On using Textural information from Longitudinal PET scans for Predicting Tumor Response

Jose George[1,2,6⋆], **Kathleen Vunckx**[1,3], **Sabine Tejpar**[4], **Christophe M. Deroose**[3], **Johan Nuyts**[1,3], **Dirk Loeckx**[1,2,5,6] & **Paul Suetens**[1,2,6]

[1]Medical Imaging Research Center, UZ Leuven [2]ESAT/PSI/MIC, [3]Nuclear Medicine, [4]Gastroenterology, KU Leuven, Belgium
[5]icoMetrix NV, Leuven, Belgium [6]IBBT-KU Leuven Future Health Department ⋆jose.george@uzleuven.be

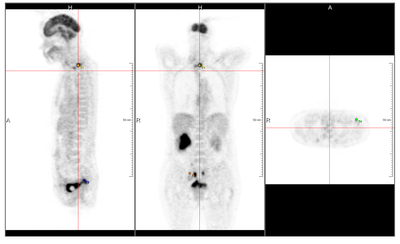
AACR Annual Meeting 2012 — March 31- April 4 • Chicago, IL
Abstract # 4544 — Poster Session: Clinical Research 18 — Poster Section 25 (4) — PO.CL01.03. Imaging and Molecular Diagnostics — Tue, Apr 3, 1:00 - 5:00 PM

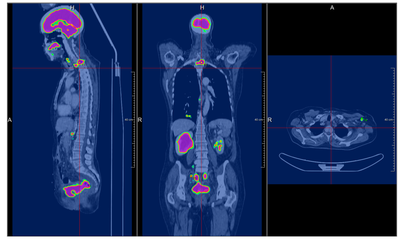
Contribution: Lesion Heterogeneity Change for Predicting Response

Early therapy response prediction, employing biomarkers such as [^{18}F]-fluorodeoxyglucose (**FDG**) followed with positron emission tomography (**PET**) imaging the tumor metabolic activity, turned out to be the de facto mode of evaluation in longitudinal studies. Traditionally, the first order intensity based feature estimates extracted from the longitudinal scans are used for the response evaluations. Nonetheless, lesion heterogeneity computed from the higher order textures of the baseline scan prior treatment has been recently found to be correlating well with the final outcome [1]. The aim of this study is to establish the significance of tumor texture in follow up evaluations.

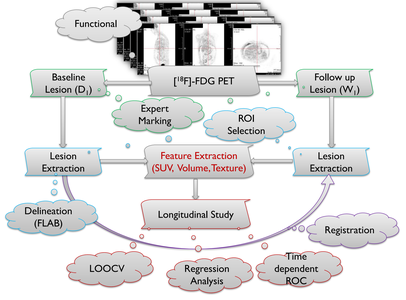
Clinical Lesions: [^{18}F]-FDG PET/CT Scans

15 chemo-naive first-line patients suffering from metastatic colorectal cancer (**mCRC**) and treated with cetuximab were selected for the study. Baseline [^{18}F]-**FDG PET** scans were acquired prior to treatment and the follow up scanned at 1 week post treatment was investigated for early response. 1-5 lesions were expert identified per patient and analyzed. Sagittal, coronal and axial orthogonal slices of the **PET** (left) and blended **PET**/**CT** (right) of a subject in the cohort is displayed below. Expert marking is also visible in the slices.





Longitudinal Evaluation



A block overview of various modules investigated in this longitudinal study is schematically represented above.

PET Lesion Delineation

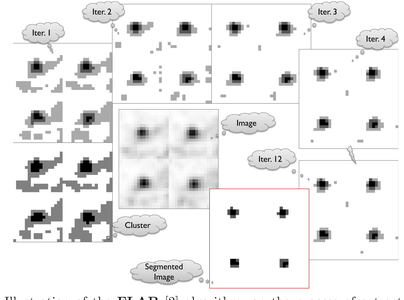
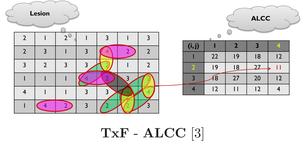
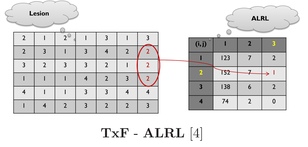


Illustration of the **FLAB** [2] algorithm, on the process of extracting a typical PET lesion, with initialization followed by step-by-step iterations demonstrated in a clock-wise manner.

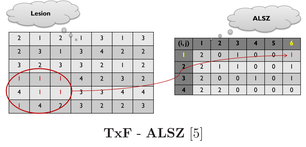
PET derived Covariates: Textural Matrices



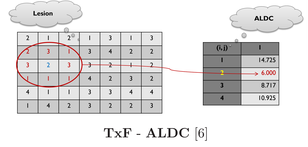
TxF - ALCC [3]



TxF - ALRL [4]



TxF - ALSZ [5]



TxF - ALDC [6]

Covariates / Feature Estimates

⇒ Traditional feature (**TrF**) estimates employed in this longitudinal study included 4 SUV-based, 1 volume-based and their 2 combinations constituting 7 **TrF** covariates.

⇒ Textural features (**TxF**) employed in this longitudinal study included 7 from **ALCC**, 11 from **ALRL**, 11 from **ALSZ**, 5 from **ALDC** and 19 from **ALHS** constituting 53 **TxF** covariates.

⇒ Activity level co-occurrence matrix [3] entries $\mathbf{ALCC}_{ij}$ express the frequency of co-occurrences of (i, j) activity level pairs.

⇒ Activity level run length matrix [4] elements $\mathbf{ALRL}_{ij}$ represent the frequency of i^{th} activity level's j runs in 13 directions.

⇒ Activity level size zone matrix [5] entries $\mathbf{ALSZ}_{ij}$ is updated when i^{th} activity level has a zone with j voxels.

⇒ Activity level difference matrix [6] **ALDC** consists of mean deviation of activity level relative to the neighbourhood.

⇒ Activity level histogram descriptive statistics **ALHS** were also computed and investigated.

Data Analysis

⇒ To compute a marker per subject, response indices (**RI**) computed for each lesion feature were accumulated either taking the **max**imum relative change ($\max\{\delta\text{FS}_{(\%)}\}_{lesions}$) or the relative change **cum**ulated ($\delta\sum\{\text{FS}\}_{lesions(\%)}$) over all lesions.

⇒ The baseline dynamic range of **TrF** estimates were investigated.

⇒ A Cox proportional hazards regression model was fit with **RI**s and analyzed. The regression coefficients were projected back to the feature set so as to compute the 'marker'.

⇒ The time dependent receiver operating characteristics (**ROC**) [7] were determined relating the continuous 'marker' with the continuous 'outcome'. The time to progression (**TTP**) was used as the 'outcome' information in our studies.

⇒ For more realistic evaluation, a leave one-out cross validation (**LOOCV**) was performed to investigate the model fitting.

⇒ In **LOOCV**, the regressor was fit with features from all subjects except one. The regressor was then tested on the remaining subject. This procedure was repeated for each subject to find the respective marker value.

⇒ The concordance measure (C_τ) [8] was determined by weighted integration of the individual area under curve (**AUC**).

⇒ **AUC** at **TTP** > 4 months (AUC_{TTP4}) was also computed.

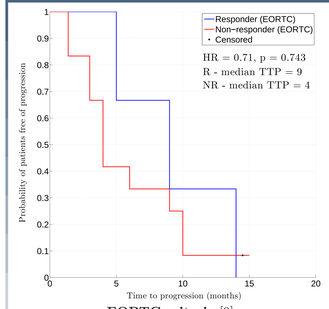
⇒ For multivariate feature selection (**MVFS**), univariate analysis was performed using a Cox model and significant features giving $p < 0.25$ were selected. Selected features (**FS**) are then used for creating the multivariate Cox model.

⇒ The current clinical standard **EORTC criteria** [9] and **PERCIST 1.0** [10] (using SUV - body weight) were studied.

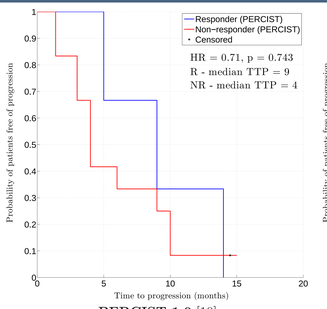
⇒ With **MVFS**, 2 predictive models, **pMVFS-max** and **pMVFS-cum**, were presented.

⇒ The statistical difference of survival curves was analyzed using logrank test.

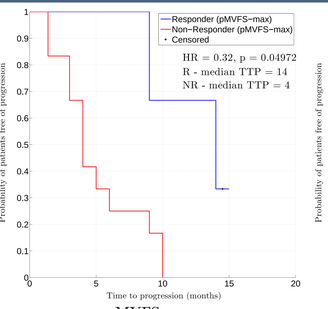
Survival Analysis: Kaplan-Meier Plot



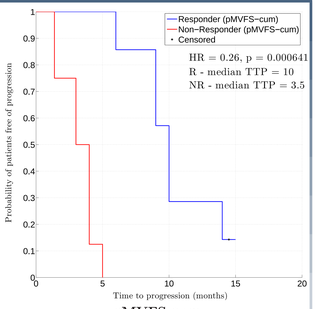
EORTC criteria [9]
$\max\{\delta\text{SUV}_{\max(\%)}\}_{lesions}$
$C_\tau = 0.6044$, $\text{AUC}_{\text{TTP4}} = 0.7045$
$\text{LOOCV} - C_\tau = 0.5659$
$\text{R} \Rightarrow \text{Marker} < -25\%$



PERCIST 1.0 [10]
$\max\{\delta\text{SUV}_{\text{peak}(\%)}\}_{lesions}$
$C_\tau = 0.6657$, $\text{AUC}_{\text{TTP4}} = 0.7955$
$\text{LOOCV} - C_\tau = 0.6509$
$\text{R} \Rightarrow \text{Marker} < -30\%$



pMVFS-max
$\text{Cox Coeff}_{\text{FS}} \times \max\{\delta\text{FS}_{(\%)}\}_{lesions}$
$C_\tau = 0.8313$, $\text{AUC}_{\text{TTP4}} = 0.9545$
$\text{LOOCV} - C_\tau = 0.8047$
$\text{R} \Rightarrow \text{Marker} < -2.1701\%$



pMVFS-cum
$\text{Cox Coeff}_{\text{FS}} \times \delta\sum\{\text{FS}\}_{lesions(\%)}$
$C_\tau = 0.9615$, $\text{AUC}_{\text{TTP4}} = 1.000$
$\text{LOOCV} - C_\tau = 0.9644$
$\text{R} \Rightarrow \text{Marker} < 8.0057\%$

Multivariate Feature Selection (MVFS)

2 **MVFS** predictive models are presented below:

Selected Features	Univariate p-value	Multivariate Cox Coeff.	Hazard Ratio	Multivariate p-value
pMVFS-max $\Rightarrow \text{Cox Coeff}_{\text{FS}} \times \max\{\delta\text{FS}_{(\%)}\}_{lesions}$				
SUV_{peak}	0.1529	0.0412	1.0421	0.1495
ALCC_{con}	0.1711	-0.1271	0.8807	0.2817
ALCC_{epy}	0.1992	-0.0391	0.9617	0.4153
ALCC_{dis}	0.1654	0.1843	1.2024	0.1130
ALRL_{sre}	0.2473	-0.6104	0.5431	0.0328
ALRL_{rp}	0.2424	0.5219	1.6852	0.0315
ALDC_{con}	0.1985	-0.002	0.9998	0.6759
pMVFS-cum $\Rightarrow \text{Cox Coeff}_{\text{FS}} \times \delta\sum\{\text{FS}\}_{lesions(\%)}$				
ALCC_{sam}	0.2283	-0.0919	0.9122	0.0387
ALCC_{idm}	0.1606	-3.0988	0.0451	0.0649
ALCC_{hom}	0.1645	4.2081	67.2291	0.0500
$\text{ALRL}_{\text{lrhae}}$	0.2313	-0.0596	0.9422	0.3385
ALSZ_{sze}	0.1557	0.0079	1.0080	0.6447
$\text{ALSZ}_{\text{szhae}}$	0.2246	-0.0236	0.9767	0.3404
ALDC_{cca}	0.1861	-0.0758	0.9270	0.0313
ALDC_{con}	0.0885	0.1044	1.1101	0.0466
ALDC_{cpx}	0.1159	-0.1928	0.8247	0.0579
ALDC_{txs}	0.1748	0.0367	1.0374	0.1134
$\text{ALHS}_{\text{aadmod}}$	0.1125	0.0503	1.0515	0.0585

Baseline Dynamic Range

The baseline dynamic range of **TrF** are listed below:

TrF estimates	Minimum	Maximum	Median	SD†
$\text{SUV}_{\max}$ (g/mL)	2.40	12.28	6.33	2.34
SUV_{peak} (g/mL)	1.64	10.52	4.86	2.01
SUV_{mean} (g/mL)	1.61	8.39	4.05	1.52
TLV (mL)	0.48	1,182.42	5.71	155.91
TLA (g)	2.11	6,620.88	29.58	878.20
TLE (g^2/mL)	3.55	38,963.30	152.79	5,138.32

†SD: Standard Deviation

Results

⇒ Both **EORTC criteria** & **PERCIST 1.0** resulted in identical subject dichotomization.

⇒ **PERCIST 1.0** gave better C_τ, AUC_{TTP4} & $\text{LOOCV} - C_\tau$ with respect to **EORTC criteria**.

⇒ For both **EORTC criteria** & **PERCIST 1.0**, hazard ratio (**HR**) of 0.71 means responders have a 29% less chance for progression (**TTP**) compared to non-responders.

⇒ However, both **EORTC criteria** & **PERCIST 1.0** based dichotomization are less statistically significant (p=0.743).

⇒ In **pMVFS-max**, a positive Cox regression coefficient indicated that, an increase in max $\delta\text{SUV}_{\text{peak}(\%)}$ for a subject will result in an increase in hazard for progression.

⇒ In **pMVFS-max**, ALRL_{sre} and ALRL_{sp} were found to be statistically significant in the multivariate Cox model.

⇒ In **pMVFS-cum**, ALCC_{idm} obtained higher statistically significance and lower **HR** (95.5% less chance for progression). Hence a subject with a high $\delta\sum\text{ALCC}_{\text{idm}}\%$ will have less chance for progression (or high chance of response to treatment).

⇒ In **pMVFS-cum**, ALCC_{hom} obtained both high **HR** and statistically significance. Hence a subject with a high $\delta\sum\text{ALCC}_{\text{hom}}\%$ will have high chance for progression.

⇒ Both **pMVFS-max** (p=0.0497) and **pMVFS-cum** (p=0.0006) provided statistically significant cohort dichotomization.

⇒ In **pMVFS-max**, responders were having 68% (HR=0.32) less chance for progression compared to non responders.

⇒ In **pMVFS-cum**, responders were having 74% (HR=0.26) less chance for progression compared to non responders.

⇒ Both **pMVFS-max** and **pMVFS-cum** provided better C_τ, AUC_{TTP4} & $\text{LOOCV} - C_\tau$ estimates compared to **EORTC criteria** and **PERCIST 1.0**.

Conclusion

The results with the current data denoted an added value in using textural information for early therapy response evaluation. Moreover, therapy response could be predicted as early as 1 week post treatment using **TrF** & **TxF** feature estimates.

References:

[1] F. Tixier, C. C. Le Rest, M. Hatt, N. Albarghach, O. Pradier, J. P. Metges, L. Corcos, and D. Visvikis, "Intratumor heterogeneity characterized by textural features on baseline ^{18}F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer," *J. Nucl. Med.*, vol. 52, no. 3, pp. 369–378, Mar 2011.

[2] M. Hatt, C. C. Le Rest, A. Turzo, C. Roux, and D. Visvikis, "A fuzzy locally adaptive Bayesian segmentation approach for volume determination in PET," *IEEE Trans. Med. Imag.*, vol. 28, no. 6, pp. 881–893, 2009.

[3] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst., Man and Cybern.*, vol. 3, no. 6, pp. 610–621, nov. 1973.

[4] M. M. Galloway, "Texture analysis using gray level run lengths," *Comput. Graph. Image Process.*, vol. 4, no. 2, pp. 172–179, June 1975.

[5] G. Thibault, B. Fetil, C. Navarro, S. Pereira, P. Cau, N. Levy, J. Sequeira, and J. Mari, "Texture indexes and gray level size zone matrix: application to cell nuclei classification," *Pattern Recognition In Process.*, pp. 140–145, 2009.

[6] M. Amadasun and R. King, "Textural features corresponding to textural properties," *IEEE Trans. Syst., Man and Cybern.*, vol. 19, no. 5, pp. 1264–1274, sep/oct 1989.

[7] P. J. Heagerty, T. Lumley, and M. S. Pepe, "Time-dependent ROC curves for censored survival data and a diagnostic marker," *Biometrics*, vol. 56, no. 2, pp. 337–344, 2000.

[8] P. J. Heagerty and Y. Zheng, "Survival model predictive accuracy and ROC curves," *Biometrics*, vol. 61, no. 1, pp. 92–105, 2005.

[9] H. Young, R. Baum, U. Cremerius, K. Herholz, O. Hoekstra, A.A. Lammertsma, J. Prium, and P. Price, "Measurement of clinical and subclinical tumour response using [^{18}f]-fluorodeoxyglucose and positron emission tomography: review and 1999 EORTC recommendations. European Organization for Research and Treatment of Cancer (EORTC) PET Study Group.," *Eur. J. Cancer*, vol. 35, pp. 1773–1782, 1999.

[10] R. L. Wahl, H. Jacene, Y. Kasamon, and M. A. Lodge, "From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors," *J. Nucl. Med.*, vol. 50, pp. 122S–150S, 2009.

Funding: IBBT, IWT-TBM, FWO

This work is financially supported by KU Leuven's Concerted Research Action GOA/11/006, IWT - TBM project 070717 & Research Foundation - Flanders (FWO).