



## Natural Language Processing for Automated Correlation of Radiology and Pathology Reports in Prostate Cancer Detection

Anthony T. Wu, University of California, Irvine; Gavin Shu, MS; Ryan O'Connell, MD; Peter Chang, MD; Robert Edwards, MD, PhD; Sungmee Park, MD; Roozbeh Houshyar, MD

### Introduction/Background

Prostate cancer (PCa) is the second most common cancer among men and has the second highest mortality rate. With an aging U.S. population, the demand for accurate PCa detection is outpacing the diagnostic radiology workforce, particularly in remote areas. Though data-driven self-improvement in radiologist PCa detection can help alleviate workforce constraints, automated pathology correlation to prostate MRI reports (RRs) is currently lacking, limiting this approach's feasibility. Herein we propose a novel natural language processing algorithm (NLP) for automated radiology-pathology report correlation using a 12-core biopsy template (12cBT) in real time.

### Methods/Intervention

Radiology reports (RRs) and their corresponding pathology reports (PRs) from UCI Health (October 2013-October 2023) were retrieved from a HIPAA-compliant data warehouse, totaling 1162 pairs across 1093 patients. A random 10% subset was labeled by medical students and verified by physicians. RRs and PRs were analyzed separately (Figure 1). Regex expressions extracted lesion locations and PI-RADS scores from RRs (Figure 2a), while core biopsy regions and Gleason scores were extracted from PRs (Figure 2b). A custom spell-check addressed domain-specific errors (Figure 1). Lesions were mapped to 12cBTs, and with NLP mapping performance evaluated on the test set.

### Results/Outcome

The NLP achieved 97.4% accuracy in detecting significant PI-RADs ( $\geq 3$ ) in RRs and 100% accuracy in detecting significant Gleason scores ( $\geq 3+3$ ) in PRs. Mapping of 12cBT regions for RRs and PRs yielded 89.6% and 89.4% overall accuracy, respectively (Figure 3a). PI-RADS v2 demonstrated 60.8% accuracy in detecting PCa, with 12cBT regional sensitivities shown in Figure 3b.

### Conclusion

Our NLP system effectively mapped radiology reports (RRs) to pathology reports (PRs) in near-real time. To our knowledge, this is the first instance of an automated radiology-pathology correlation for PCa. In addition, we found that while radiologists faced challenges in pinpointing exact PCa locations, they were able to detect the general region of PCa with relatively high sensitivity.

### Statement of Impact

Our tool provides radiologists with a means for self-improvement by delivering feedback as soon as pathology reports are available if integrated with hospital electronic health record systems. Additionally, we assessed the performance of PI-RADS v2 at UCI Health in detecting PCa across over 1000 reports and patients.

**Fig 1**

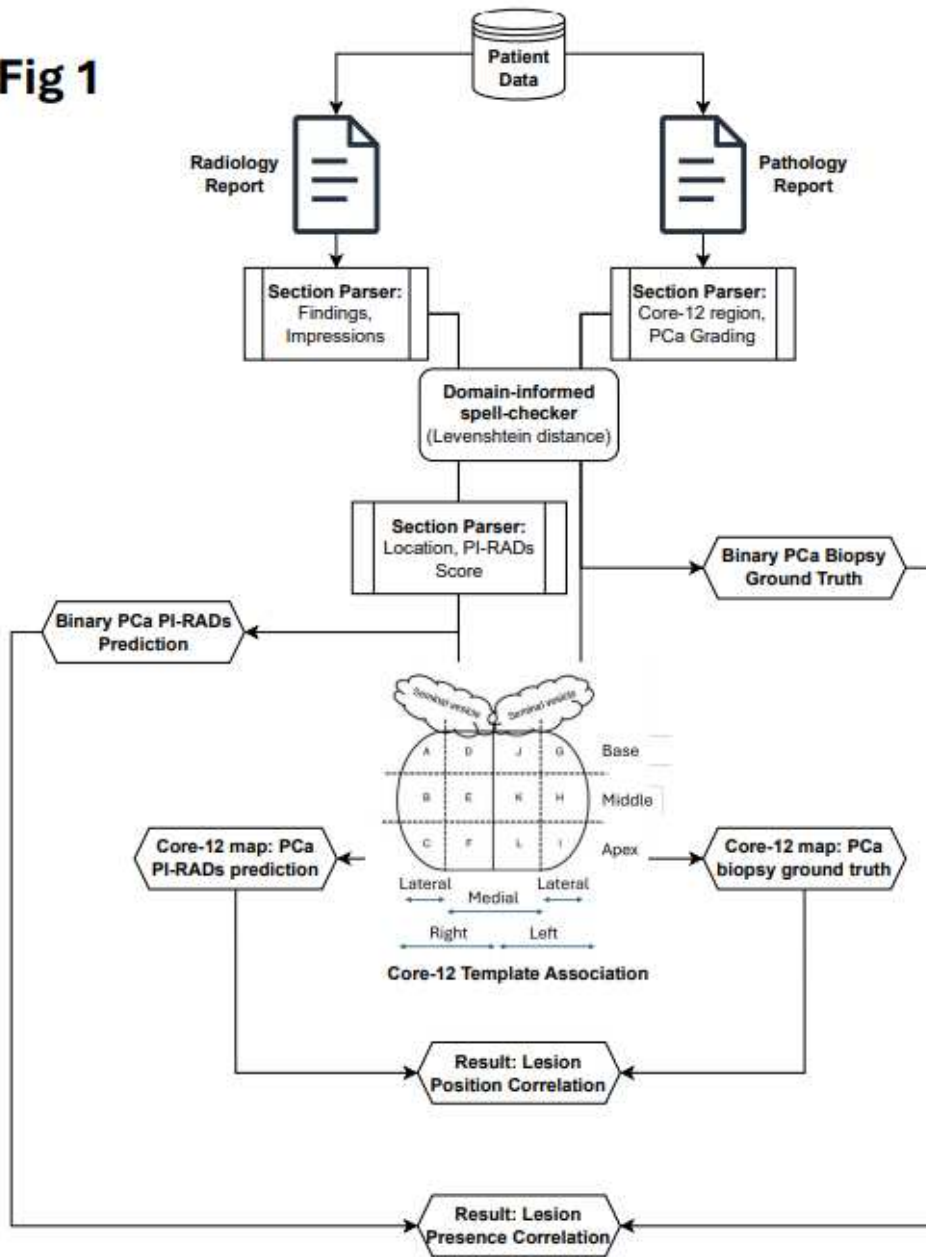


Figure 1: overall pipeline of radiology-pathology report correlation. Note that inferior/posterior keywords in radiology reports are ignored, as biopsies run along that axis. The peripheral zone is designated as sectors A, B, C, F, L, G, H, and I, while transition and central zones are designated as sectors D, E, J and K. Prostate figure borrowed from You et. al's publication (DOI: 10.14366/usg.15065) under the Creative Commons Attribution Non-Commercial License. Research reported in this publication was supported by the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health under Award Number T35DK128788 and the Radiological Society of North America Research & Education Foundation 2023 Medical Student Grant. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or Radiological Society of North America.

Fig 2a

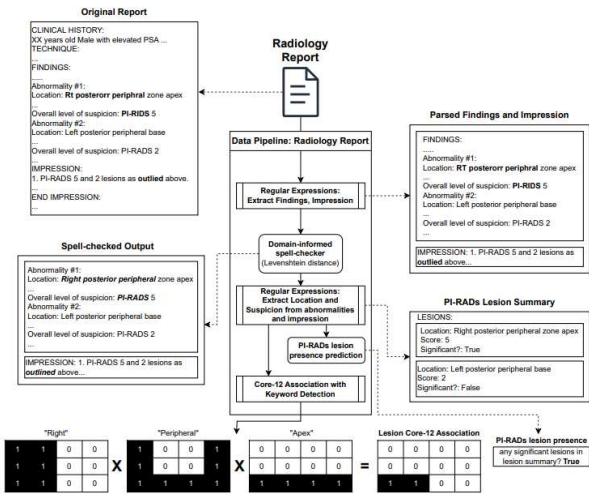


Fig 2b

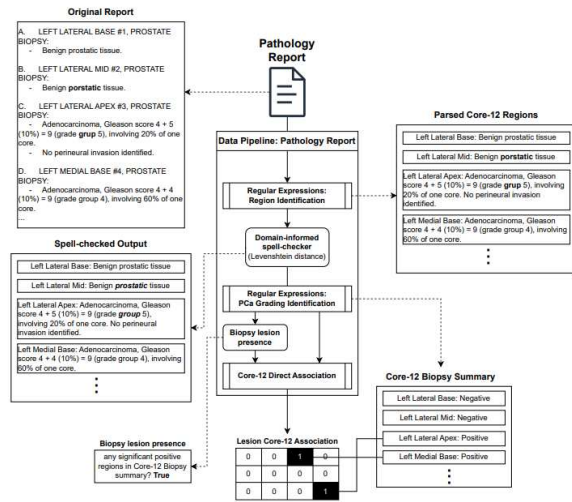
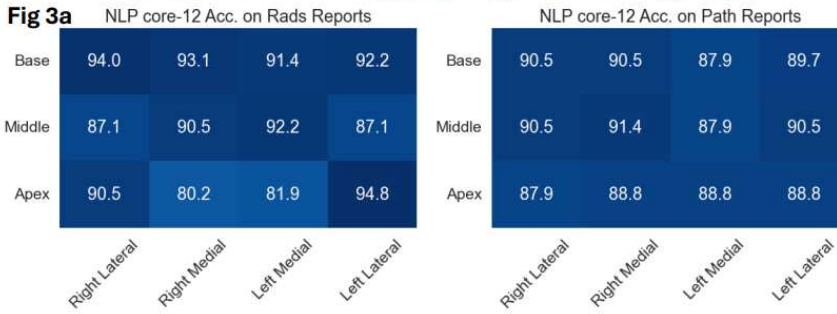


Figure 2a, b: Abbreviated Examples of radiology and pathology report processing. Dotted arrows indicate report snapshots at major processing events. Bolded words in example reports indicate misspellings of words or abbreviations that are rectified by or expanded upon by spell-checker (bold and italicized). Spell-checker (minimization of Levenshtein distance across a custom, medically informed dictionary) uses Trie Nodes to achieve sub-second identification of spelling errors.

NLP performance in Parsing Radiology and Pathology Reports



PI-RADS v2 Sensitivities in Core-12 lesion localization

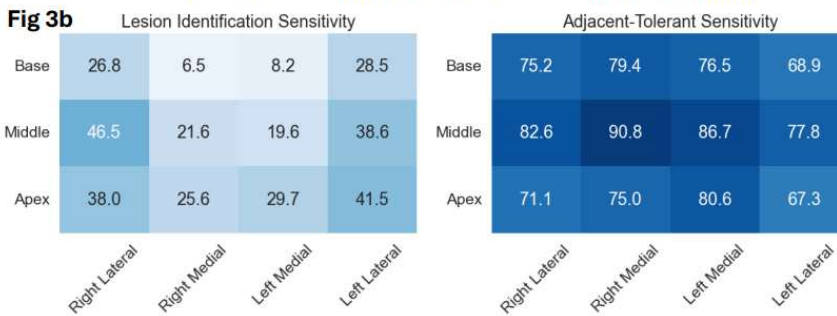


Fig 3a: NLP performance in parsing and localizing radiology and pathology reported lesions, performing with similar accuracy across all 12cBT regions. Fig 3b: Sensitivities of PI-RADS v2 in lesion localization at UCI. While radiological localization correlates poorly with biopsies (left), the predicted lesions generally fall within the broader region defined by adjacent cores (right).

Keywords

Natural Language Processing; Prostate Cancer Detection; Radiology Reports; 12-core Pathology Reports; Automated Correlation; Physician self-improvement