

Iowa Initiative for Artificial Intelligence

Final Report

Project title:	Sarcopenia quantification using novel automated software to predict patient outcomes		
Principal Investigator:	Peter Nau		
Prepared by (IIAI):	Yanan Liu, Honghai Zhang		
Other investigators:			
Date:			
Were specific aims fulfilled:		Y	
Readiness for extramural proposal?		Y/N	
If yes ... Planned submission date		Y	
Funding agency		SAGES	
Grant mechanism			
If no ... Why not? What went wrong?			

Brief summary of accomplished results:

The highest prediction correctness of 75% for home discharge was achieved with all available patient information combined with 3D radiomic features. When comparing the prediction value of 2D vs. 3D RFs in the tested prediction models, which differed in the richness of the input data but always included 2D or 3D radiomic features, 3D psoas muscle analyses and associated 3D radiomic features consistently and statistically significantly outperformed that of 2D analyses and 2D radiomic features ($p < 0.01$ for both correctness and F1 score comparisons).

Research report:

Aims (provided by PI):

Aim 1: Develop a quantitative index of sarcopenia severity based on 2D as well as 3D measurements of psoas anatomy using an existing population of 350 patients who have undergone assessments using both techniques – based on the measurements, associate each patient with a sarcopenia quartile of severity.

Aim 2: A population of critically ill patients will be evaluated using the current semi-automated Psoas Analyzer program. This program will calculate the size, shape and texture features using PyRadiomics.

Aim 3: Explore the possibility of automated psoas muscle segmentation using Deep Logismos. Deep Logismos is the combination of deep learning automatic intelligence and graph search segmentation. This will be used to produce data which can yield more accurate results while also allowing for smart editing of localized errors. Demonstrate that the fully automated 3D analysis achieves at least the same prediction accuracy of clinical outcomes as the semi-automated approach evaluated in Aim 2.

Data:

Psoas muscles from 317 patients were segmented in whole body CT scans by our LOGISMOS software. Psoas muscles were separately analyzed in 2D (single slice analysis at L3 level yielded total psoas area

TPA) and 3D (total psoas volume TPV). 2D and 3D radiomic features (RF) were extracted for TPA and TPV using PyRadiomics package. Patient information included patient characteristics (PC), injury severity score (SS), time in ICU (TII) and time until discharge (TUD).

Then more patients were collected. Psoas muscles from 628 patients were segmented in whole body CT scans by our LOGISMOS software. Psoas muscles were analyzed in 3D to generate total psoas volume (TPV). 3D radiomic features (RF) were extracted for TPV using PyRadiomics package. Patient information included patient characteristics (PC), pre operation information (POI), surgical information (SI) and blood transfusion information (BT).

AI/ML Approach:

Random forest classifier was trained to predict clinical outcomes. Prediction performance was assessed using 5-fold-cross-validation. Performance of predicting the following clinical outcomes was evaluated by quantifying prediction correctness (mean $\pm$ standard deviation) and F1-score (1 is best).

Results:

2D VS. 3D ... The highest prediction correctness of 74% was achieved with all available patient information combined with 3D radiomic features (Table 1). When comparing the prediction value of 2D vs. 3D RFs in the tested prediction models, which differed in the richness of the input data but always included 2D or 3D radiomic features, 3D psoas muscle analyses and associated 3D radiomic features consistently and statistically significantly outperformed that of 2D analyses and 2D radiomic features ($p < 0.01$ for both correctness and F1 score comparisons).

Table 1 Prediction of patient home discharge without assist Accuracy (mean $\pm$ -std) and F1 score of different input feature combination (3D and 2D)

	Patient-specific features used in respective prediction models					
Predicting patient home discharge without assist	RF only	RF+PC	RF+SS	RF+PC+SS	RF+PC+SS+TII	RF+PC+SS+TII+TUD
3D prediction correctness [mean $\pm$ stdev %]	57 $\pm$ 2	60 $\pm$ 3	64 $\pm$ 4	65 $\pm$ 3	67 $\pm$ 4	74 $\pm$ 3
2D prediction correctness [mean $\pm$ stdev %]	56 $\pm$ 6	55 $\pm$ 2	59 $\pm$ 3	62 $\pm$ 5	64 $\pm$ 4	72 $\pm$ 5
3D F1-score	0.644	0.672	0.702	0.718	0.728	0.772
2D F1-score	0.63	0.624	0.647	0.672	0.697	0.762

Prediction of Clinical outcomes

Quantitative indices of achieved performance and definition of individual performance tasks is given in Table 2.

Table 2 Prediction of Clinical outcomes Accuracy (mean+/-std) and F1 score of different input feature combination

	Clinical outcome prediction performances					
	Discharge to home (313p/573)	Discharge to other facilities (194p/574)	In hospital>=15days (193p/603)	Blood transfusion (173p/553)	Post op septic shock (160p/554)	Not be discharged to a facility and not go into septic shock (259 p/504)
All features [mean±stdev %]	75±4	69±2	67±2	77±3	75±3	69±3
All features+BT [mean±stdev %]	73±4	71±3	66±2		75±3	70±4
All features F1	0.772	0.407	0.214	0.528	0.450	0.680
All features+BT F1	0.752	0.452	0.215		0.456	0.698