The reliability and precision of OsiriX® imaging software in the assessment of psoas muscle surface area from computed tomography scans in advanced cancer patients

Stefanie Fallone*[1,4], Noor Mady*[2,3,4], Robert D. Kilgour[1,4], Leonard Rosenthall[5], Sarah Khan[4], Antonio Vigano[4]

[1]Department of Exercise Science; [2]Department of Psychology; [3]Science College, [4]McGill Nutrition and Performance Laboratory, McGill University Health Centre; [5]Department of Radiology, McGill University. *Both authors contributed equally to this study.

Introduction

- Assessment of skeletal muscle using computed tomography (CT) is important in determining sarcopenia and cachexia in advanced cancer[1].

- Medical imaging software programs such as OsiriX are available to quantify paraspinal skeletal muscle from CT scans[2,3].

Aim

- To determine the reliability and precision of OsiriX medical software in assessing the cross-sectional surface area of the psoas muscle at the 4th lumbar vertebra (L4).

- It is hypothesized that there are no significant intra- and inter-rater reliability differences for measurements of the surface area of the psoas muscle at L4 over time and among the top, middle and bottom scan measures.

Methods

- Psoas muscle cross sectional area was measured using OsiriX software from CT scans obtained from 19 advanced cancer patients.

- Reliability measurements were done over time (t=0, 24 hrs, 2 wks) and over three regions of L4 (top, mid, bottom) by two separate raters (SF & NM).
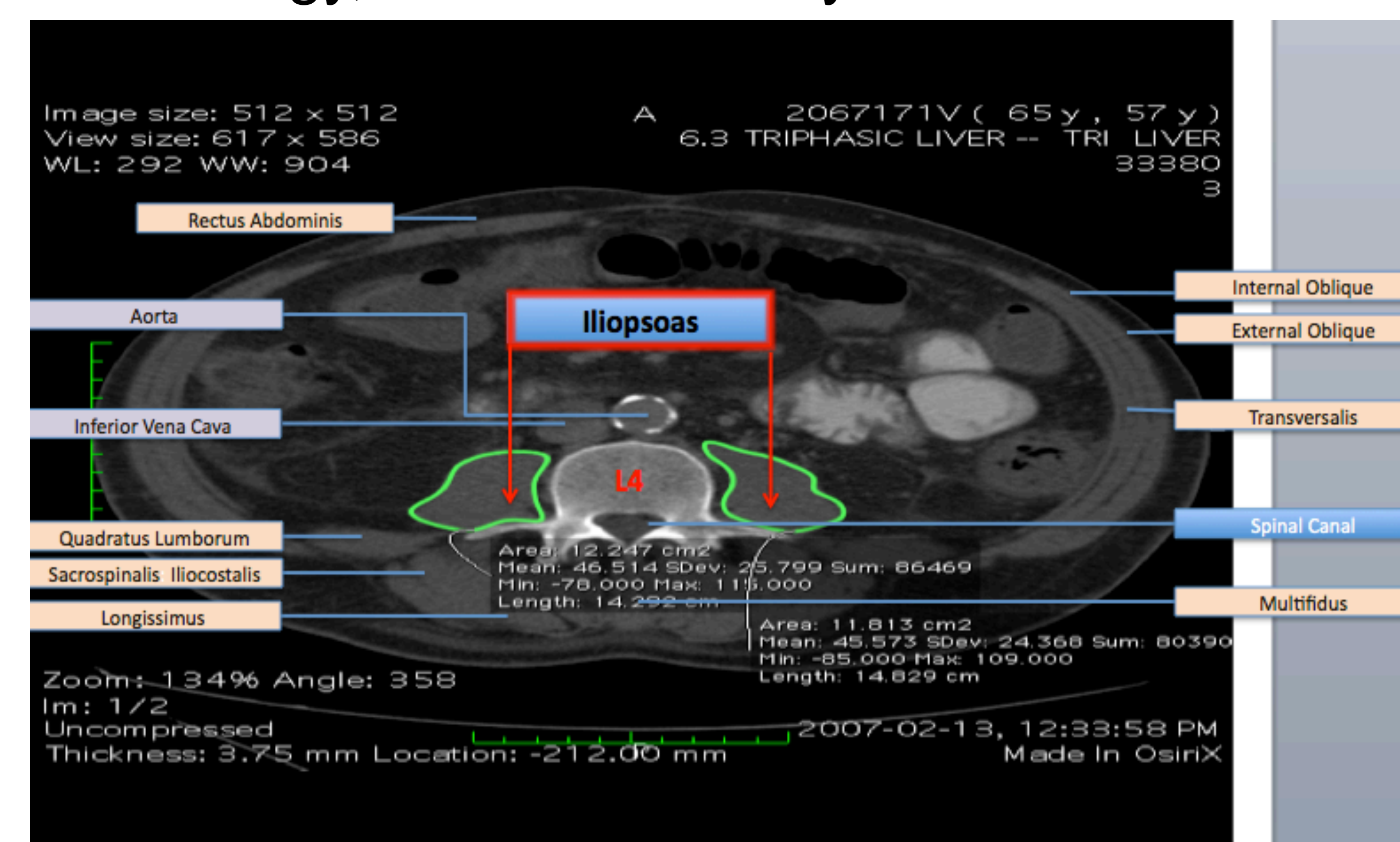
Results

Table 1. Inter-rater differences in psoas cross-sectional area measurements

Measurement time intervals	Pearson correlations between Rater 1 and 2
t=0 hours	0.990
t=24 hours	0.983
t=2 weeks	0.987

Table 2. %CV and SDcm² comparisons between raters at the different L4 levels

| L4 level	Rater 1 %CV	Rater 2 %CV	P-value*	Rater 1 SDcm²	Rater 2 SDcm²	P-value*
Top	4.244	3.750	0.518	0.723	0.606	0.337
Middle	3.385	3.030	0.753	0.586	0.532	0.577
Bottom	4.199	4.201	0.996	0.794	0.790	0.974

*Paired t-test; %CV is the coefficient of variation; SDcm² is the standard deviation.

Table 3. Multiple analyses for each regression line (Rater 1 vs Rater 2)

L4 regions	RMSE	Correlation	Limits of Predictability
Top	0.772	0.992	±1.514
Middle	0.777	0.993	±1.522
Bottom	0.870	0.992	±1.711

RMSE=root mean squared error
Limits of Predictability= 1.96 x RMSE or the dispersion of points above and below the line of regression

Figure 1. Scatter plots of rater 1 versus rater 2 depicting the line of regression and limits of prediction for the three L4 levels

L4 Top   L4 Middle   L4 Bottom
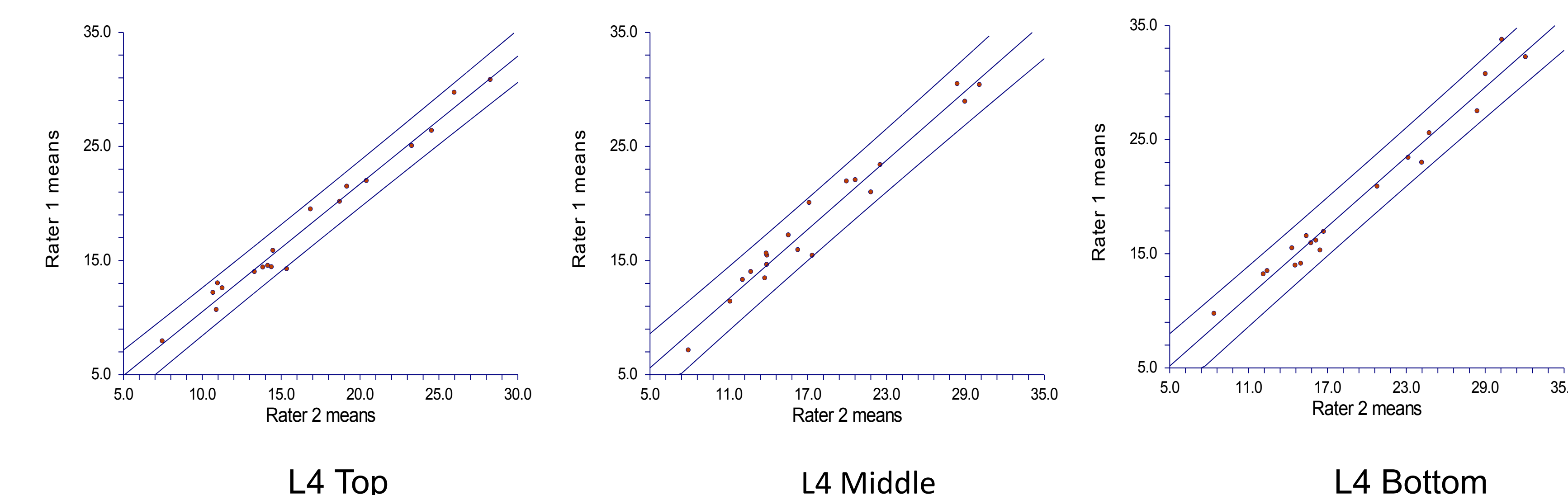
Discussion

- Inter-rater correlations demonstrated a high reliability over time and at each of the three levels of L4.
- No differences in the level of precision (%CV and SDcm²) between raters.
- OsiriX medical imaging software is shown to be reliable and precise and can be used by multiple raters to assess psoas muscle surface area in axial CT scans.
- Future studies can take advantage of OsiriX technology to determine the presence of muscle wasting conditions such as sarcopenia and cachexia using CT.

References

[1] Baracos V, Kazemi-Bajestani S. Clinical outcomes related to muscle mass in humans with cancer and catabolic illnesses. *International Journal of Biochemistry & Cell Biology* [serial online].October 2013; 45(10):2302-2308. Available from: Academic Search Complete, Ipswich, MA. Accessed August 17, 2014.

[2] Rosset, A., Spadola, L., and Ratib, O. (2004). OsiriX: An Open-Source Software for Navigating in Multidimensional DICOM Images. Journal of Digital Imaging, 17(3), 205-216.

[3] Fortin, M., and Battie, M. C. (2012). Quantitative Paraspinal Muscle Measurements: Inter-Software Reliability and Agreement Using OsiriX and ImageJ. Phys. Ther, 92, 853-864.