Evaluation of the Stability of Radiomics Features using 4D-CT and across Radiomics Platforms for Lung and Liver Tumors

Xiao Wang [1], Chi Ma [1], Hao Wang [2], Yin Zhang [1], Ning J. Yue [1] and Ke Nie [1]

[1] Department of Radiation Oncology, Rutgers-Cancer Institute of New Jersey, Rutgers-Robert Wood Johnson Medical School, New Brunswick, NJ

[2] Department of Radiation Oncology, Shanghai Chest Hospital, Shanghai Jiaotong University, Shanghai, China

RUTGERS Cancer Institute of New Jersey

INTRODUCTION

Radiomics to extract quantitative imaging features has been widely discussed as a potential tool in making clinical judgements, with open-source software available. The stability and reproducibility of radiomics features though, still needs to be verified before incorporating radiomics in clinical trials. The test-retest scans can be used for this purpose, but it may not be clinically available all the time. On the other hand, 4D-CT which contains images acquired at different time stamps may serve as an alternative to mimic the test-retest scenario with each phase as a separate scan. In addition, since many open-source software are available to the public, the variation in the algorithm implementation across platforms also needs to be tested. Our study included time dependence, contour dependence, disease site dependence, and also radiomics platform dependence, when we evaluated feature stability.

AIM

To evaluate the stability of radiomics features using 4D-CT as an alternative to test-retest CT scans, and the consistency across radiomics platforms.

METHOD

4D-CT images with 10 breathing phases were acquired for 10 patients with lung tumors and 10 patients with liver tumors. For each patient, two contours (GTV and GTV-1mm) were delineated on each breathing phase. GTV-1mm was generated by subtracting a 1mm inner margin from GTV to investigate the impact of contouring accuracy. A total of 16 first-order histogram-based and second-order texture-based features (Gray Level Co-occurrence Matrix as GLCM, and Gray Level Run Length Matrix as GLRLM) were extracted using two open-source radiomics platforms. The intraclass correlation coefficient (ICC) of each feature was calculated from all breathing phases to evaluate test-retest stability. The concordance correlation coefficient (CCC), the ICC(A,1) with an absolute agreement definition, the ICC(C,1) with a consistency definition, as well as Pearson correlation coefficient (PCC) were calculated to evaluate the agreement/consistency of the two radiomics platforms. Features with ICC/CCC above 75% were deemed stable features.

RESULTS

Regardless of difference in contours, all 16 features were identified as stable in test-retest CT scans for both radiomics platforms, with ICC ranging from 0.785 to 1. The two radiomics platforms only concorded on one feature (Correlation in GLCM) for GTV in liver tumors (ICC(A,1)=0.88, CCC=0.879). The two radiomics platforms showed consistency on two features: short run emphasis (in GLRLM) for GTV-1mm in lung tumors (ICC(C,1)=0.813, PCC=0.813), and entropy (in 1st order) in liver tumors for GTV (ICC(C,1)=0.999, PCC=0.999), and for GTV-1mm (ICC(C,1)=0.993, PCC=0.994).

	lung				liver			
	IBEX		CERR		IBEX		CERR	
	GTV	GTV-1	GTV	GTV-1	GTV	GTV-1	GTV	GTV-1
energy	0.983	0.980	0.969	0.971	0.995	0.989	0.958	0.960
contrast	0.994	0.993	0.981	0.979	0.967	0.880	0.973	0.962
dissimilarity	0.994	0.994	0.985	0.987	0.981	0.974	0.982	0.976
corr	0.995	0.980	0.978	0.983	0.945	0.973	0.923	0.978
sre	0.985	0.994	0.987	0.989	0.999	0.991	0.966	0.972
lre	0.993	0.992	0.975	0.982	0.998	0.991	0.858	0.874
gln	0.999	0.999	0.997	0.999	1.000	1.000	0.974	0.973
rln	1.000	1.000	0.995	0.999	1.000	1.000	0.992	0.995
rp	0.992	0.995	0.999	0.999	0.999	0.993	0.993	0.989
lglre	0.998	0.997	0.980	0.980	0.785	0.995	0.974	0.926
hglre	0.996	0.997	0.971	0.975	0.986	0.994	0.994	0.934
srlgle	0.998	0.997	0.981	0.980	0.891	0.995	0.921	0.806
srhgle	0.995	0.997	0.972	0.969	0.989	0.987	0.997	0.949
lrlgle	0.996	0.996	0.961	0.967	0.980	0.991	0.820	0.821
lrhgle	0.996	0.994	0.981	0.983	0.998	0.992	0.849	0.870
entropy	0.974	0.989	0.992	0.958	0.980	0.986	0.982	0.985

Table 1. *ICC between 10 phases CT for all features with different contours, disease site and platforms.*

	ICC(A,1)				CCC				ICC(C,1)				PCC			
	lung		liver		lung		liver		lung		liver		lung		liver	
	GTV	GTV-1	GTV	GTV-1	GTV	GTV-1	GTV	GTV-1	GTV	GTV-1	GTV	GTV-1	GTV	GTV-1	GTV	GTV-1
energy	0.101	0.264	-0.011	-0.053	0.100	0.262	-0.011	-0.053	0.188	0.336	-0.029	-0.070	0.795	0.645	-0.327	-0.279
contrast	0.007	0.014	-0.118	-0.548	0.007	0.013	-0.116	-0.539	0.077	0.059	-0.356	-0.551	0.544	0.554	-0.631	-0.627
dissimilarity	0.019	0.059	-0.117	-0.455	0.019	0.058	-0.115	-0.449	0.429	0.307	-0.437	-0.452	0.632	0.717	-0.559	-0.806
corr	0.179	0.193	0.880	0.495	0.177	0.192	0.879	0.492	0.648	0.467	0.879	0.584	0.662	0.468	0.879	0.678
sre	0.292	0.742	0.005	-0.188	0.290	0.740	0.005	-0.186	0.477	0.813	0.011	-0.187	0.649	0.813	0.024	-0.449
lre	0.242	0.551	-0.003	-0.006	0.240	0.549	-0.002	-0.006	0.351	0.602	-0.004	-0.007	0.668	0.669	-0.294	-0.340
gln	0.037	0.052	0.063	0.074	0.037	0.051	0.062	0.073	0.050	0.067	0.112	0.116	0.990	0.992	0.874	0.872
rln	0.210	0.188	0.078	0.074	0.209	0.186	0.078	0.074	0.309	0.241	0.095	0.096	0.961	0.986	0.580	0.536
rp	0.000	0.000	0.001	0.000	0.000	0.000	0.001	0.000	0.002	0.002	0.160	0.070	0.019	0.060	0.438	0.505
lglre	0.031	0.059	0.000	0.000	0.031	0.059	0.000	0.000	0.152	0.224	0.000	0.000	0.463	0.785	0.149	0.306
hglre	0.011	0.018	0.000	0.000	0.011	0.018	0.000	0.000	0.060	0.102	-0.038	0.212	0.438	0.707	-0.048	0.283
srlgle	0.029	0.057	0.001	0.000	0.028	0.057	0.001	0.000	0.143	0.219	0.002	0.000	0.447	0.760	0.273	0.076
srhgle	0.006	0.012	0.000	-0.001	0.006	0.012	0.000	-0.001	0.038	0.088	-0.022	-0.158	0.300	0.638	-0.031	-0.286
lrlgle	0.026	0.051	0.000	0.000	0.026	0.050	0.000	0.000	0.098	0.197	0.000	0.000	0.371	0.778	-0.193	-0.255
lrhgle	0.091	0.074	-0.055	-0.018	0.090	0.073	-0.055	-0.018	0.222	0.153	-0.358	-0.273	0.734	0.633	-0.364	-0.384
entropy	0.146	0.384	0.261	0.306	0.145	0.381	0.259	0.304	0.318	0.450	0.999	0.993	0.614	0.498	0.999	0.994

Table 2. *ICC(A,1), CCC, ICC(C,1), and PCC for all features with different contours and disease site.*

CONCLUSIONS

The results show that the stability of radiomics features can be evaluated using 4D-CT as an alternative to test-retest CT scans. The two radiomics platforms show some consistency, but low agreement. The stability evaluation also depends on the contouring accuracy and disease site. Overall, we are able to identify stable radiomics features in test-retest and across platforms which may have the potential to be used for radiomics-guided clinical studies.

REFERENCES

Nie K et al. NCTN assessment on current applications of Radiomics in oncology. *International Journal of Radiation Oncology* Biology* Physics. 2019 Jun 1;104(2):302-15.*

CONTACT INFORMATION

xw240@cinj.rutgers.edu