BMJ Open

Is it possible to automatically assess pretreatment digital rectal examination documentation using natural language processing? A single-centre retrospective study

Selen Bozkurt,[1,2] Kathleen M Kan,[3] Michelle K Ferrari,[4] Daniel L Rubin,[1,5] Douglas W Blayney,[6] Tina Hernandez-Boussard,[1,2,7] James D Brooks [4]

For numbered affiliations see end of article.

Correspondence to
Dr James D Brooks;
jdbrooks@stanford.edu

ABSTRACT

Objectives To develop and test a method for automatic assessment of a quality metric, provider-documented pretreatment digital rectal examination (DRE), using the outputs of a natural language processing (NLP) framework.
Setting An electronic health records (EHR)-based prostate cancer data warehouse was used to identify patients and associated clinical notes from 1 January 2005 to 31 December 2017. Using a previously developed natural language processing pipeline, we classified DRE assessment as documented (currently or historically performed), deferred (or suggested as a future examination) and refused.
Primary and secondary outcome measures We investigated the quality metric performance, documentation 6 months before treatment and identified patient and clinical factors associated with metric performance.
Results The cohort included 7215 patients with prostate cancer and 426 227 unique clinical notes associated with pretreatment encounters. DREs of 5958 (82.6%) patients were documented and 1257 (17.4%) of patients did not have a DRE documented in the EHR. A total of 3742 (51.9%) patient DREs were documented within 6 months prior to treatment, meeting the quality metric. Patients with private insurance had a higher rate of DRE 6 months prior to starting treatment as compared with Medicaid-based or Medicare-based payors (77.3% vs 69.5%, p=0.001). Patients undergoing chemotherapy, radiation therapy or surgery as the first line of treatment were more likely to have a documented DRE 6 months prior to treatment.
Conclusion EHRs contain valuable unstructured information and with NLP, it is feasible to accurately and efficiently identify quality metrics with current documentation clinician workflow.

INTRODUCTION

Prostate cancer is the most common malignant tumour among men and over 85% of prostate cancers are localised at diagnosis.[1] Digital rectal examination (DRE) is part of a physical exam routinely used for prostate cancer screening and pretreatment assessment. Studies have shown that DRE is an important tool for early diagnosis of prostate cancer, especially when serum levels of prostate-specific antigen are low.[2–4] DRE is used to assess the need for a needle biopsy and for clinical tumour staging and for treatment decisions once the diagnosis of prostate cancer is made.[5–7] DRE findings are useful in surgical planning, such as the need to remove a neurovascular bundle, and can suggest additional diagnostic imaging in patients with locally advanced disease who are at greater risk for metastatic disease and can help guide imaging for radiation treatment planning.

Performance of a digital rectal exam has been identified as an important quality metric in prostate cancer care,[8 9] and most clinical guidelines include DRE as part of a comprehensive pretreatment assessment.[6 7] However, DRE results are often not systematically recorded nor included in the billing or claims datasets used in outcomes research. Several studies have conducted manual chart reviews of patients with prostate cancer to assess DRE documentation.[2 10 11] These reviews found

that while DRE was well documented, clinical details from the exam were often missing.

As we move into the era of digital healthcare, adoption and implementation of electronic health records (EHRs) have become widespread in clinical practice.[12] Some studies have used EHRs to evaluate DRE as a screening tool and showed that, while DRE is important for early detection of prostate cancer,[3] it has modest rates of documentation.[13] Furthermore, the use of EHRs to assess DRE from the quality metric perspective, including documentation, quality of the information documented and appropriate timing for a pretreatment assessment is limited, as the DRE-related information is available only in free-text clinical narratives which thwarts computerised applications. New studies demonstrate the ability to leverage the information available in EHRs in order to improve the clinicians' processes of recording care in a digital format, and the way they conduct clinical research, moving from labour intensive manual chart reviews to automated and efficient text processing.[14] Therefore, advanced computational techniques, such as natural language processing (NLP), are needed to extract DRE information from EHRs for use in intelligent applications.

NLP has emerged as a potential solution for bridging the gap between free-text and structured representations of clinical information. Since biomedical domains exhibit high degree of terminological variation, NLP's precision becomes even more valuable based on its ability to automatically recognise all variants of domain language.[15] There are a few studies[16–20] using NLP techniques to extract concepts related to prostate cancer from EHRs for decision support, but the accuracy of DRE extraction was reported only in two studies.[16 21] The features related to DRE reporting was also missing in those studies.

To enhance availability of quantifiable measures of quality care delivery, we have developed a system to use routinely collected electronic clinical text data to measure and improve prostate cancer care. The primary aim of this study was automatic assessment of pretreatment DRE documentation using the outputs of an NLP framework. Our previous work demonstrates that an NLP pipeline can effectively capture the documentation, timing and exam results from unstructured text.[22] In this paper, we used this NLP pipeline to investigate DRE documentation and an academic institution's performance on this endorsed quality metric. Our approach opens new opportunities for scalable and automated information extraction of quality measures from unstructured clinical data and highlights areas for focused quality improvement efforts.

METHODS
Data source
Data were collected from a prostate cancer data warehouse, which is described in detail elsewhere.[23] In brief, data were collected from a tertiary care academic healthcare system Epic EHR system (Epic Systems, Verona, Wisconsin, USA) and linked to the California Cancer
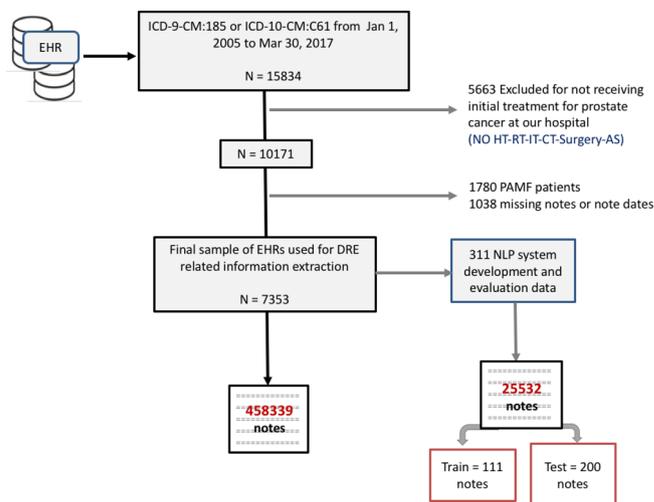

Figure 1 Flow chart of the selection process used to identify patients with prostate cancer for DRE information extraction from notes. DRE, digital rectal examination; EHR, electronic health records, ICD, International Classification of Diseases; NLP, natural language processing.

Registry.[24] The registry contains structured data on diagnosis, histology, cancer stage, treatment and outcomes, including information from healthcare organisations across California. Data were included from 1 January 2005 to 31 December 2017. Our study design is illustrated in online supplementary figure 1.

Study cohort
Patients with prostate cancer >34 years of age were identified using International Classification of Diseases (ICD) diagnostic codes, ICD-9-CM:185 and ICD-10-CM:C61. Patients were excluded if they did not receive initial treatment for prostate cancer (prostatectomy, radiotherapy, hormone therapy or chemotherapy) in our health system, as we would not have their detailed initial pretreatment medical records. We further excluded patients who received cystoprostatectomy (ICD-9 procedure code: 57.71), as these patients had a primary diagnosis of bladder cancer with prostate cancer detected incidentally on surgical pathology (figure 1).

Patient and public involvement
No patients were involved in setting the research question, study design or outcome measures. Patients will be informed by press reports and the institute's homepage following this publication.

The NLP pipeline
We decomposed the NLP problem into a set of subtasks (ie, preprocessing, dictionary mapping, etc), and created a processing pipeline comprising a set of sequentially executed processing modules to automatically recognise the pertinent named entities in clinical narratives. The preprocessing steps included sentence boundary detection, tokenisation and removing all the stop words, punctuation characters and two-letter words. The integer and floating-point numbers were converted to the

Example text snippets from clinical notes	Output of the NLP pipeline
On <u>digital rectal examination</u>, the prostate bed is scaphoid.	DRE: performed and documented
Prostate nodule on <u>DRE</u>.	DRE: performed and documented
<u>DRE</u> : <u>deferred</u>	DRE: not performed but documented
HISTORY OF PRESENT ILLNESS: The patient **was found** to have a palpable right-sided firmness on <u>digital rectal examination</u>.	DRE: not performed but documented as historical

Figure 2 Examples of DRE documentation in clinical notes and output of the NLP pipeline. DRE, digital rectal examination; NLP, natural language processing.

corresponding string representation. In order to show how DREs are documented in clinical notes and the appearance of outputs from our NLP pipeline, unselected examples of extracted text are displayed in figure 2.

In order to define name entities, a domain-specific dictionary was created based on domain knowledge experts (KMK, MF, JDB), matches with existing ontologies and using distributional information on words and phrases from all notes in the database. We used a distributional semantics method which can be easily trained from a large corpus and derive representations for words in such a way that words occurring in similar contexts will have similar representations.[25–27] To carry out the distributional semantics method, we used the word2vec package created by Mikolov *et al* to build vector representations of all terms in our corpus.[28] The skip-gram model with vector length of 100 and a context window width of 5 was used to recognise biomedical synonyms from among the options. Because word2vec constructs a separate vector for each token in the text, there is no intuitive way to compose the vectors for 'rectal' and 'exam' into a combined vector for the term 'rectal exam'. We thus needed to create single entities for multiword phrases before running word2vec using bigrams and trigrams of words.[25 29] In order to quantify closeness of the word vectors, we used the cosine similarity between two terms. Based on the calculated word similarity metric, additional synonyms and other lexical variants were identified in the clinical notes and terms list was expanded after the expert confirmation of candidate terms. We also implemented the ConText algorithm (with small modifications)[22] within our NLP system in order to

identify three contextual values: negation, hypothetical and historical.[30]

We evaluated the pipeline on a set of patient reports that had been manually annotated by two domain experts (MF and JDB). As described elsewhere, our pipeline had 95% precision and 90% recall.[22] Using the NLP pipeline, we classified patient-specific DREs as documented (current or performed historically), deferred (or suggested as a future examination) or refused. In addition, we investigated the timing of DRE documentation in accordance with the quality metric: 6 months before treatment.

Statistical analysis

Statistical analysis of clinical characteristics between two groups consisted of unpaired t-tests and Mann-Whitney U test, whereas the X^2/Fisher's exact tests were used for categorical variables. All statistical tests were two-sided with a threshold of p≤0.05 for statistical significance and carried out using R project statistical environment (http://www.r-project.org).

RESULTS

The final cohort consisted of 7215 patients and 426 227 unique clinical notes. A total of 5958 (82.6%) patients had at least one DRE documented in their medical record at any time during their course of care. In addition, 737 (10.2%) patients' reports state directly that the patient deferred or refused DRE before treatment. Among deferred/refused cases (737), 357 (48.4%) had a DRE documented after treatment. A total of 3742 (51.9%) patients had a DRE documented 6 months prior to initial treatment (table 1).

Characteristics of the patients based on the DRE documentation (performed or not performed/mentioned) are provided in table 2. While age, first line of treatment and stage were significantly different (p<0.001 for each characteristics), the insurance payor type, ethnicity and race were not different across the groups. Patients initially managed with active surveillance, surgery and chemotherapy had a higher rate of DRE prior to starting

Table 1 DRE documentation statistics before initial treatment

Stage*	Documented (performed)		Documented (not performed)	Not documented	Total
	6+ months before Tx	≤6 months before Tx	Deferred/Refused	No mention	
1, n (%)	199 (28.5)	332 (47.5)	79 (11.3)	89 (12.7)	699
2, n (%)	720 (18.3)	2102 (53.4)	374 (9.5)	743 (18.9)	3939
3, n (%)	110 (11.8)	607 (65)	65 (7)	152 (16.3)	934
4, n (%)	137 (21.4)	286 (44.8)	96 (15)	120 (18.8)	639
N/A†, n (%)	313 (31.2)	415 (41.3)	123 (12.3)	153 (15.2)	1004
Total	1479 (20.5)	3742 (51.9)	737 (10.2)	1257 (17.4)	7215

*Stage=overall stage.
†Stage information is missing in the database.
DRE, digital rectal examination; Tx, treatment.

Table 2 Characteristics for patients based on digital rectal examination documentation as performed or not performed

Patient characteristics	Not performed or mentioned n=1994 (27.6%)	Performed n=5221 (72.4%)	P value
Age at diagnosis (years), $\bar{x} \pm SD$	67.46±9.52	66.30± 9.03	<0.001
Insurance payor type, n (%)			
Private	539 (25.5)	1576 (74.5)	0.236
Medicare/Medical	1199 (26.9)	3264 (73.1)	
Ethnicity, n (%)			
Non-Hispanic	1815 (27.5)	4784 (72.5)	0.432
Hispanic	159 (29.1)	388 (70.9)	
Race, n (%)			
Asian	224 (28.1)	574 (71.9)	0.317
White	1480 (27.3)	3933 (72.7)	
Other	136 (30.6)	308 (69.4)	
First line of treatment, n (%)			
Active surveillance	154 (21.8)	554 (78.2)	<0.001
Chemotherapy	70 (24.9)	211 (75.1)	
Hormone therapy	577 (32.1)	1223 (67.9)	
Radiation therapy	431 (34.1)	832 (65.9)	
Surgery	762 (24.1)	2401 (75.9)	
Stage, n (%)			
1	168 (24.0)	531 (76.0)	<0.001
2	1117 (28.4)	2822 (71.6)	
3	217 (23.2)	717 (76.8)	
4	216 (33.8)	423 (66.2)	

treatment as compared with hormone and radiation therapy (78.2%, 75.9%, 75.1% vs 67.9% and 65.9%, p<0.001). Likewise, stage 1 and 3 patients had a higher rate of DRE prior to starting treatment as compared with stage 2 and 4 patients (76%, 76.8% vs 71.6%, 66.2%, p<0.001).

Characteristics for patients receiving a DRE, stratified by quality metric adherence (ie, 6 months prior to treatment) are summarised in table 3. Patients with private insurance had a higher rate of DRE prior to starting treatment compared with patients with Medicaid or Medicare (77.3 vs 69.5, p<0.001). Patients undergoing chemotherapy, radiation therapy or surgery as the first line of treatment were more likely to have a documented DRE 6 months prior to treatment. Likewise, stage 2 and 3 patients had a higher rate of DRE prior to starting therapy as compared with stage 1 and 4 patients.

Documentation of DREs improved slightly over time from an initial rate of 67% in 2005 to 75% in 2009, and 87% in 2016/2017 (online supplementary figure 2).

DISCUSSION

Systematic monitoring and reporting of DRE documentation has been limited in the past due to its non-standardised capture and storage as unstructured, free text in EHRs. We developed a computational approach to identify and extract the reporting of DRE in virtually all patients seen and treated for prostate cancer at a tertiary care academic healthcare system within a prespecified time period. We found that the majority of patients had a documented DRE in their clinical records, and that the DRE was performed prior to treatment, satisfying the endorsed quality metric. To our knowledge, this is the largest study on DRE documentation using real-world unstructured text data and the proposed framework opens opportunities for improved quality metric monitoring and reporting within the current clinician documentation workflow and without increased documentation burden.

Our evaluation of over 7215 patients undergoing treatment for prostate cancer revealed that over 82.6% of patients had a documented DRE in the EHR and 72.4% of these were prior to initiating treatment. DRE compliance among patients in the screening setting has been investigated using population-based surveys, which are compromised by recall bias, respondent bias or reporting errors.[31 32] Studies on provider compliance with performance of DRE for screening purposes consist of much smaller groups of patients and primarily use manual chart reviews.[13 33] One of the larger studies reviewed compliance

Table 3 Characteristics for patients receiving a digital rectal examination, stratified by quality metric timing

Patient characteristics	≤6 months before Tx n=3742	6+ months before Tx n=1479	P value
Age at diagnosis (years), $\bar{x}\pm$SD	65.81±8.83	67.54±9.41	
Insurance payor type, n (%)			
Private	1219 (77.3)	357 (22.7)	<0.001
Medicare/Medical	2269 (69.5)	995 (30.5)	
Ethnicity, n (%)			
Non-Hispanic	3431 (71.7)	1353 (28.3)	0.767
Hispanic	281 (72.4)	107 (27.6)	
Race, n (%)			
Asian	423 (73.7)	151 (26.3)	0.523
White	2809 (71.4)	1124 (28.6)	
Other*	222 (72.1)	86 (27.9)	
First line of treatment, n (%)			
Active surveillance	248 (44.8)	306 (55.2)	<0.001
Chemotherapy	137 (64.9)	74 (35.1)	
Hormone therapy	768 (62.8)	455 (37.2)	
Radiation therapy	612 (73.6)	220 (26.4)	
Surgery	1977 (82.3)	424 (17.7)	
Stage, n (%)			
1	332 (62.5)	199 (37.5)	<0.001
2	2102 (74.5)	720 (25.5)	
3	607 (84.7)	110 (15.3)	
4	286 (67.6)	137 (32.4)	

Tx, treatment.

in the American College of Surgeons National Cancer Database using an extraction tool found an 80% rate of compliance for this measure.[11]

Patients with private insurance had a higher rate of DRE documentation 6 months prior to starting treatment as compared with those with Medicaid or Medicare. We did not find differences among our patient population in terms of ethnicity and race. Patients undergoing surgery and radiation as the first line of treatment were the most likely to have a documented DRE 6 months prior to treatment. DRE assessment of the prostate likely was highest in these settings since the information gained has direct implications for treatment planning, such as guiding nerve-sparing techniques in surgery and determining prostate volume in radiation. It is not clear why stage 3 patients had the highest rates of DRE recorded. By definition, most patients with stage 3 cancer will have palpable abnormalities on DRE, and the higher rates might reflect a bias in physicians recording positive findings on physical examination or a higher likelihood of performing an exam in a patient referred to their practice with a previously documented abnormality. Patients undergoing chemotherapy or hormone therapy nearly always have metastatic disease where local stage is far less

likely to influence clinical decision-making. Interestingly, patients on active surveillance had the highest overall rate of documented DRE but failed to meet the quality metric of having their DRE within 6 months of diagnosis. This failure is likely because we defined the start date of active surveillance based on their date of diagnosis. Most patients placed on active surveillance had a DRE prior to or at the time of their diagnostic biopsy and would therefore meet the quality metric. The high rates of DRE after diagnosis in these patients is because DRE is part of routine follow-up in nearly all active surveillance protocols. As active surveillance becomes more widely adapted in practice, our algorithms will need to be modified to capture prediagnostic examinations.

We found an overall improvement in compliance at our institution over time, with documented DREs increasing from an initial rate of 67% in 2005 to 87% in 2017. We suspect this improvement is due to widespread education on published quality metrics and institutional encouragement of the providers treating patients with prostate cancer to improve their compliance with published guidelines. In the future, information captured by mining EHRs could be used to provide individual providers feedback on their adherence to guidelines. Eventually, these

tools could be incorporated into the EHR to promote guideline adherence and documentation in real time.

Several guidelines recommend performance of a DRE prior to initiation of therapy, but are mute on whether and how this quality metric needs to be documented. The rapid rise of the EHR offers new opportunities for measuring compliance with diverse quality metrics that was previously impossible with paper records. We have focused on using the EHR to address the need to measure compliance with several quality metrics in prostate cancer care, including DRE, and have made these algorithms freely available so that other healthcare systems can modify and use them.[17 34] In the near future, measurement of compliance with quality metrics, with the goal of improving patient outcomes, will be used to monitor and quantify healthcare quality. Compliance with quality metrics could be directly tethered to reimbursements, in which case, building structured fields for documentation of quality metrics will increase since these will improve compliance with, and documentation of quality metrics like performance of a DRE prior to initiating therapy.

On the other hand, many of the >100 quality metrics proposed in prostate cancer are based on expert opinion and other low-level evidence.[8] Given the widespread concerns about 'click fatigue' and physician burn-out related to documentation in the EHR (PMID: 29801050), required structured fields need to be deployed wisely and judiciously. As we move to a learning-based healthcare system with the EHR at its hub, capture of compliance and non-compliance with quality metrics with the automated methods such as ours will provide opportunities to assess the true effects of compliance with disease-related outcomes (such as higher rates of positive surgical margins and recurrence in patients undergoing surgery who did not receive a DRE) as well patient-centred outcomes. Ideally, quality metrics that significantly effect clinical end points and patient-centred outcomes can be tested and incorporated as structured fields, while those that do not can be discarded to decrease documentation burdens on physicians. In addition, understanding which quality metrics lead to improved outcomes and patient experience would allow insurers to tailor reimbursements to incentivise improvements in quality.

Our study has some limitations. Our algorithm was developed at a single academic site, where unique vocabularies might be used in recording DREs. Since our data are from a single institution, the rates of DRE performance might not generalise to other populations. Validation in an external cohort could address these issues to a great extent, although external validation poses several practical challenges. Sites often are reluctant to provide access to their EHR because of patient privacy concerns, and to avoid comparisons of adherence to quality metrics with other institutions. Sites also lack the information technology infrastructure and linkage to a data warehouse that makes this work possible.

CONCLUSIONS

In this study of documentation of DRE within 6 months prior to prostate cancer treatment using NLP, we found EHRs contain valuable free unstructured text information on this clinical assessment for the majority of patients with prostate cancer. With NLP techniques, it is feasible to accurately and efficiently identify adherence to this quality metric without increasing documentation burden to clinicians. By leveraging big data for quality assessment, this methodology allows us to systematically paint a picture of our institution's clinical practice by mining free-text clinical narrative text embedded in the EHR. Other settings can benefit from such algorithms to improve quality assessment while not further burdening clinicians. As healthcare moves forward with public reporting of quality metrics and federal incentives are attaching value to these metrics, it is important to produce precise, valid measures that reflect best medical practices that are considered important by clinicians. By doing so, timely and efficient feedback can be provided to continually improve patient care.

Author affiliations
[1]Biomedical Data Science, Stanford University, Stanford, CA, USA
[2]Medicine (Biomedical Informatics), Stanford University, Stanford, CA, USA
[3]Urology, Stanford Lucile Salter Packard Children's Hospital, Stanford, CA, USA
[4]Urology, Stanford University, Stanford, CA, USA
[5]Radiology, Stanford University, Stanford, CA, USA
[6]Medicine (Oncology), Stanford University, Stanford, CA, USA
[7]Surgery, Stanford University, Stanford, CA, USA

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. *CA Cancer J Clin* 2015;65:5–29.
2. Palmerola R, Smith P, Elliot V, *et al*. The digital rectal examination (DRE) remains important - outcomes from a contemporary cohort of men undergoing an initial 12-18 core prostate needle biopsy. *Can J Urol* 2012;19:6542–7.
3. Walsh AL, Considine SW, Thomas AZ, *et al*. Digital rectal examination in primary care is important for early detection of prostate cancer: a retrospective cohort analysis study. *Br J Gen Pract* 2014;64:e783–e787.
4. Kowalik CG, Gerling GJ, Lee AJ, *et al*. Construct validity in a high-fidelity prostate exam simulator. *Prostate Cancer Prostatic Dis* 2012;15:63–9.
5. Heidenreich A, Bastian PJ, Bellmunt J, *et al*. EAU guidelines on prostate cancer. part 1: screening, diagnosis, and local treatment with curative intent-update 2013. *Eur Urol* 2014;65:124–37.
6. Thompson I, Thrasher JB, Aus G, *et al*. Guideline for the management of clinically localized prostate cancer: 2007 update. *J Urol* 2007;177:2106–31.
7. Mohler JL, Armstrong AJ, Bahnson RR, *et al*. Prostate Cancer, Version 1.2016. *J Natl Compr Canc Netw* 2016;14:19–30.
8. Gori D, Dulal R, Blayney DW, *et al*. Utilization of Prostate Cancer Quality Metrics for Research and Quality Improvement: A Structured Review. *Jt Comm J Qual Patient Saf* 2019;45.
9. Litwin MS, Steinberg M, Malin J, *et al*. Prostate cancer patient outcomes and choice of providers: development of an infrastructure for quality assessment. *Rand Corp Santa Monica Ca* 2000.
10. Sayre EC, Bunting PS, Kopec JA. Reliability of self-report versus chart-based prostate cancer, PSA, DRE and urinary symptoms. *Can J Urol* 2009;16:4463–71.
11. Spencer BA, Miller DC, Litwin MS, *et al*. Variations in quality of care for men with early-stage prostate cancer. *J Clin Oncol* 2008;26:3735–42.
12. Adler-Milstein J, DesRoches CM, Kralovec P, *et al*. Electronic Health Record Adoption In US Hospitals: Progress Continues, But Challenges Persist. *Health Aff* 2015;34:2174–80.
13. Federman DG, Pitkin P, Carbone V, *et al*. Screening for prostate cancer: are digital rectal examinations being performed? *Hosp Pract* 2014;42:103–7.
14. Pan HY, Shaitelman SF, Perkins GH, *et al*. Implementing a Real-Time Electronic Data Capture System to Improve Clinical Documentation in Radiation Oncology. *J Am Coll Radiol* 2016;13:401–7.
15. Spasić I, Livsey J, Keane JA, *et al*. Text mining of cancer-related information: review of current status and future directions. *Int J Med Inform* 2014;83:605–23.
16. Hernandez-Boussard T, Kourdis P, Dulal R, *et al*. A natural language processing algorithm to measure quality prostate cancer care. *American Society of Clinical Oncology* 2017.
17. Hernandez-Boussard T, Kourdis PD, Seto T, *et al*. Mining Electronic Health Records to Extract Patient-Centered Outcomes Following Prostate Cancer Treatment. *AMIA Annu Symp Proc* 2017;2017:876–82.
18. Thomas A, Zheng C, Jung H, *et al*. 83 Validity of natural language processing to identify patients with prostate cancer. *J Urol* 2013;189:e34.
19. Thomas AA, Zheng C, Jung H, *et al*. Extracting data from electronic medical records: validation of a natural language processing program to assess prostate biopsy results. *World J Urol* 2014;32:99–103.
20. Gregg JR, Lang M, Wang LL, *et al*. Automating the Determination of Prostate Cancer Risk Strata From Electronic Medical Records. *JCO Clin Cancer Inform* 2017;1:1–8.
21. Hong SN, Son HJ, Choi SK, *et al*. A prediction model for advanced colorectal neoplasia in an asymptomatic screening population. *PLoS One* 2017;12:e0181040.
22. Selen Bozkurt JIP, Kathleen Mary Kan MF, Daniel L Rubin, James D Brooks, Tina Hernandez-Boussard, *et al*. Ed. An Automated Feature Engineering for Digital Rectal: In. *An Automated Feature Engineering for Digital Rectal Examination Documentation using Natural Language Processing*. San Francisco, CA, 2018.
23. Seneviratne MG, Seto T, Blayney DW, *et al*. Architecture and Implementation of a Clinical Research Data Warehouse for Prostate Cancer. *eGEMs* 2018;6. eGEMs (Generating Evidence & Methods to improve patient outcomes).
24. (CDSRB) CDoPHsCDSaRB. California Cancer Registry (CCR). [http://www.ccrcal.org/ (cited 15 Aug 2018).
25. Percha B, Zhang Y, Bozkurt S, *et al*. Expanding a radiology lexicon using contextual patterns in radiology reports. *J Am Med Inform Assoc* 2018;25:679–85.
26. Gupta A, Banerjee I, Rubin DL. Automatic information extraction from unstructured mammography reports using distributed semantics. *J Biomed Inform* 2018;78:78–86.
27. McGregor S, Agres K, Purver M, *et al*. From distributional semantics to conceptual spaces: A novel computational method for concept creation. *Journal of Artificial General Intelligence* 2015;6:55–86.
28. Mikolov T, Sutskever I, Chen K, Corrado GS, Dean J, ed. *Distributed representations of words and phrases and their compositionality*. Advances in neural information processing systems. 2013.
29. Banerjee I, Chen MC, Lungren MP, *et al*. Radiology report annotation using intelligent word embeddings: Applied to multi-institutional chest CT cohort. *J Biomed Inform* 2018;77:11–20.
30. Chapman WW, Chu D, Dowling JN, eds. *ConText: An algorithm for identifying contextual features from clinical text. Proceedings of the workshop on BioNLP 2007: biological, translational, and clinical language processing*: Association for Computational Linguistic, 2007.
31. Ross LE, Taylor YJ, Richardson LC, *et al*. Patterns in prostate-specific antigen test use and digital rectal examinations in the Behavioral Risk Factor Surveillance System, 2002-2006. *J Natl Med Assoc* 2009;101:316–24.
32. Tourville EA, Nguyen MM. Prostate cancer detection by using digital rectal examination: contemporary practice patterns in the United States. *Clin Genitourin Cancer* 2013;11:263–9.
33. Murthy GD, Byron DP, Pasquale D. Underutilization of digital rectal examination when screening for prostate cancer. *Arch Intern Med* 2004;164:313–6.
34. Banerjee I, Li K, Seneviratne M, *et al*. Weakly supervised natural language processing for assessing patient-centered outcome following prostate cancer treatment. *JAMIA Open* 2019;2:150–9.