# Bladder Cancer Staging in CT Urography: Effect of Stage Labels on Statistical Modeling of a Decision Support System

Dhanuj Gandikota, Lubomir Hadjiiski, Kenny H. Cha, Heang-Ping Chan, Elaine M. Caoili, Richard H. Cohan, Alon Weizer, Ajjai Alva, Chintana Paramagul, Jun Wei, Chuan Zhou

Department of Radiology, University of Michigan, Ann Arbor, MI 48109

## ABSTRACT

In bladder cancer, stage T2 is an important threshold in the decision of administering neoadjuvant chemotherapy. Our long-term goal is to develop a quantitative computerized decision support system (CDSS-S) to aid clinicians in accurate staging. In this study, we examined the effect of stage labels of the training samples on modeling such a system. We used a data set of 84 bladder cancers imaged with CT Urography (CTU). At clinical staging prior to treatment, 43 lesions were staged as below stage T2 and 41 were stage T2 or above. After cystectomy and pathological staging that is considered the gold standard, 10 of the lesions were upstaged to stage T2 or above. After correcting the stage labels, 33 lesions were below stage T2, and 51 were stage T2 or above. For the CDSS-S, the lesions were segmented using our AI-CALS method and radiomic features were extracted. We trained a linear discriminant analysis (LDA) classifier with leave-one-case-out cross validation to distinguish between bladder lesions of stage T2 or above and those below stage T2. The CDSS-S was trained and tested with the corrected post-cystectomy labels, and as a comparison, CDSS-S was also trained with understaged pre-treatment labels and tested on lesions with corrected labels. The test AUC for the CDSS-S trained with corrected labels was 0.89 ± 0.04. For the CDSS-S trained with understaged pre-treatment labels and tested on the lesions with corrected labels, the test AUC was 0.86 ± 0.04. The likelihood of stage T2 or above for 9 out of the 10 understaged lesions was correctly increased for the CDSS-S trained with corrected labels. The CDSS-S is sensitive to the accuracy of stage labeling. The CDSS-S trained with correct labels shows promise in prediction of the bladder cancer stage.

**Keywords:** Bladder Cancer Staging, Radiomics, Classification, Segmentation, Computer-Aided Diagnosis, CT Urography, Feature Extraction.

## 1. INTRODUCTION

Urinary bladder cancer is of the most common forms of cancer that affects both men and women. In 2017 there were an estimated 79,030 new cases of bladder cancer with 16,780 estimated deaths [1]. The correct staging of bladder cancer is crucial for the decision of administering neoadjuvant chemotherapy.

The treatment of neoadjuvant chemotherapy is only recommended for patients who are clinically diagnosed with a stage T2-T4 muscle invasive operable urothelial carcinoma of the bladder. Multiple practices are utilized in the clinical staging to determine the diagnosis of the bladder cancer. Clinical staging not only includes a comprehensive physical examination, but resection bimanual examination while under anesthesia and pathology information derived from TURBT (transurethral resection of bladder tumor). Cross-sectional imaging of the bladder is also often performed and used in the diagnosis. A CTU slice with bladder cancer is shown in Figure 1. A patient's pre-existing conditions of poor renal function may exclude a patient from chemotherapy. Otherwise, neoadjuvant chemotherapy is recommended when muscle invasion is present in the pathology, a significant mass is discovered in the bimanual exam, and/or cross-sectional imaging indicates a disease of T2 or greater [2-5]. At clinical staging, approximately 30% of patients are under-staged or over-staged. The inaccuracy in staging may be partly attributed to the subjectivity and variability of clinicians in utilizing available diagnostic information. In our previous pilot study, we have developed a decision support system for staging of bladder cancer trained with the pre-treatment clinical staging information [6-8].

The pathological staging after cystectomy (the removal of the bladder) is considered the gold standard in current clinical practice. An objective decision support system trained with the correct staging information may be useful for assisting clinicians in making more accurate and consistent staging assessments. Our goal is to develop a quantitative computerized decision support system (CDSS-S) that may be used to aid clinicians in staging of bladder cancer in CT Urography (CTU) and can reduce the diagnostic error during clinical staging.

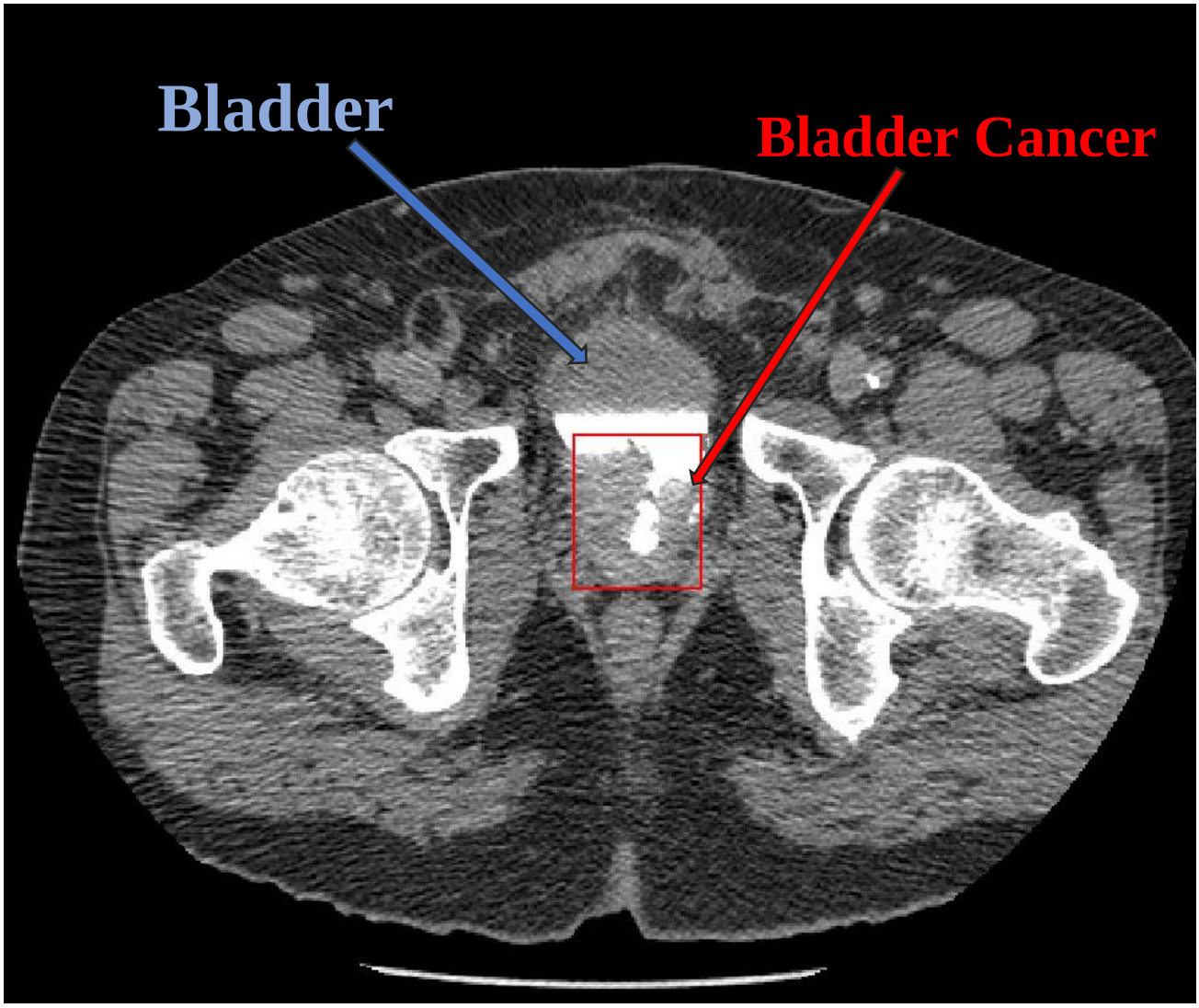


**Figure 1**: Cross-sectional image, CT slice, of the urinary bladder. The bladder cancer (Stage T1) is enclosed by the red box provided by an experienced abdominal radiologist.

## 2. MATERIALS AND METHODS

### 2.1 Data set

We used a data set of 84 bladder cancers from 76 CTU examinations that were all clinically staged prior to treatment. Each lesion was categorized into to one of two classes: (1) below stage T2 or (2) at or above stage T2. Of the 84 bladder cancers, 43 of the lesions were staged as below stage T2, and 41 were stage T2 or above. After cystectomy and pathological staging, 10 of the 84 bladder cancers were upstaged to stage T2 or above. We corrected the labels for these 10 understaged lesions. After correcting the stage labels for the understaged lesions, 33 of the 84 lesions were below stage T2, and 51 were stage T2 or above.

### 2.2 Segmentation and Feature Extraction of Bladder Cancer Lesions

Segmentation of bladder lesions poses many challenges as lesions can often be small and have noisy boundaries. In addition, a lesion located in the bladder's non-contrast-enhanced region can be difficult to distinguish with high accuracy. In our approach, the CTU volume with a bounding box enclosing the lesion provided by radiologist is used as input for our previously developed auto-initialized cascaded level sets (AI-CALS) segmentation pipeline. The input volume of interest is pre-processed and the initial contour is estimated. The initial contour is used to initialize the cascaded 3-D/2-D level set and a 3-D lesion outline is obtained as shown in Figure 2. Eighty-nine radiomics features are extracted from each of the 84 segmented lesions. The feature set includes 24 gray level descriptors, the lesion volume, and 64 shape descriptors.

### 2.3 Classification

A leave-one-case-out cross-validation was conducted with the 76 cases. Feature selection was performed to identify key features within the data set. Three features were selected including the Fourier descriptor, texture, and gradient magnitude profile features. We trained a linear discriminant analysis (LDA) classifier to distinguish between bladder lesions that were diagnosed as stage T2 or above and those below stage T2. For each training/test leave-one-case-out cycle, the trained classifier outputted the likelihood of a lesion being stage T2 or above for the lesions in the left-out test case.

As a comparison to our CDSS-S trained and tested with the corrected post-cystectomy labels, a second CDSS-S was trained with the understaged pre-treatment labels and tested on the lesions corrected for the understaged labels. Again, feature selection and leave-one-case-out cross-validation were performed for training and testing the LDA classifier. The selected features were texture, gradient magnitude, and gradient magnitude profile features.

The CDSS-S performance was evaluated by the receiver operating characteristics (ROC) analysis, and the classification performance was quantified by the area under the ROC curve (AUC).

## 3. RESULTS

For the CDSS-S trained with the corrected stage labels, the test AUC was $0.89 \pm 0.04$ (Table 1 and Figure 3). For the CDSS-S that was trained with the understaged pre-treatment labels and tested on the data set corrected for the understaged lesions, the test AUC was $0.86 \pm 0.04$ (Table 2 and Figure 3). The differences did not reach statistical significance ($p > 0.05$). When the CDSS-S was trained with the corrected stage labels, the CDSS-S produced correctly higher likelihood scores of stage T2 or above for 9 of the 10 understaged lesions. Examples of CTU scans of bladder cancers with stages $\geq$ T2 or $<$ T2 and the corresponding CDSS-S scores are presented in Figure 2.

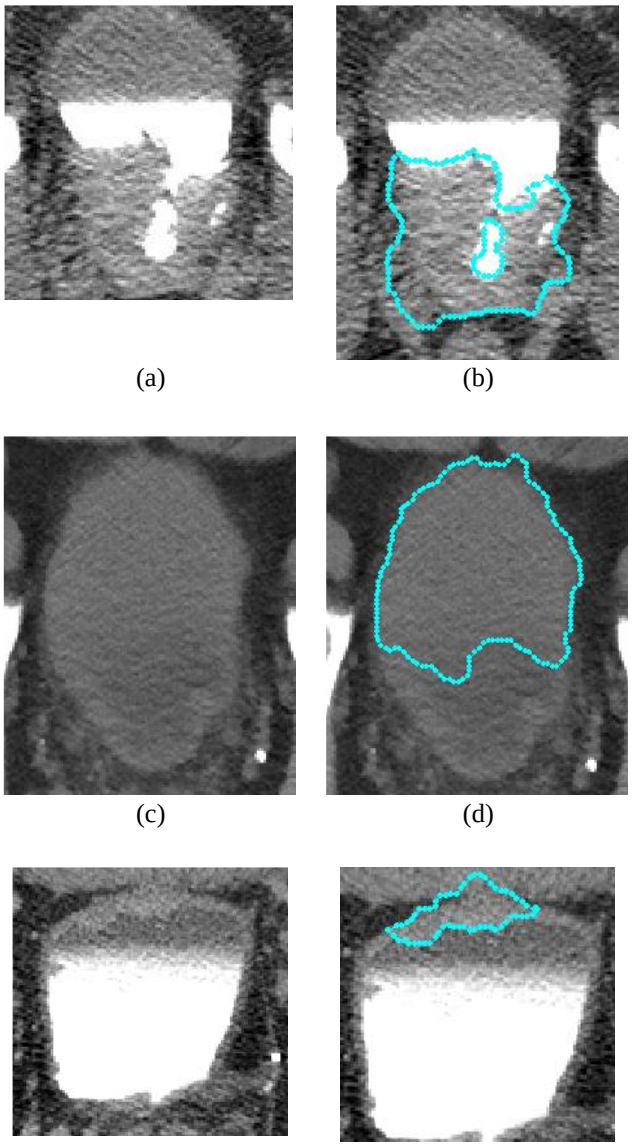


**Figure 2**: Examples of CTU scans of bladder cancers with stages ≥ T2 or < T2. The blue outlines are the AI-CALS segmentation. The following reported scores are the test scores. (a) and (b) show a T1 stage cancer that was properly classified with a score of -4.47. (c) and (d) show a T2 stage case that was properly classified with a score of 1.37. (e) and (f) show a case that was clinically identified as Ta (< T1) pre-treatment but was identified as a T2 stage cancer post-cystectomy. The classifier trained with understaged labels identified the cancer as <T2 with a score of -1.11 and the classifier trained with corrected labels identified the cancer correctly as ≥T2 with a higher score of 0.094

**Table 1**. Summary table of the classification results from the leave-one-case-out experiment trained with corrected stage labels.

| Data Partitions | AUC |
|---|---|
| Training (Corrected Stage Labels) | 0.89 |
| Testing | 0.89 |

**Table 2**. Summary table of the classification results from the leave-one-case-out experiment trained with understaged pre-treatment labels.

| Data Partitions | AUC |
|---|---|
| Training (Pre-treatment labels) | 0.94 |
| Testing | 0.86 |

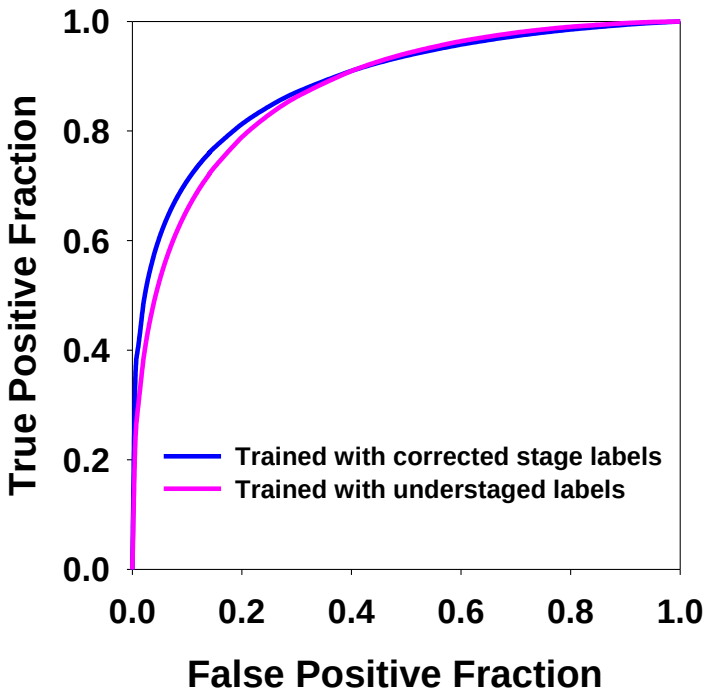


**Figure 3**: Test ROC curves for computerized decision support system (CDSS-S) that was trained with corrected stage labels (AUC = 0.89 ± 0.04) and the CDSS-S trained with the understaged labels (AUC = 0.86 ± 0.04).

## 4. CONCLUSION

We demonstrate the promise of using radiomic features automatically extracted from CTU and correct staging information to build a statistical predictive model for staging of bladder cancer. The improvement of the AUC scores when the decision support system was trained with more accurate labels affirms the importance of collecting reliable training data for statistical modeling. Further work includes improving the radiomics features and the predictive model through the inclusion of clinical and molecular data and the collection of a larger data set.


## ACKNOWLEDGMENT

This work is supported by National Institutes of Health Grant number U01CA179106.


## REFERENCES


[1] American Cancer Society, www.cancer.org 2017, "Cancer Facts & Figures 2017" (2017).

[2] Grossman H. B., Natale R. B., Tangen C. M., et al. "Neoadjuvant chemotherapy plus cystectomy compared with cystectomy alone for locally advanced bladder cancer," N Engl J Med 349, 859-866 (2003).

[3] Sternberg C. N., "The treatment of advanced bladder cancer," Ann Oncol 6, 113-126 (1995).

[4] Hadjiiski L., Mukherji S. K., Gujar S. K., et al. "Treatment response assessment of head and neck cancers on CT using computerized volume analysis," AJNR 31, 1744-1751 (2010).

[5] Meeks JJ., Bellmunt J., Bochner BH., et al. "A systematic review of neoadjuvant and adjuvant chemotherapy for muscle invasive bladder cancer," Eur Urol 62, 5230533 (2012).

[6] Hadjiiski L., Chan H-P., Caoili E. M., Cohan R. H., Wei J., Zhou C. "Auto-initialized cascaded level set (AI-CALS) segmentation of bladder lesions on multi-detector row CT urography," Acad Radiol 20, 148-155 (2013).

[7] Garapati SS, Hadjiiski L, Cha K, Chan HP, Caoili EM, Cohan RH, Weizer A, Alva A, Paramagul C, Wei J, Zhou C. “Automatic staging of bladder cancer on CT urography.” Proc. SPIE Medical Imaging; 9785: 97851G-1,6 (2016).

[8] Garapati SS, Hadjiiski L, Cha KH, Chan HP, Caoili EM, Cohan RH, Weizer A, Alva A, Paramagul C, Wei J, Zhou C. “Urinary Bladder Cancer Staging in CT Urography using Machine Learning.” Med Phys; 44(11): 5814-5823 (2017).