A Machine-learning Signature Performs as Well as Experts' Consensus to Diagnose Pelvic Lymph Node Involvement in Bladder Cancer on [18F]-FDG PET/CT: A Pathology-controlled Study

Antoine GIRARD (a.girard@rennes.unicancer.fr)
Centre Eugene Marquis  https://orcid.org/0000-0002-9472-9980

Laurent DERCLE
Columbia University Medical Center: Columbia University Irving Medical Center

Helena VILA-REYES
Columbia University Medical Center: Columbia University Irving Medical Center

Lawrence H. SCHWARTZ
Columbia University Medical Center: Columbia University Irving Medical Center

Astrid GIRMA
Hospital Foch - Suresnes: Hopital Foch

Marc BERTAUX
Hospital Foch - Suresnes: Hopital Foch

Camelia RADULESCU
Hospital Foch - Suresnes: Hopital Foch

Thierry LEBRET
Hospital Foch - Suresnes: Hopital Foch

Olivier DELCROIX
University Hospital Centre Brest: CHRU de Brest

Mathieu ROUANNE
Hospital Foch - Suresnes: Hopital Foch

Research Article

Keywords: Bladder cancer, [18F]-FDG PET/CT, Machine Learning, Staging, Artificial Intelligence

DOI: https://doi.org/10.21203/rs.3.rs-520536/v1

License: ☺ ☀ This work is licensed under a Creative Commons Attribution 4.0 International License.  Read Full License
Abstract

Purpose

To develop and validate an objective machine-learning-based signature for the diagnosis of pelvic lymph node (LN) involvement based on \[^{18}\text{F}\]-FDG PET/CT before radical cystectomy in patients with muscle-invasive bladder cancer (MIBC).

Methods

Between 2010-2017, consecutive patients with localized MIBC were retrospectively and randomly assigned to a training (n=115) and a validation (n=58) sets. The reference standard used for LN involvement was pathological status after extended pelvic LN dissection. In the training set a signature was obtained with a random forest algorithm built from 21 morphological and metabolic imaging features. The signature performance was evaluated using areas under ROC curves (AUCs).

Results

The prevalence of pelvic LN involvement was 26% (n=30/115) in the training set, and 21% (n=12/58) in the validation set. In the training set, the top 5 features selected by the algorithm were derived from pelvic LNs (SUVmax, sum of the products of diameters, product of diameters of the largest LN), primary bladder tumor (tumor largest diameter), and urine (SUVmax). In the training set, the AUC of the signature was 0.79 (95%CI 0.71-0.88), versus 0.72 (95%CI 0.62-0.81) with the consensus of experts, p=0.24. In the validation set, AUCs were 0.67 (95%CI 0.50-0.84) and 0.62 (95%CI 0.48-0.77) p=0.70, respectively.

Conclusion

The diagnostic performance of a machine-learning signature was not inferior to the expert consensus review for pelvic LN staging, providing an objective and accurate tool for this task.

Introduction

Bladder cancer is the tenth most common cancer worldwide, with more than half a million new cases diagnosed globally and 200,000 related deaths per year [1]. At diagnosis, approximately 25% of the patients have muscle invasive bladder cancer (MIBC). This diagnosis is achieved through transurethral resection of the bladder tumor (TURBT). MIBC is an aggressive disease characterized by a high rate of nodal and/or distant metastases [2]. Accurate preoperative staging with abdomen CT scan or MRI is standard practice. In patients with clinically localized disease (cT2-T4aN0M0) and adequate renal function, the current standard of care is radical cystectomy (RC) with neoadjuvant cisplatin-based chemotherapy (NAC). For patients with more advanced disease (i.e. node positive and/or solid organ metastases) cytotoxic chemotherapy remains the standard of care for first-line systemic treatment, while immune checkpoint blockade antibodies have been recently approved as first-line treatment in cisplatin-ineligible patients [3].

There is increasing data supporting the role of 2-deoxy-2-\[^{18}\text{F}\]fluoro-D-glucose (\[^{18}\text{F}\]-FDG) positron emission tomography/computed tomography (PET/CT) for the evaluation of metastatic urothelial carcinoma [4–11]. For detection of pelvic lymph node (LN) involvement this modality has a sensitivity and specificity of 50-55% and 92%, respectively [12], though there is high inter-center variability due to heterogeneous interpretation criteria. Machine-learning (ML) techniques are able to use data gathered from PET images to extract objective imaging biomarkers that are otherwise invisible to the naked eye [13]. ML defined diagnostic criteria for pelvic LN staging may improve reproducibility, provided that such criteria do not lower diagnostic performance for the detection of pelvic LN disease spread.

The aim of this study was to compare the diagnostic performance obtained with ML defined criteria based on \[^{18}\text{F}\]-FDG PET/CT to an expert consensus review for initial pelvic LN staging of MIBC.

Materials And Methods

Selection of patients

A total of 173 consecutive patients diagnosed with MIBC between October 2010 and May 2017 (≥pT2 proven by TURBT) were retrospectively included. Inclusion criteria were: ≥18 years of age; history of RC with extended pelvic lymph node dissection (ePLND); available \[^{18}\text{F}\]-FDG PET/CT completed for initial staging ≤90 days prior to surgery. Exclusion criteria were history of previous pelvic or genitourinary cancer and prior neoadjuvant chemotherapy (NAC). This study was approved by an institutional ethical committee (IRB00012437). All the patients included were informed and did not object to participate in this study.
**[18F]-FDG PET/CT protocol**

A 6-h fasting and blood glucose level of <12 mM/L were required before intravenous injection of [18F]-FDG (4 MBq/kg). Furosemide (20 mg) IV was administered immediately after [18F]-FDG infusion in patients without contraindications such as urinary incontinence. Sixty minutes post-injection and after the patient voided, PET and CT images were acquired in supine position, with the arms above the head, from vertex to mid-thigh, on integrated PET/CT systems (Discovery 610 Elite and Discovery STE, GE Healthcare) with iterative reconstruction using CT-based attenuation correction. No contrast agent was used for CT.

**Clinical staging and surgical technique**

In addition to [18F]-FDG PET/CT, preoperative clinical staging included bimanual examination, cystoscopy, and contrast-enhanced CT scan. All patients underwent standardized open radical cystectomy and ePLND. LNs were dissected along the ilio-obturator region and from the internal and external iliac arteries up to the common iliac vessels and the aortic bifurcation, according to the template defined by the European Urology Association [3].

**Pathology analysis**

Pathological review was centralized and performed by a specialized uropathologist (C.R.). Each LN was isolated from the fresh sample and processed according to accepted standards [14]. LNs were fixed overnight in 4% formalin and embedded in paraffin. Samples were sectioned into 3-mm thick tissue slices and stained with hematoxylin, eosin and saffron. Step sectioning to 2.5µm-thick slices were examined for evidence of cancer involvement. The number of LNs positive for metastatic disease and the total number of LNs examined microscopically were reported.

**Image analysis**

All PET/CT images were independently reread by two senior nuclear medicine physicians (A.G and O.D.) who were blinded to the pathological results. They subjectively assessed for metastatic involvement of the LNs in each pelvic area based on both metabolic and morphological criteria. Additional metabolic and morphological features of the bladder tumor, LNs and urine were also described (Supplementary Material 1). In the case of disagreement regarding LN involvement between the two readers, consensus was reached through joint review.

**Statistical analysis**

Patients were randomly assigned to the training (n=115) or the validation (n=58) set. Statistical analyses were performed using R version 4.0.4 (The R Foundation for Statistical Computing), and MedCalc® version 12.5.0.0 (Medcalc Software, Ostend, Belgium). The random forest signature was generated with the “RandomForest” package with a forest of 500 decision trees, and explained with the “randomForestExplainer” package in R. Twenty-one [18F]-FDG PET and CT features were initially assessed (Supplementary Material 1). Areas under the receiver operating characteristics curves (AUCs) for each parameter to predict pathologic pelvic LN involvement were created. Pearson's correlation coefficients were calculated between each pair of features. Statistical significance was set at p=0.05 (two-tailed). In the training and the validation sets, diagnostic performance (AUC) for detection of pelvic LN involvement provided by ML-based criteria was compared to the one obtained by a consensus of experts.

**Results**

**Patients' characteristics**

There were no statistically significant differences in preoperative characteristics between patients in the training and validation sets (Table 1). The mean age of the complete cohort (n=173) was 70 years old, with a standard deviation (SD) of 10 years old, and a male-to-female sex ratio of 7.2. The mean time between [18F]-FDG PET/CT and RC plus ePLND was 31 ± 21 days. The mean number of LNs yielded per patient was 16 ± 8 LNs. LN metastases from bladder cancer were pathologically identified in 42 (24.3%) patients.

**LN staging based a consensus of experts**

Subjective [18F]-FDG PET/CT interpretation in the whole dataset provided a sensitivity of 0.45 [95%CI 0.30-0.61], a specificity of 0.92 [95%CI 0.86-0.96], a positive predictive value of 0.66 [95%CI 0.46-0.82] and a negative predictive value of 0.84 [95%CI 0.77-0.90] for pelvic LN staging.
Signature predicting pathological LN status

Model building in the training set

Several models were created utilizing up to 21 quantitative and qualitative imaging features related to the morphological and metabolic characteristics of the LNs. There was no significant improvement in performance by using more than 5 variables (Fig. 1), presumably due to co-linearity between variables (Fig. 2). The 5 most important variables for the diagnosis of LN involvement as ranked by the ML algorithm were: 1) the product of diameters of the largest pelvic LN; 2) the sum of the products of diameters of all pelvic LNs; 3) the SUVmax of pelvic LNs; 4) SUVmax of urine; and 5) the largest diameter of residual primary tumor (Table 2, Supplementary Material 2). The final signature was obtained from a random forest using regression methods in 500 trees and evaluating the 5 variables mentioned above in each split to explain 63.3% of the variance with a mean of the squared residuals of 0.0092. Using these 5 variables, the signature reached an AUC of 0.79 (95%CI 0.71-0.88) for the prediction of LN status. This was not significantly different (p=0.24) from the performance of consensus of experts, which resulted in an AUC of 0.72 (95%CI 0.62-0.81).

Performance evaluation in the validation set

In the validation set, the signature reached an AUC of 0.67 (95%CI 0.50-0.84) which was not significantly different (p=0.70) from the consensus of experts with an AUC of 0.62 (95%CI 0.48-0.77). The performance of each variable included in the signature, as well as optimal thresholds, are included in Table 3. An example of discordance between subjective and ML-based conclusion for LN staging is described in Fig. 3.

Signature predicting the experts’ conclusion regarding LN status

As an ancillary study, we evaluated the performance of ML algorithms for the prediction of pelvic LN status using as a reference standard the consensus of the two expert radiologists. In doing so we essentially trained the ML algorithm to mimic the interpretation practices of expert radiologists. The model achieving the highest performance in the training set was selected as the signature. The machine learning algorithm ranked the importance of each variable for the prediction of LN status (Supplementary Material 3). In the validation set, performance provided by the ML-based signature provided an AUC of 1.00 (0.99-1.00) to mimic the conclusion from the consensus of experts (Supplementary Material 4).

Discussion

In this study, we created a ML-based signature to predict pelvic LN metastases using preoperative [18F]-FDG PET/CT images in patients with MIBC. The signature was externally validated in an independent data set. Preoperative LN staging with [18F]-FDG PET/CT based on ML-defined criteria showed non-inferior diagnostic performances compared to a consensus review by experts. Specifically, reported AUCs in the validation set were 0.67 with ML-based criteria and 0.62 with a subjective human interpretation. To the best of our knowledge, similar results have not been reported previously in the literature.

Bladder cancer is an aggressive disease, and LN involvement confers a particularly bad prognosis even when only one LN is affected [15]. In clinically localized tumors (pT2-4 pNO cM0), 10-year cancer-specific survival after RC without NAC or adjuvant chemotherapy may reach 77%, whereas LN involvement (pN1-3) dramatically drops it to 17% [15]. Therefore, an accurate detection of LN metastases at diagnosis has a major impact on treatment decisions and future prognosis. Current guidelines recommendations [3] are in favor of targeting micrometastases by administering NAC to all cisplatin-eligible patients who have clinically localized MIBC. However, these guidelines still suffer from poor adherence in routine clinical practice [16]. This non-compliance with recommended practice is mostly the result of the reported low benefit of NAC in terms of survival for ≤pT2N0 tumors [17], as well as the limited use of highly reliable imaging tools for nodal staging. Due to the modest ability of morphological imaging to detect LN involvement, [18F]-FDG PET/CT may play a decisive role in patient management [5]. However, heterogeneous interpretation criteria among institutions hamper the reproducibility of [18F]-FDG PET/CT, and thus limits its widespread use.

In the present study, an ML-based signature and an expert consensus diagnostic performance did not show statistically significant difference for LN staging using [18F]-FDG PET/CT. These data suggest that ML-based signatures may improve the reproducibility of [18F]-FDG PET/CT image interpretation without loss of performance. Moreover, the developed ML-based signature could increase accuracy in non-expert centers it can complement subjective human interpretation.

This signature relied on five selected top variables measured from pelvic LNs (the SUVmax of pelvic LNs, the product of diameters of the largest pelvic LN, and the sum of the products of diameters of all pelvic LNs), the primary bladder tumor (the largest diameter of residual primary tumor after TURBT), and the urine (SUVmax of urine). Metabolic (SUVmax) and morphological (product of diameters of the largest
LN) characteristics of pelvic LNs have already been reported in several studies as objective criteria for LN staging with [18F]-FDG PET/CT [6,18]. To our knowledge, the sum of the products of diameters of pelvic LNs has not previously been reported in the literature as a criterion of interest. It nevertheless seems logical that this variable related to the pelvic LN tumor volume was included in the top 5 features. Interestingly, the two remaining top 5 features are not directly measured from pelvic LN. First, the largest diameter of the primary bladder tumor was a selected predictor, and patients with primary tumor larger than 49-57 mm had higher risk of pelvic LN involvement. This is in line with previously published studies that reported the primary bladder tumor size to be associated with the risk of LN involvement [19] and distant metastases [20]. Of note, the largest diameter of primary tumor was measured when visible on PET and/or non-contrast enhanced CT from [18F]-FDG PET/CT performed after TURBT. Thus, it may slightly differ from the total tumor size before resection. Finally, the radioactivity of urine (SUVmax of urine) was unexpectedly one of the top five features. It does not have an obvious pathophysiological explanation but might be explained by the PET protocol itself. This latter included diuretic administration for all patients with no urine incontinence (45% of patients in this study), to lower urine radioactivity due to physiological urinary excretion of [18F]-FDG in order to optimize detection of primary tumor and pelvic LN involvement. Thus, patients with continence problems did not receive diuretic, and as a consequence had higher [18F]-FDG concentration in urine. One hypothesis is that patients with advanced disease might be more often subject to continence disorders, did not receive diuretic before [18F]-FDG PET/CT, and thus might have had a higher SUVmax in urine.

Additionally, we developed another ML-based signature to mimic the conclusion of a consensus of two expert radiologists. Interestingly, the top 5 selected variables differed from the ones used to predict pathological pelvic LN involvement. Metabolic characteristics of pelvic LNs (pelvic LN [18F]-FDG uptake higher than background, and their SUVmax) was deemed most important by the experts, followed by the size of the tumor (small diameter of residual primary tumor (mm)) and of the pelvic LN (small diameter of the largest pelvic LN, and the product of diameters of the largest pelvic LN). Interestingly, metabolic characteristics of pelvic LNs were by far the most important factors to state the conclusion of experts.

This study has several limitations. First, the retrospective design of the study and the moderate size of the sample may hamper the statistical power of the comparison between the ML signature and the experts consensus. However, this is the largest cohort of MIBC patients studied with [18F]-FDG PET/CT who did not receive NAC, which allowed accurate matching of LNs between imaging and to pathological analysis. Second, emphasis should be given to the imbalance between positive (pN+) and negative (pN-) subjects in both datasets, related to the small number of events (n=42). This has led us to underfit the ML model by using a limited subset of features. The prevalence of pathological pelvic LN involvement in this sample (24%) was in the same range as those reported in the literature [15]. Finally, the size of the tumor has been described as one of the most relevant factors for predicting the LN positivity. However, the tumor size measurement included only the remaining tumor visible on PET/CT performed after TURBT, rather than the whole tumor. This probably had a limited impact on the results since the exophytic volume of the tumor is not related with its degree of invasion. Additionally, measuring only the remaining tumor accurately reflects real clinical practice wherein the tumor is initially resected partially or completely.

Conclusions

In this study, we developed a ML-based signature for the diagnostic of pelvic LN metastasis with [18F]-FDG PET/CT imaging in patients with MIBC. This signature relied on five key metabolic and morphological features related to pelvic LNs, primary tumor and urine. In an independent validation set, the diagnostic performance of the signature was found to be non-inferior to a consensus review by experts. Such an objective ML-based interpretation tool provides good accuracy and may enhance reproducibility of [18F]-FDG PET/CT interpretation for pelvic LN staging in MIBC.

Declarations

Funding: Not applicable

Conflicts of interest/Competing interests: The authors do not have any conflict of interest to disclose

Availability of data and material: Available by request to the corresponding author

Code availability: Not applicable

Ethics approval: This retrospective study was approved by an institutional review board (IRB00012437).

Consent to participate/Consent for publication: Since this study was retrospective, each participant received written information about the study and publications and was offered the opportunity to refuse to participate at any time.
Acknowledgements:

We would like to thank Dr. G. Joel DeCastro who accepted to review and relevantly edit this article.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicians. 2018;68:394–424.
2. Kamat AM, Hahn NM, Efstathiou JA, Lerner SP Malmström P-U, Choi W, et al. Bladder cancer. Lancet. 2016;388:2796–810.
3. Witjes JA, Bruins HM, Cathomas R, Compérat EM, Cowan NC, Gakis G, et al. European Association of Urology Guidelines on Muscle-invasive and Metastatic Bladder Cancer: Summary of the 2020 Guidelines. Eur Urol. 2020;
4. Salau̇n P-Y, Abgral R, Malard O, Querellou-Lefranc S, Quere G, Wartski M, et al. Good clinical practice recommendations for the use of PET/CT in oncology. Eur J Nucl Med Mol Imaging. 2020;47:28–50.
5. Voskuilen CS, van Gennep EJ, Einerhand SMH, Vegt E, Donswijk ML, Bruining A, et al. Staging 18F-fluorodeoxyglucose Positron Emission Tomography/Computed Tomography Changes Treatment Recommendation in Invasive Bladder Cancer. European Urology Oncology. 2021;S2588931121000298.
6. Girard A, Rouanne M, Taconet S, Radulescu C, Neuzillet Y, Girma A, et al. Integrated analysis of 18F-FDG PET/CT improves preoperative lymph node staging for patients invasive bladder cancer. Eur Radiol. 2019;29:4286–93.
7. Rouanne M, Girma A, Neuzillet Y, Vilain D, Radulescu C, Letang N, et al. Potential impact of 18F-FDG PET/CT on patients selection for neoadjuvant chemotherapy before radical cystectomy. European Journal of Surgical Oncology (EJSO). 2014;40:1724–30.
8. Goodfellow H, Viney Z, Hughes P, Rankin S, Rottenberg G, Hughes S, et al. Role of fluorodeoxyglucose positron emission tomography (FDG PET)-computed tomography (CT) in the staging of bladder cancer: FDG pet in the staging of bladder cancer. BJU Int. 2014;n/a-n/a.
9. Hitier-Berthault M, Ansquer C, Branchereau J, Renaudin K, Bodere F, Bouchot O, et al. 18F-fluorodeoxyglucose positron emission tomography-computed tomography for preoperative lymph node staging in patients undergoing radical cystectomy for bladder cancer: A prospective study: PET scan for bladder cancer. Int J Urol. 2013;20:788–96.
10. Swinnen G, Maes A, Pottel H, Vanneste A, Billiet I, Lesage K, et al. FDG-PET/CT for the Preoperative Lymph Node Staging of Invasive Bladder Cancer. European Urology. 2010;57:641–7.
11. Kibel AS, Dehdashti F, Katz MD, Klim AP, Grubb RL, Humphrey PA, et al. Prospective Study of [18 F]Fluorodeoxyglucose Positron Emission Tomography/Computed Tomography for Staging of Muscle-Invasive Bladder Carcinoma. JCO. 2009;27:4314–20.
12. Girard A, Vila Reyes H, Shaish H, Grellier J-F, Derde L, Salaün P-Y, et al. The Role of 18F-FDG PET/CT in Guiding Precision Medicine for Invasive Bladder Carcinoma. Front Oncol. 2020;10:565086.
13. Aktolun C. Artificial intelligence and radiomics in nuclear medicine: potentials and challenges. Eur J Nucl Med Mol Imaging. 2019;46:2731–6.
14. ADASP Committee. The Association of Directors of Anatomic and Surgical Pathology. ADASP recommendations for processing and reporting of lymph node specimens submitted for evaluation of metastatic disease. Mod Pathol. 2001;14:629–32.
15. Hautmann RE, de Petriconi RC, Pfeiffer C, Volkmer BG. Radical Cystectomy for Urothelial Carcinoma of the Bladder Without Neoadjuvant or Adjuvant Therapy: Long-Term Results in 1100 Patients. European Urology. 2012;61:1039–47.
16. Martini T, Gilfich C, Mayr R, Burger M, Pycha A, Aziz A, et al. The Use of Neoadjuvant Chemotherapy in Patients With Urothelial Carcinoma of the Bladder: Current Practice Among Clinicians. Clinical Genitourinary Cancer. 2017;15:356–62.
17. Niegisch G, Lorch A, Droller MJ, Lavrey HJ, Stensland KD, Albers P. Neoadjuvant Chemotherapy in Patients with Muscle-invasive Bladder Cancer: Which Patients Benefit? European Urology. 2013;64:355–7.
18. Aljabery F, Lindblom G, Skoog S, Shabo l, Olsson H, Rosell J, et al. PET/CT versus conventional CT for detection of lymph node metastases in patients with locally advanced bladder cancer. BMC Urol. 2015;15:87.
19. Ou N, Song Y, Liu M, Zhu J, Yang Y, Liu X. Development and Validation of a Nomogram to Predict Lymph Node Metastasis in Patients With T1 High-Grade Urothelial Carcinoma of the Bladder. Front Oncol. 2020;10:532924.
20. Cheng L, Neumann RM, Scherer BG, Weaver AL, Leibovich BC, Nehra A, et al. Tumor size predicts the survival of patients with pathologic stage T2 bladder carcinoma: a critical evaluation of the depth of muscle invasion. Cancer. 1999;85:2638–47.

Tables
Table 1. Patients’ characteristics

| Characteristics                                      | Training set | Validation set | p value |
|------------------------------------------------------|--------------|----------------|---------|
| Number of patients                                   | 115          | 58             |         |
| Age (years), \( mean \pm SD \)                       | 68.9 ± 10.7  | 71.1 ± 8.6     | 0.15    |
| Sex (males)                                          | 100 (87.0%)  | 52 (89.7%)     | 0.79    |
| Body mass index (kg/m\(^2\)), \( mean \pm SD \)     | 25.9 ± 4.2   | 25.3 ± 4.7     | 0.38    |

\[^{18}F\]-FDG PET/CT parameters

- Injection of diuretic
- Injected activity (MBq), \( mean \pm SD \)

|                  | Training set | Validation set | p value |
|------------------|--------------|----------------|---------|
| Injection of diuretic |              |                |         |
| Injected activity (MBq) |              |                |         |
| 53 (46.1%)       | 25 (43.1%)   | 0.83           |
| 338 ± 65         | 321 ± 83     | 0.18           |

Pathology results

- pT stage on RC sample:
  - \( \leq T1 \)
  - T2
  - T3
  - T4a

|                  | Training set | Validation set | p value |
|------------------|--------------|----------------|---------|
| \( \leq T1 \)     | 36 (31.3%)   | 17 (29.3%)     | 0.55    |
| T2               | 22 (19.1%)   | 17 (29.3%)     |         |
| T3               | 43 (37.4%)   | 19 (32.8%)     |         |
| T4a              | 14 (12.2%)   | 5 (8.6%)       |         |

- WHO 2016 classification:
  - Urothelial cell carcinoma
  - Other histology

|                  | Training set | Validation set | p value |
|------------------|--------------|----------------|---------|
| Urothelial cell carcinoma |              |                | 0.92    |
| Other histology   | 111 (96.5%)  | 55 (94.9%)     |         |
|                  | 4 (3.5%)     | 3 (5.1%)       |         |

- LN status on pelvic LN dissection sample:
  - Number of pelvic LNs dissected (per patient), \( mean \pm SD \)
  - pN stage on extended pelvic LN dissection sample
  - pN0
  - pN1
  - pN2
  - pN3
  - LN density (N+/N) among patients with LN involvement, \( mean \pm SD \)

|                  | Training set | Validation set | p value |
|------------------|--------------|----------------|---------|
| Number of pelvic LNs dissected (per patient) | 16 ± 8       | 16 ± 7         | 0.93    |
| pN0              | 85 (73.9%)   | 46 (79.3%)     | 0.43    |
| pN1              | 16 (13.9%)   | 4 (6.9%)       |         |
| pN2              | 9 (7.8%)     | 5 (9.6%)       |         |
| pN3              | 5 (4.4%)     | 3 (5.2%)       |         |
| LN density (N+/N) among patients with LN involvement | 20% ± 22%    | 23% ± 28%      | 0.70    |

\[^{18}F\]-FDG PET/CT to surgery (days), \( mean \pm SD \)

|                  | Training set | Validation set | p value |
|------------------|--------------|----------------|---------|
| \[^{18}F\]-FDG PET/CT to surgery (days) | 30 ± 22      | 32 ± 18        | 0.45    |

Note: SD standard-deviation, RC radical cystectomy, LN lymph node

Table 2. Metrics outputted by the random forest algorithm allowing for the quantification of the importance of each variable in order to predict lymph node status on pathology
Table 2 legend. A variable of importance has lower mean minimal depth, higher number of nodes, lower mean squared error increase, higher node purity increase, higher number of trees, and higher times at root.

Table 3 legend. Our primary endpoint was to evaluate the performance of the signature using AUC. The interest of imaging biomarkers is that they are continuous values, hence they class patients along a continuum of probability to be positive on pathology. The aim of this study was not to determine a specific decision threshold since in practice, some clinicians may emphasize the importance of high sensitivity vs. high specificity. We compared the optimal thresholds to guide clinical decisions. Using optimal threshold in both datasets, we see that while some values are “clearly” positive and “clearly negative”, there is a “grey area” for most biomarkers such as LN SUVmax (1.35-1.98) and Tumor largest diameter (48.8mm-56.8mm).

| Variables                                | TRAINING SET (n=115) | VALIDATION SET (n=58) |
|------------------------------------------|----------------------|-----------------------|
|                                          | AUC (%) (95%CI)      | Sensitivity (%)       | Specificity (%) | AUC (%) (95%CI) | Sensitivity (%) | Specificity (%) |
| LN_SUVmax                                | 72.4 (60.1-84.7)     | 1.98                  | 89.4            | 57.1 (37.5-76.6)| 1.35            | 54.3            | 66.7            |
| Urine_SUVmax                             | 59.5 (48.1-70.9)     | 35.7                  | 80.0            | 72.5 (56.7-88.2)| 10.1            | 76.1            | 66.7            |
| LN_Sum_Product_of_diameters (mm^2)       | 62.5 (53.2-77.2)     | 76.7                  | 63.5            | 66.1 (44.7-87.5)| 170.3           | 78.3            | 66.7            |
| LN_Max_Product_of_diameters (mm^2)       | 65.0 (52.9-77.1)     | 71.8                  | 68.2            | 58.1 (37.1-79.0)| 153.8           | 93.5            | 33.3            |
| T_Largest_diameter (mm)                  | 65.6 (53.6-77.6)     | 48.8                  | 70.6            | 63.7 (45.8-81.6)| 56.8            | 82.6            | 50.0            |
| SIGNATURE                                | 79.2 (70.5-87.9)     | 0.16                  | 62.3            | 66.7 (49.7-83.6)| 0.30            | 73.9            | 58.3            |

*values higher than “optimal thresholds” were associated with higher risk of pathological LN involvement.
Patients with a diagnosis of muscle-invasive bladder cancer with [18F]-FDG PET/CT images acquired before radical cystectomy n=173 pts

Training set n=115 pts

Validation set n=58 pts

Validation set: Signature validation

1

Extraction of image features n=21 features

Training set: Signature selection n=5 features

2

BLADDER TUMOR
Tumor largest diameter

URINE
SUVmax

LYMPH NODES
SUVmax
Sum of the products of diameters
Product of diameters of the largest lymph node

3

Prediction of pelvic lymph node involvement

Figure 1
Workflow
Figure 2

Correlogram (Pearson coefficient). Paired correlation with p-value > 0.05 are hidden.
Figure 3

[18F]-FDG PET/CT images from a patient who was falsely classified cN+ based on subjective human interpretation, whereas he was classified cN0 with machine learning-based pelvic LN staging. Primary tumor uptake was SUVmax 7.6 with 116 mm of largest axial diameter, and one pelvic lymph node was visible in the left external iliac area (red arrow head) with moderate uptake of SUVmax 2.0 and 10 mm x 7 mm axial diameters. SUVmax of urine was 5.8.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SM2.tif
- SM4.tif
- SupplementaryMaterial1.docx
- SupplementaryMaterial3.docx