to tell the most cohesive story of the Protocol. Remember that non-highlighted Protocol steps will remain in the manuscript, and therefore will still be available to the reader.
We have highlighted Steps 2.2.1, 2.2.2, 3.6.1, 3.6.2, 3.7.2, 4.2.1, 4.2.2, and Step 5 in it entirety.

13. Please ensure that the highlighted steps form a cohesive narrative with a logical flow from one highlighted step to the next. Please highlight complete sentences (not parts of sentences). Please ensure that the highlighted part of the step includes at least one action that is written in imperative tense.

14. As we are a methods journal, please revise the Discussion to explicitly cover the following in detail in 3-6 paragraphs with citations:
a) Critical steps within the protocol:
We have noted the importance of choosing appropriate features.

b) Any modifications and troubleshooting of the technique.
We have noted that an segmentation software may be used in lieu of itk-SNAP.

c) Any limitations of the technique
We have noted that the method requires feature selection and that deep learning is likely a better approach, but requires more data.

d) The significance with respect to existing methods
We have noted that existing methods only consider a few number of features and/or are constrained to include only-imaging or only-clinic data.

e) Any future applications of the technique
We have noted that the method is extensible beyond interventional radiology as it may be applied to any treatment modality with pre- and post- intervention clinical and imaging features.

15. Please do not abbreviate journal titles.
We have reviewed and eliminated our use of abbreviations.

Reviewer #1:

Manuscript Summary:

The paper's is devoted to the very actual topic in medical machine learning (classifiers of responders/non-responders to certain treatment methods of cancer), but the overall presentation style of the current version of the manuscript fits more for a user's manual rather than a scientific paper. To improve the comprehensiveness of the papers, a major revision is needed.

Major Concerns:

1. Lines 41-44. The short abstract is too short. It needs extension.
Abstract has been extended while staying under the 50-word limit.

2. Lines 119-152. Protocol: Workstation setup for machine learning. The style of this section resembles more a user's manual than a scientific report. Installation instructions seem redundant.
The authors agree that this section is more instructional than novel or specific to the method. However, reproduction of methods is often highly dependent on equivalent software setups. The authors feel that some instruction should be provided on how to mirror the tool setup used for the novel results. In particular, explaining that Anaconda's repositories should be use for package installation (in the Installation Instructions section) can avoid end-user headaches related to software dependences.

All actions should be describe in past indicative rather than imperative mood.
The expected mood of the article is imperative, as confirmed by the Editor.

3. Lines 161-205. Feature extraction for case histories and DICOM images. Requires a flowchart/diagram.
Flowchart added.

4. Lines 235-263. Requires a flowchart/diagram, along with a code sample.
Flowchart added and we have provided a Python script for extracting imaging features.

5. Lines 265-296. Results section. Please provide model figures on model training/validation results (e.g., correlation plots for fitted/trained and LOO-validated results etc.). Please also extend figure legends and their description in the main text. Please provide principal data used for machine learning as Supplementary materials, since this is more methodical paper than a result analysis one.
Figure 1 is the primary result graph illustrating the performance of the LOO-validated models in aggregate. We do not provide the performance of each individual model, which would merely consist of 36 accuracies validated only against one patient. We can provide this if strongly desired. We have written out individual instructions (Step 5.2.1, 5.2.2, and 5.3) that document how the reader may obtain the individual accuracies of each LOO model.

Regarding supplemental materials, we have now provided extensive code demonstrating how to complete Step 4 and Step 5.

Reviewer #2:

Manuscript Summary:

Excellent innovative idea in using standard criteria (including imaging and clinical information) to predict patient response and ultimately survival. Great start into mostly unchartered territory in IR. Overall the manuscript is concise and your methods are easy to follow even for someone not experienced in programming language.

Major Concerns:

1. Your focus on the article is clearly defined in using machine learning as a tool to initially predict a response to intra-arterial therapy, and the bulk of the article is focused on the IT protocol and programming language. However, you need a bit more background to say way in your introduction as this applies specifically to your hypothesis. mRECIST (or qEASL) and other predictors exist AFTER treatment to predict response. Make a little more of a case here for why this matters initially for IA therapy. A paragraph about IA therapy in IR and why it is so variable to begin with might be helpful in addition to your examples of use in oncology, using the MELD, etc.

We have expanded the first paragraph of the discussion to include detail about qEASL as a post-treatment response criteria.

Table 1

Patient #	Feature A	Feature B	Feature C	Target
1	0	0	1	0
2	0	1	0	0
3	0	1	0	1
4	1	1	1	0
5	1	1	1	0
6	0	1	0	1
7	0	0	0	0
8	1	0	0	1
9	0	1	0	1
10	1	0	0	0