AWARD NUMBER: W81XWH-15-1-0110

TITLE: Noninvasive Characterization of Indeterminate Pulmonary Nodules Detected on Chest High-Resolution Computed Tomography

PRINCIPAL INVESTIGATOR: Fabien Maldonado, M.D.

CONTRACTING ORGANIZATION: Vanderbilt University Medical Center
Nashville, TN 37232

REPORT DATE: October 2018

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command

Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
October 2018	Annual	09/30/2017 – 9/29/2018

4. TITLE AND SUBTITLE

Noninvasive Characterization of Indeterminate Pulmonary Nodules Detected on Chest High-Resolution Computed Tomography

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-15-1-0110

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Fabien Maldonado, M.D.

E-Mail: Fabien.Maldonado@vumc.org

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Vanderbilt University Medical Center
1161 21st Ave S, Ste D3300 MCN
Nashville, TN 37232-0011

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The widespread implementation of lung cancer screening, following favorable results of the National Lung Screening Trial (NLST) and more recently the European NELSON trial, will likely continue to exacerbate the widespread clinical problem of indeterminate pulmonary nodules, which were detected in 40% of high-risk individuals screened by low dose high-resolution computed tomography (HRCT) in the NLST. Because 96% of these nodules were benign, the issue of diagnostic resolution of incidentally and screen-identified lung nodules will become increasingly important. Current clinical and radiological risk prediction models, which allow risk-stratification of patients and individualize management of pulmonary nodules, are commonly used, but remain suboptimal, and optimization of the clinical management of larger (≥ 7 mm) screen-detected nodules is urgently needed to avoid unnecessary diagnostic interventions leading to unwarranted mortality, morbidity and healthcare costs. In our project, we explore the utility of a conventional radiomic approach to the classification of screen-detected indeterminate nodules, leveraging unexploited large datasets contained on digital HRCT images to estimate the probability of malignancy based on selected predictive quantitative radiologic features.

15. SUBJECT TERMS
Lung adenocarcinoma, Radiomics, Lung cancer screening, Chest computed tomography, Biomarkers, Lung nodules.

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	18	19b. TELEPHONE NUMBER *(include area code)*

Table of Contents

Page

1. **INTRODUCTION:**

Lung cancer affects 220,000 individuals every year in the US, accounting for more deaths than the next three most common cancers combined (breast, colon and prostate), resulting in approximately160,000 deaths per year. Early detection is key to improving outcomes, as treatment for advanced stages is of limited benefit, with a disappointingly low survival rate of 16% at 5 years overall. The National Lung Screening Trial (NLST), a large randomized controlled trial on lung cancer screening, demonstrated a 20% relative reduction in lung cancer mortality with annual low-dose chest computed tomography (LDCT). These were recently confirmed in a large European study, the NELSON trial, presented at the European Respiratory Society meeting in Paris last September. These encouraging results have led to widespread endorsement of lung cancer screening, but broad-scale implementation has been hampered by the considerable number of false positive LDCT, leading to many unnecessary interventions leading to excess morbidity, mortality, patient anxiety and healthcare costs. In the NLST, 40% of patients had identifiable lung nodules, 96% of which proved benign. In addition, with an estimated 20 million new chest CTs performed every year in the US, incidental lung nodules are identified in 1.5 million patients each year, which is almost certainly an underestimate. The detection of lung nodules leads to a variety of non-invasive and invasive studies to determine whether they are benign or malignant. While guidelines have outlined strategies to approach indeterminate lung nodules based on pre-test probability of malignancy, many patients with benign nodules continue to be submitted to unnecessary procedures leading to increased morbidity, mortality and healthcare costs. Novel tools to distinguish benign from malignant nodules are needed. We have previously demonstrated that volumetric CT-based quantitative imaging for lung adenocarcinoma characterization is useful in risk-stratifying these lesions, exploiting the wealth of data points available with modern CT imaging. In this project, we are using similar quantitative imaging metrics to develop a conventional radiomics-based model that will assist radiologists and clinicians in their attempt to determine the likelihood of malignant lung nodule based on LDCT imaging. We used the available NLST dataset as a training set and are planning on using the large ongoing prospective study Detection of Early lung Cancer Among Military Personnel Study 1 (DECAMP-1) for validation, in addition to other independent validation datasets as we await the results of DECAMP-1. This project will help to limit morbidity, mortality and healthcare costs associated with the management of incidentally or screen-identified pulmonary nodules.

2. **KEYWORDS:**

Lung adenocarcinoma, Radiomics, Lung cancer screening, Chest computed tomography, Biomarkers, Lung nodules.

3. **ACCOMPLISHMENTS:**

- **What were the major goals of the project?**

Aim 1 (first year of the grant): The first aim of this grant was to develop an imaging-based approach using volumetric analysis of screen-identified lung nodules, and a combined clinical-radiologic model to differentiate benign from malignant nodules.

> Milestone: Development of optimized quantitative radiological variables predictive of the benign or malignant character of lung nodules from a cohort isolated from the NLST (12 months – October 2016)

Note that subcontracts with Brown University and Mayo Clinic (required due to relocation of the PI, Fabien Maldonado, to Vanderbilt University Medical Center) were not established until March 2016 and as such work could not be started before that time.

The identification of optimization of quantitative radiological variables was completed by October 2017.

> Milestone: development of a radiologic prediction model (12 months)
>
> -The radiologic model was completed by October 2017.
>
> Milestone: development of a combined clinical/radiologic prediction model (12 months)
>
> -The clinical/radiologic model was completed by October 2017, but addition of clinical variables did not contribute significantly to the diagnostic test performance of the model.

Aim 2 (second year of the grant): the second aim of this grant is to prospectively validate the models developed in Aim 1 in the DECAMP-1 dataset (500 patients with indeterminate pulmonary nodules**,** DECAMP PROTOCOL ACRIN 4703).

> Milestone: Validation of a radiologic and combined clinical/radiologic prediction models (Year 2 of the grant).

Enrollment for the DECAMP1 study continues to be considerably delayed. Completion of enrollment in the study was anticipated by December 2015 at the time of our application (August 2014), as 125 of the planned 500 patients had already been enrolled (see attached original support letter from DECAMP1 PI Dr. Avrum Spira). As of August 2017, DECAMP-1 study had accrued and adjudicated 274 cases including 183 malignant and 91 confirmed benign nodules. An interim blind analysis of these nodules using our radiologic model yielded a disappointingly low AUC of 0.66 (strict validation) and 0.74 (loose validation). These disappointing results were felt to be due to the very large number of malignant nodules in this cohort, likely to result to early adjudication of the most suspicious lung nodules.

We were also able to secure two alternative validation sets, including an alternative validation cohort from the lung nodule registry at Vanderbilt University Medical Center/Nashville Veterans

Administration Tennessee Valley Healthcare system (primary investigator: Dr. Pierre Massion, see below) comprised of 103 malignant lung nodules and 99 benign lung nodules, as well as the lung nodule cohort from the Lung Tissue Research Consortium, comprised of 88 benign and 89 malignant nodules. Similar to the early DECAMP-1 cohort, this cohort was considered "high-risk" as all nodules were evaluated by expert radiologists and felt to be suspicious enough for malignancy to require surgical resection (see below). Note funds from the award were not used for these external validation sets.

- **What was accomplished under these goals?**

1) Major activities:

Summary of activities that occurred during the first year of the grant:

Year 1:

Nodule selection

Participants for our project were selected from the pool of eligible participants in the NLST, who did not withdraw from follow-up, in the CT arm of the study (N=26,262) and included all screen-detected lung cancer cases. Non-lung cancer controls were selected as a stratified random sample from all participants in the pool defined above who were not found to have lung cancer during the screen or follow-up periods of the NLST in a 1:1 fashion. We restricted our analysis to nodules with a size defined by a largest diameter comprised between 7 and 30 mm as reported in the NLST database, as these represent the size criteria used as eligibility criteria in the DECAMP-1 study

The CT datasets were obtained from the Lung Screening Study core laboratory and from the American College of Radiology Imaging Network core laboratory were transferred initially via hard drive, then electronically to the investigators. Information on nodule location was available to the investigators in the NLST database and confirmed by one radiologist (B.J.B.) and two pulmonologists (F.M. and T.P.) using the CT obtained the closest in time to the diagnosis of malignant or benign lung nodules. Nodules were electronically tagged for segmentation and analysis. HRCT without visible nodules, nodules with borders indistinguishable from neighboring structures (e.g. mediastinum or pleura) and nodules without related clinical data were excluded.

Nodule segmentation and analysis

The lung nodules were segmented manually using the ANALYZE software (Biomedical Imaging Resource, Mayo Clinic, Rochester, MN), as previously reported. A comprehensive set of automatically computable, quantitative radiomic metrics was included for the development of a multivariable predictive model to discriminate benign from malignant lung nodules. Based on previous data and preliminary analysis, we considered metrics within the following categories: general characteristics of the nodule (volume and location), nodule characteristics (texture and surface characteristics) and nodule-free surrounding lung characteristics, as below:

1. Bulk metrics based on the global shape descriptors of the nodule.
2. Intensity metrics based on the CT Hounsfield units within the nodule.
3. Metrics capturing the spatial location of the nodule.
4. Nodule texture metrics based on the texture exemplar distributions within the nodule.
5. Surround texture metrics based on the parenchymal texture exemplar distributions within a region surrounding the nodule.
6. Metrics capturing the surface descriptors of the nodule.
7. Metrics capturing the distribution of the surface exemplars of the nodule.

Year 2 (for additional details see prior annual report):

Multivariate analysis was performed using least absolute shrinkage and selection operator (LASSO) method for both variable selection and regularization in order to enhance the prediction accuracy and interpretability of the multivariate statistical model. To increase the stability of the modeling, LASSO was run 1,000 times and the variables that were selected by at least 50% of the runs were included into the final multivariate model. The bootstrapping method was then applied for the internal validation, and the optimism-corrected AUC was reported for the final model.

Results:

After exclusion of cases lacking HRCT data, cases with no apparent lesion on last HRCT prior to the cancer diagnosis, cases with nodules invading the mediastinum, cases with missing outcome data, and lesion with size < 7mm or >30 mm, 408 LDCT scans with malignant nodules were selected and analyzed. A stratified random sample of non-lung cancer controls (nodules with size comprised between 7 and 30 mm) was selected on a 1:1 basis, and after exclusion of HRCT containing more than one nodule, 318 nodules were selected and included in the analysis.

The demographic and clinical characteristics of individuals included in the study are summarized below:

Demographics and Clinical Characteristics of Cancer and Control (n = 726)

	Lung Cancer Cases (n=408)	Nodule-Positive Controls (n=318)	p Value
Age, mean ± SD, y	63 .7 ± 5.3	61.2 ± 5.0	<0.001
Sex, n (%)			0.45
Male	230 (56.4)	189 (59.4)	
Female	178 (43.6)	129 (40.6)	
Race, n (%)			0.03
White	385 (94.4)	286 (89.9)	
Black, Asian, other	23 (5.6)	32 (10.1)	
Ethnicity, n (%)			0.31
Hispanic or Latino	405 (98.4)	313 (99.3)	
Neither Hispanic nor Latino	3 (1.6)	5 (0.7)	
Smoking, n (%)			0.37
Current	221 (54.2)	161 (50.6)	
Former	187 (45.8)	157 (49.4)	

Pack-years smoked, mean ± SD			
Current smokers	64.8 ± 25.8	55.5 ± 20.9	<0.001
Former smokers	66.7 ± 30.6	55.2 ± 26.9	<0.001
Self-reported history of COPD, n (%)			
Yes	43 (10.5)	18 (5.7)	0.02
No	365 (89.5)	300 (94.3)	
FH of lung cancer, n (%)			0.08*
Yes	113 (28.9)	69 (22.8)	
No	278 (71.1)	233 (77.2)	
Missing	n=17	n=16	
Stage, n (%)			—
I	298 (73.0)	—	
II	29 (7.1)	—	
III	55 (13.5)	—	
IV	20 (5.0)	—	
Carcinoid, unknown	6 (1.5)	—	
Histologic subtype, n (%)			—
Adenocarcinoma	290 (71.1)	—	
Squamous cell carcinoma	81 (19.9)	—	
Other, NOS, unknown	37 (9.1)	—	

P Values calculated using Fisher's exact test for categorical variables, Student's t test for continuous variables.
* P value for family history of lung cancer was calculated without missing data.

AUC analysis across cancers and controls.

ID	Variables	Cancer_mean(SD)	Control_mean(SD)	AUC	P value
1	Centroid_x	154.78 (74.5)	142.21 (78.73)	0.56	0.02837
2	Centroid_y	143.95 (47.18)	151.84 (55.47)	0.47	0.03916
3	Centroid_Z	203.38 (60.1)	186.88 (65.91)	0.57	0.00052
4	Volume	3985.59 (13526.02)	344.48 (818.4)	0.9	0
5	SurfaceArea	1841.06 (3508.55)	344.12 (501.43)	0.87	0
6	Sphericity	0.51 (0.21)	0.6 (0.29)	0.58	1.00E-05
7	SphereFitFactor	6.82 (8.31)	5.28 (5.82)	0.58	0.00668
8	Radius_Estimated	7.61 (3.99)	3.59 (1.57)	0.9	0
9	Min.Enclosing.Brick_x	19.82 (12.12)	9.46 (5.51)	0.84	0
10	Min.Enclosing.Brick_y	19.63 (12.13)	10.11 (6.72)	0.82	0
11	Min.Enclosing.Brick	16.49 (14.51)	4.97 (2.65)	0.92	0
12	Max.Bricklength	24.08 (16.27)	11.31 (7.04)	0.84	0
13	Elongation	-0.25 (0.4)	-0.31 (0.47)	0.57	0.07783

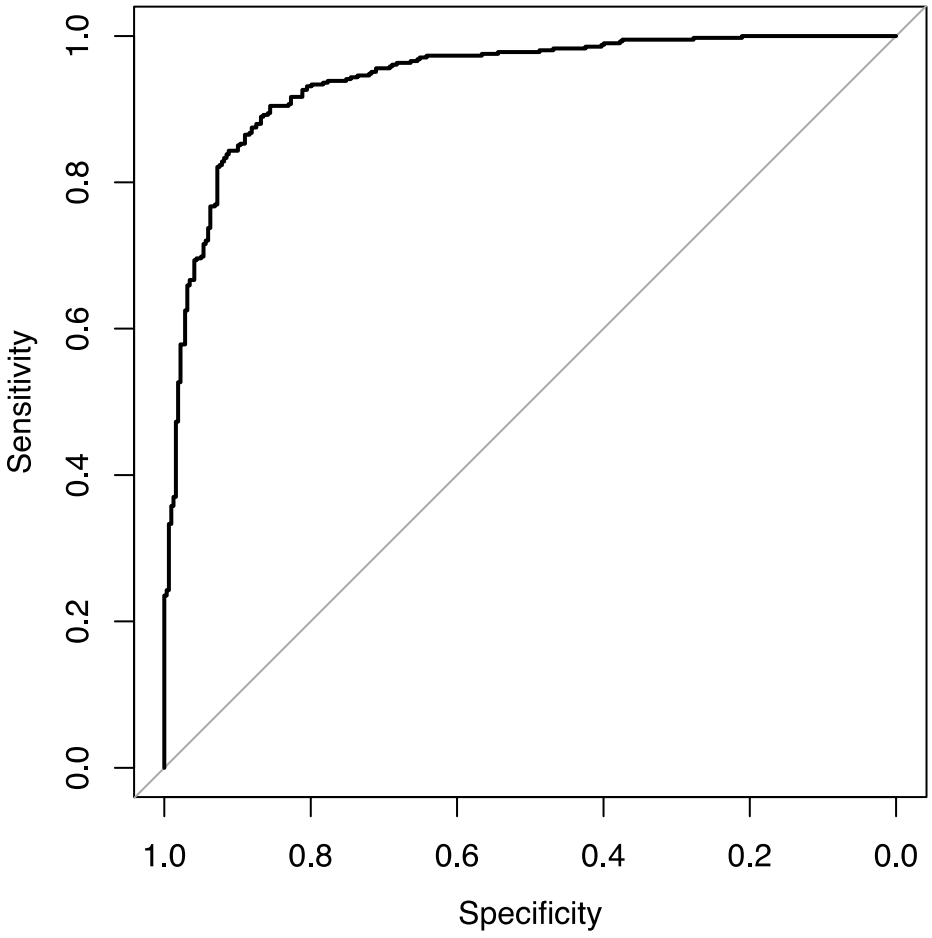
14	Flatness	-0.56 (0.99)	-1.01 (1.05)	0.66	0
15	HU_mean	-209.18 (163.55)	-465.23 (201.91)	0.83	0
16	HU_var	614546.92 (3444392.14)	295011.7 (609422.64)	0.56	0.09419
17	HU_skew	-2.64 (10.09)	-2.39 (1.2)	0.57	0.66095
18	HU_kurt	133.91 (2032.65)	10.54 (10.04)	0.74	0
19	HU_entropy	7.89 (1.77)	6.76 (1.76)	0.82	0
20	Location	6.37 (3.42)	7.06 (3.16)	0.56	0.00558
21	SILA_Tex	122.91 (34.32)	58.62 (38.1)	0.88	0
22	Tex_Risk	2.17 (0.57)	1.36 (0.54)	0.82	0
23	Ves_.	1.88 (2.8)	0.75 (1.29)	0.74	0
24	Bgnd_.	9.49 (9.56)	9.59 (11.25)	0.52	0.89459
25	SILA_Fib	32.32 (17.84)	27.42 (22.96)	0.57	0.00136
26	SILA_Laa	35.54 (16.33)	32.69 (19.86)	0.55	0.03461
27	Num.Vertices	2711.4 (4745.67)	515.25 (697.45)	0.88	0
28	Num.Faces	5419.18 (9488.83)	1026.56 (1395.09)	0.88	0
29	WBE_2	1574.75 (3792.16)	480.61 (721.39)	0.75	0
30	WBE	2269.82 (6283.03)	802.67 (1116.04)	0.7	0
31	Min_MeanCurv	-0.92 (0.65)	-0.28 (0.46)	0.82	0
32	Max_MeanCurv	3.57 (2.44)	3.27 (1.82)	0.5	0.0694
33	Avg_PosMeanCurv	0.34 (0.11)	0.58 (0.2)	0.87	0
34	Skew_PosMeanCurv	2.89 (2.04)	2.01 (1.2)	0.66	0
35	Min_GCurv	-1.01 (0.87)	-0.87 (0.84)	0.58	0.03424
36	Max_GCurv	15.43 (30.41)	12.6 (21.14)	0.51	0.16811
37	Avg_PosGCurv	0.29 (0.29)	0.61 (0.52)	0.79	0
38	Skew_PosGCurv	7.57 (3.82)	4.66 (2.09)	0.78	0
39	Min_Sharp	0 (0)	0 (0)	0.79	0
40	Max_Sharp	38.99 (62.98)	22.44 (52.57)	0.59	0.00026
41	Avg_Sharp	0.59 (0.43)	1.01 (0.78)	0.71	0
42	Skew_Sharp	7.95 (7.45)	4.25 (3.53)	0.72	0
43	Min_Curved	0.01 (0.03)	0.07 (0.1)	0.82	0
44	Max_Curved	5.72 (4.21)	4.8 (3.05)	0.53	0.00131
45	Avg_Curved	0.58 (0.19)	0.96 (0.32)	0.87	0
46	Skew_Curved	2.87 (2.26)	1.79 (1.25)	0.69	0
47	Min_SI	-0.98 (0.01)	-0.98 (0.02)	0.63	0
48	Max_SI	0.98 (0.16)	0.55 (0.61)	0.82	0
49	Avg_SI	-0.29 (0.18)	-0.55 (0.13)	0.88	0
50	Skew_SI	1.63 (0.91)	1.72 (1.42)	0.54	0.3307
51	ICI	37.78 (118.81)	15.7 (21.56)	0.64	0
52	ECI	113.69 (284.16)	39.41 (57.05)	0.73	0
53	SILA_T	36.02 (11.24)	19.71 (12.61)	0.84	0
54	AvgCrv_T1	0.74 (0.23)	1.05 (0.32)	0.81	0
55	SkewCrv_T1	2.33 (1.73)	1.57 (1.04)	0.66	0

56	Avg_LocalSILA	27.65 (8.71)	15.3 (9.26)	0.84	0
57	Skew_LocalSila	0.71 (0.42)	0.49 (0.68)	0.6	0

Multivariate analysis

In order to select the optimal variables, adjust the regression coefficients to optimize the transportability (external validity) of the model and determine the degree of optimism of the model and perform optimism-corrected analysis of the performance of the model by ROC analysis, all selected 57 quantitative imaging variables were included in the LASSO regression model. Multivariate analysis using LASSO on all features yielded a multivariate model with 8 selected features (selected with frequency > 50% after introducing bootstrap to reduce variability after 1000 runs) with an AUC estimate of 0.941. These 8 features include: 1. centroid_Z, 2. Min Enclosing Brick, 3. flatness, 4. SILA_Tex, 5. Max_SI, 6. Avg_SI, 7. Avg_PosMeanCurv and 8. Min_MeanCurv, all with $P<0.01$. To correct overfitting (internal validation) we used the bootstrapping technique to estimate the optimism of the AUC. **The optimism-corrected AUC is 0.939**.

Centroid_z captures the location of the nodule in the lung (vertical axis), the **minimal enclosing brick** and **flatness** capture volume and shape, respectively, **Sila_Tex** is a summary variable capturing the degree of abnormality based on texture density within the nodule, **maximum** and **average shape index (Max_SI and Avg_SI)** capture the complexity of the nodule surface and **Average positive mean curvature and (Avg_PosMeanCurv)** and **Minimum mean curvature (Min_MeanCurv)** represents the degree of curvature of the outer surface of the nodule.



We then added variables independently associated with an increased risk of lung cancer in our cohort (age and pack-years). The optimism-corrected AUC for model 2 is 0.941.

Model 2 – radiomic features + clinical variables

```
Coefficients:
                    Estimate Std. Error z value Pr(>|z|)
(Intercept)        -8.189458   2.390602  -3.426 0.000613 ***
Centroid_Z          0.005665   0.002095   2.704 0.006846 **
Min.Enclosing.Brick 0.178434   0.057463   3.105 0.001901 **
Flatness            0.390379   0.227995   1.712 0.086855 .
SILA_Tex            0.023527   0.004142   5.680 1.35e-08 ***
Min_MeanCurv       -0.332742   0.385821  -0.862 0.388454
Avg_PosMeanCurv    -1.425776   1.131270  -1.260 0.207550
Max_SI              0.663254   0.425328   1.559 0.118904
Avg_SI              1.759540   1.844368   0.954 0.340080
age                 0.063890   0.024851   2.571 0.010143 *
pkyr                0.011214   0.005324   2.106 0.035171 *
```

AUC: 0.944

Optimism correction using bootstrap

Mean of Bootstrap AUC is 0.947
Mean of Test AUC is 0.944

The difference is 0.003

Optimism-corrected AUC for Model 2:

0.944 – 0.003 = **0.941**

Validation

Due to considerable delay in enrollment of the DECAMP1 study (see above), validation of our model on a prospective cohort of screened individuals similar to those enrolled in the NLST is still pending. Application to access this dataset was completed and submitted to and approved by the DECAMP biomarker committee for image transfer.

DECAMP1 dataset validation

The most recent status update from DECAMP1 as of September 9, 2018 is as follow:

ENROLLMENT	
Accrual Goal	*500*
Enrolled (as of 09SEPT2018)	425
Average Monthly Accrual (based on past 6 months)	4
Projected Completion Date (based on past 6 months)	April 2020
ADJUDICATION	
Adjudication Review Complete	358
*Lung Cancer Status Useable (Cancer vs Benign)	335

* Metastatic and Off-Study Cases not included

An early validation on the initial DECAMP1 data, as previously reported, was disappointing:

Section 1. Summary of the DECAMP Data

BUcancer	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Benign	64	26.78	64	26.78
Lung Cancer	143	59.83	207	86.61
Metastatic Cancer	3	1.26	210	87.87
OFF-STUDY	4	1.67	214	89.54
Presumed Benign	10	4.18	224	93.72
Presumed Lung Cancer	15	6.28	239	100.00

Frequency Missing = 11

Section 2: Strict validation (i.e., using the 8-feature logistic model developed from the NLST data to impute the probability of lung cancer occurrence for each DECAMP case).

For the dataset "dod_Decamp1BenignMalignant.csv"

1. Only include benign vs. lung cancer;

   AUC: 0.6567

2. Add the presumed benign (as "benign") and the presumed lung cancer (as "lung cancer").

   AUC: 0.6484

Section 3: Loose validation (i.e., using the same 8 features identified from the NLST data to then re-fit the logistic regression, plus Bootstrap correction)

For the dataset "dod_Decamp1BenignMalignant.csv"

1. Only include benign vs. lung cancer;

   AUC without correction via bootstrap: 0.7415

   AUC with correction via bootstrap: 0.72326

2. Add the presumed benign (as "benign") and the presumed lung cancer (as "lung cancer").

   Auc without correction using boostrap: 0.7245 with correction 0.70725

For the dataset "dod_Decamp1BenignMalignant-Curated.csv"

1. Only include benign vs. lung cancer;

AUC without correction via bootstrap: 0.7234

AUC with correction via bootstrap: 0.6954

2. Add the presumed benign (as "benign") and the presumed lung cancer (as "lung cancer").

AUC without correction using bootstrap: 0.6853

AUC with correction via bootstrap: 0.6511

Section 4: Following the strict validation, we first pinpointed an optimum cutoff from the 8-feature logistic model developed from the NLST data, and then calculated the sensitivity and specificity at this cutoff.

1. 8-feature model in the NLST data
   Cutoff via Youdan's index: 0.4783
   Sensitivity: 0.9039
   Specificity: 0.8553

2. For the dataset "dod_Decamp1BenignMalignant.csv"
   a. Only include benign vs. lung cancer
      Sensitivity: 0.25
      Specificity: 0.9434
   b. Add the presumed benign (as "benign") and the presumed lung cancer (as "lung cancer").
      Sensitivity: 0.2644
      Specificity: 0.9435

3. For the dataset "dod_Decamp1BenignMalignant-Curated.csv"
   a. Only include benign vs. lung cancer
      Sensitivity: 0.1633
      Specificity: 0.9431
   b. Add the presumed benign (as "benign") and the presumed lung cancer (as "lung cancer").
      Sensitivity: 0.1636
      Specificity: 0.9429

As we are awaiting full recruitment and adjudication of the DECAMP1 dataset, alternative validation datasets were pursued, curated and analyzed. *Note that no funding from this grant was used for these analyses*.

<u>Lung Tissue Research Consortium validation</u>

The radiomic model was validated using the Lung Tissue Research Consortium dataset, comprised of 88 benign and 89 malignant nodules. This cohort was considered "high-risk" as all

nodules in this cohort were evaluated by expert radiologists and felt to be suspicious enough for malignancy to require surgical resection (i.e. a nodules, benign and malignant, were resected lung nodules and therefore with a high pre-test probability than typical screen- or incidentally identified lung nodules) Hence, this is a very different cohort than the cohort on which our radiologic model was derived (NLST), and we did not expect that it would perform as well.

Using these 177 nodules, the results were as follow:

Sensitivity: 87.6%

Specificity: 68.2%

PPV: 73.6%

NPV: 84.5%

Negative likelihood ratio 0.18 (95% CI 0.10-0.32)

Positive likelihood ratio 5.51 (95% CI 3.11-9.77)

While the results are clearly inferior to those expected based on our internal validation, the nature of the LTRC database comprised of nodules with a very high pretest probability of malignancy make these results encouraging as we are in the process of validating these results on the more similar Vanderbilt and DECAMP1 database.

Vanderbilt nodule cohort validation

In addition, we secured an alternative validation cohort from the lung nodule registry at Vanderbilt University Medical Center/Nashville Veterans Administration Tennessee Valley Healthcare system (primary investigator: Dr. Pierre Massion). All CT datasets have now been de-identified with corresponding clinical data recorded on a database and password-protected at Vanderbilt University. We performed quality control on these CT datasets to ensure that they met minimum criteria for radiomic analysis and tagged the nodules for analysis. This was approved by both Mayo Clinic and Vanderbilt University's respective institutional review boards.

A total of 84 benign nodules and 92 malignant nodules (total 176 nodules), all histology-proven (post-resection or biopsy). Application of the original 8 variable model yielded an AUC of 0.894.

- **What opportunities for training and professional development has the project provided?**

  Nothing to report.

- **How were the results disseminated to communities of interest?**

  Nothing to report.

- **What do you plan to do during the next reporting period to accomplish the goals?**

The development and internal validation of our radiological model using quantitative radiologic variables was completed with an optimism-corrected area under the receiver operating curve of 0.939. As we await full enrollment of the DECAMP1 study, we started validating this radiomic model on other datasets. Initial data from the LTRC and Vanderbilt datasets are promising. We are currently calculating the pre-test probability of malignancy for all the nodules in the NLST and in the Vanderbilt cohort using the Brock model to determine whether our model performs better than a commonly used clinical tool. We are planning on submitting the results of this work to be presented at the 2019 American Thoracic Society meeting. Due to considerable delays in recruitment in the DECAMP1 study, we have not yet been able to externally validate our results using the proposed validation set. Nonetheless, we are hopeful that recruitment will be completed and all benign and malignant nodules adjudicated to allow for formal validation.

While addition of clinical variables in our model 2 (clinical-radiological model) did not appear to provide superior performance of the model, it is possible that validation may be improved with model 2 and we plan on validating model as well and compare these two models.

4. IMPACT:

- **What was the impact on the development of the principal discipline(s) of the project?**

An estimated 1.5 million new lung nodules are identified via chest CT annually in the US, which is likely an underestimate given the ever-expanding use of HRCT in the US and in the world. This is also likely to increase markedly with implementation of lung cancer screening for high-risk individuals, with a number of individuals eligible for lung cancer screening estimated around 10 million in the US alone. Only approximately 10,000 individuals have been screened based on Medicare data as of May 2017. The large number of individual with false positive screening CTs, approximately 40% in the NLST, is likely to result in unnecessary invasive diagnostic interventions with excessive morbidity, mortality, patient stress and healthcare expenses.

We have previously demonstrated that volumetric CT-based quantitative characterization can risk-stratify lung nodules of the adenocarcinoma spectrum. This approach eliminates the intra- and inter-observer variability and subjectivity of CT image interpretation by trained radiologists. In addition, modern digital CT images include a large amount of valuable high-dimensional data not currently utilized to assist in diagnosis. This invaluable unexploited resource can be leveraged by modern quantitative imaging methods. Radiomic approaches to lung nodule analysis consist of extracting reproducible and objective quantitative radiological variables from CT datasets, reducing large volumes of complex data into manageable and clinically relevant information. These quantitative imaging techniques have been proposed to facilitate the development of diagnostic and prognostic models in lung imaging, allowing for example the risk-stratification of lung adenocarcinomas, the classification of screen-or incidentally detected lung nodules and the characterization of lung cancer subtypes and tumor heterogeneity. We used to the NLST dataset to develop and internally validate a radiological multivariate model that

include quantitative radiological features distinguishing malignant from benign CT-screen detected indeterminate pulmonary nodules. If this model is externally validated on a broad scale, it could lead to substantial improvement in lung nodule management, available to a large audience of clinicians and radiologists as a software-based image analytical tool which could substantially reduce error and reduce the risk of unnecessary invasive and non-invasive procedures.

- **What was the impact on other disciplines?**

  Nothing to report.

- **What was the impact on technology transfer?**

  Nothing to report.

- **What was the impact on society beyond science and technology?**

  Our project is not completed yet, but if successful could have a major impact on lung nodule management, by offering clinicians and radiologists reproducible tools to assist in the management of incidentally or screen-identified lung nodules, a major healthcare problem that affects Veteran and non-Veteran populations. Quantitative nodule analysis can be applied to existing CT scans obtained for screening or clinical indications and do not require additional testing beyond software application of image analytics. Our quantitative analytics tool could help standardize the management of lung nodules and lead to a substantial reduction in the unnecessary morbidity, mortality and healthcare costs.

5. **CHANGES/PROBLEMS:**

- **Changes in approach and reasons for change**

  There hasn't been a major change in approach, except for the pursuit of additional validation sets given the considerable delays in accumulating enough cases in the DECAMP1 dataset to allow for enough power. None of the Department of Defense funds allocated to these analyses. We are awaiting the DECAMP1 final results.

- **Actual or anticipated problems or delays and actions or plans to resolve them**

  This award was effective on September 30, 2015, but because of the relocation of the grant PI (Fabien Maldonado) from Mayo Clinic, Rochester, MN to Vanderbilt University, Nashville, TN, substantial delays were incurred from the need to establish subcontracts between the three partnering institutions (Mayo Clinic, Brown University and Vanderbilt University), which were eventually finalized in April 2016. This resulted in a significant delay for case selection and image transfer from the ACRIN and LSS core labs and our work on the development and optimization of discriminative radiological quantitative variables.

  However, the variables were developed and optimized by the end of 2016 and both model 1 (radiological model) and model 2 (clinical-radiological model) were developed and internally

validated using LASSO for variable penalization and selection and bootstrapping for internal validation. External validation, however, has been hampered by delays in recruitment in our planned validation dataset, the DECAMP1 dataset (PI: Dr. Avrum Spira). Accordingly, we have pursued additional validation cohorts and were able to validate our radiological model using the LTRC and Vanderbilt datasets. The former dataset, however, is significantly different than our derivation dataset in that all nodules were resected because of high suspicion of malignancy, explaining the decreased diagnostic test performance of our radiomic model. The Vanderbilt dataset was more similar and accordingly yielded better results with an AUC of 0.894.

- **Changes that had a significant impact on expenditures**

  Nothing to report.

- **Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents**

  Nothing to report.

- **Significant changes in use or care of human subjects**

  Nothing to report.

- **Significant changes in use or care of vertebrate animals**

  Nothing to report.

- **Significant changes in use of biohazards and/or select agents**

  Nothing to report.

6. **PRODUCTS:**

- **Publications, conference papers, and presentations**
  Conference paper:
  Computed tomography-based radiomic classifier distinguishes malignant from benign nodules in the national screening trial

  18th World Conference on Lung Cancer

  October 15 - 18 2017 | Yokohama, Japan http://wclc2017.iaslc.org/

- **Journal publications.**

  Journal publication:

  Peikert T, Duan F, Rajagopalan S, Karwoski RA, Clay R, Robb RA, Qin Z, Sicks J, Bartholmai BJ, **Maldonado F**. Novel high-resolution computed tomography-based radiomic classifier for

screen-identified pulmonary nodules in the National Lung Screening Trial. *PLoS One*. 2018 May 14;13(5):e0196910.

- **Books or other non-periodical, one-time publications.**
  Nothing to report.

- **Other publications, conference papers, and presentations.**
  Nothing to report.

- **Website(s) or other Internet site(s)**
  Nothing to report.

- **Technologies or techniques**
  Novel CT-based quantitative analytics to distinguish benign from malignant nodules. How this novel analytical tool will be shared has not yet been determined.

- **Inventions, patent applications, and/or licenses**
  Nothing to report.

- **Other Products**
  Nothing to report.

7. **PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS**

- **What individuals have worked on the project?**

Name: Tobias Peikert
Project Role: PD/PI
Research Identifier: N/A
Nearest Person Months: 1.0
Contribution to the Project: Mayo Clinic PI, administrative leadership at Mayo Clinic, review and selection of all benign NLST (nodules) training set and benign and malignant DECAMP nodules. Shared supervision of Dr. Rajagopalan and Ron Karwoski with Dr. Bartholmai. Participation in weekly team videoconferences.

Name: Srinivasan Rajagopalan
Project Role: Co-Investigator
Research Identifier: 0000-0003-3286-1529
Nearest Person Months: 6.0
Contribution to the Project: Image analysis and development of imaging variables. Participation in weekly meetings.

Other Support Changes (since 2017 Annual Report)
Maldonado, Fabien, M.D.
Ended: None
New: 1 R01 EB024864-01 (Webster); W81XWH-17-1-0442 (Blackwell); 5 U01 CA196405-04 (Massion)

Srinivasan, Rajagopalan, Ph.D.
Ended: None
New: None

Tobias, Peikert, M.D.
Ended: None
New: None

- **Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?**
  No.

- **What other organizations were involved as partners?**
  Organization Name:
  Location of Organization: *(if foreign location list country)*
  Partner's contribution to the project: *(identify one or more)*
  Financial support;
  In-kind support *(e.g., partner makes software, computers, equipment, etc., available to project staff)*;
  Facilities *(e.g., project staff use the partner's facilities for project activities)*;
  Collaboration *(e.g., partner's staff work with project staff on the project)*;
  Personnel exchanges *(e.g., project staff and/or partner's staff use each other's facilities, work at each other's site)*; and
  Other.

8. **SPECIAL REPORTING REQUIREMENTS**
   N/A

9. **APPENDICES:**
   N/A