

5-12-2016

New Paradigms for Patient-Centered Outcomes Research in Electronic Medical Records: An example of detecting urinary incontinence following prostatectomy

Tina Hernandez-Boussard
Stanford University, boussard@stanford.edu

Suzanne Tamang
Stanford University, stamang@stanford.edu

Douglas Blayney
Stanford University, dblayney@stanford.edu

Jim Brooks
Stanford University, jdbrooks@stanford.edu

See next pages for additional authors

Follow this and additional works at: <http://repository.edm-forum.org/egems>

 Part of the [Health Information Technology Commons](#), [Health Services Research Commons](#), [Oncology Commons](#), [Surgery Commons](#), and the [Urology Commons](#)

Recommended Citation

Hernandez-Boussard, Tina; Tamang, Suzanne; Blayney, Douglas; Brooks, Jim; and Shah, Nigam (2016) "New Paradigms for Patient-Centered Outcomes Research in Electronic Medical Records: An example of detecting urinary incontinence following prostatectomy," *eGEMs (Generating Evidence & Methods to improve patient outcomes)*: Vol. 4: Iss. 3, Article 1.

DOI: <http://dx.doi.org/10.13063/2327-9214.1231>

Available at: <http://repository.edm-forum.org/egems/vol4/iss3/1>

This Informatics Empirical Research is brought to you for free and open access by the the Publish at EDM Forum Community. It has been peer-reviewed and accepted for publication in eGEMs (Generating Evidence & Methods to improve patient outcomes).

The Electronic Data Methods (EDM) Forum is supported by the Agency for Healthcare Research and Quality (AHRQ), Grant 1U18HS022789-01. eGEMs publications do not reflect the official views of AHRQ or the United States Department of Health and Human Services.

New Paradigms for Patient-Centered Outcomes Research in Electronic Medical Records: An example of detecting urinary incontinence following prostatectomy

Abstract

Introduction: National initiatives to develop quality metrics emphasize the need to include patient-centered outcomes. Patient-centered outcomes are complex, require documentation of patient communications, and have not been routinely collected by healthcare providers. The widespread implementation of electronic medical records (EHR) offers opportunities to assess patient-centered outcomes within the routine healthcare delivery system. The objective of this study was to test the feasibility and accuracy of identifying patient centered outcomes within the EHR.

Methods: Data from patients with localized prostate cancer undergoing prostatectomy were used to develop and test algorithms to accurately identify patient-centered outcomes in post-operative EHRs – we used *urinary incontinence* as the use case. Standard data mining techniques were used to extract and annotate free text and structured data to assess urinary incontinence recorded within the EHRs.

Results. A total 5,349 prostate cancer patients were identified in our EHR-system between 1998-2013. Among these EHRs, 30.3% had a text mention of *urinary incontinence* within 90 days post-operative compared to less than 1.0% with a structured data field for *urinary incontinence* (i.e. ICD-9 code). Our workflow had good precision and recall for urinary incontinence (positive predictive value: 0.73 and sensitivity: 0.84).

Discussion. Our data indicate that important patient-centered outcomes, such as urinary incontinence, are being captured in EHRs as free text and highlight the long-standing importance of accurate clinician documentation. Standard data mining algorithms can accurately and efficiently identify these outcomes in existing EHRs; the complete assessment of these outcomes is essential to move practice into the patient-centered realm of healthcare.

Acknowledgements

Research reported in this presentation was supported by the National Cancer Institute of the National Institutes of Health under Award Number R01CA183962. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Keywords

Electronic health records; quality improvement; patient-centered care; data mining; health services research

Disciplines

Health Information Technology | Health Services Research | Oncology | Surgery | Urology

Creative Commons License



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License](https://creativecommons.org/licenses/by-nc-nd/3.0/).

Authors

Tina Hernandez-Boussard, *Stanford University*; Suzanne Tamang, *Stanford University*; Douglas Blayney, *Stanford University*; Jim Brooks, *Stanford University*; Nigam Shah, *Stanford University*.



New Paradigms for Patient-Centered Outcomes Research in Electronic Medical Records: An Example of Detecting Urinary Incontinence Following Prostatectomy

Tina Hernandez-Boussard, PhD, MPH; Suzanne Tamang, MS; Douglas Blayney, MD; Jim Brooks, MD; Nigam Shah, MBBS, PhD¹

ABSTRACT

Introduction: National initiatives to develop quality metrics emphasize the need to include patient-centered outcomes. Patient-centered outcomes are complex, require documentation of patient communications, and have not been routinely collected by healthcare providers. The widespread implementation of electronic medical records (EHR) offers opportunities to assess patient-centered outcomes within the routine healthcare delivery system. The objective of this study was to test the feasibility and accuracy of identifying patient centered outcomes within the EHR.

Methods: Data from patients with localized prostate cancer undergoing prostatectomy were used to develop and test algorithms to accurately identify patient-centered outcomes in post-operative EHRs—we used *urinary incontinence* as the use case. Standard data mining techniques were used to extract and annotate free text and structured data to assess urinary incontinence recorded within the EHRs.

Results: A total 5,349 prostate cancer patients were identified in our EHR-system between 1998-2013. Among these EHRs, 30.3% had a text mention of *urinary incontinence* within 90 days post-operative compared to less than 1.0% with a structured data field for *urinary incontinence* (i.e. ICD-9 code). Our workflow had good precision and recall for urinary incontinence (positive predictive value: 0.73 and sensitivity: 0.84).

Discussion: Our data indicate that important patient-centered outcomes, such as urinary incontinence, are being captured in EHRs as free text and highlight the long-standing importance of accurate clinician documentation. Standard data mining algorithms can accurately and efficiently identify these outcomes in existing EHRs; the complete assessment of these outcomes is essential to move practice into the patient-centered realm of healthcare.

¹Stanford University

Introduction

Prostate cancer is the most common malignancy in men.¹ Although survival rates for prostate cancer treatment are excellent, patients may acquire treatment-related side effects, many which can only be reported by the patient (e.g., urinary incontinence, erectile dysfunction, or bowel dysfunction).²⁻⁴ Reported rates of such side effects vary depending on the population studied or treatment characteristics.⁵⁻¹⁰ The majority of research on these patient-centered outcomes stems from high-volume academic centers, which may undermine its generalizability to other settings.^{11,12} Given the current state of prostate cancer research, both patients and clinicians have limited evidence to guide their treatment choices;^{13,14} accurate and efficient measurement of outcomes other than mortality are needed to help patients make informed decisions regarding their treatment pathway.

The Patient Protection and Affordable Care Act (ACA) aims to improve the quality and efficacy of health care delivery in the United States.¹⁵ Many sections of the ACA rely on accurate quality measurement (e.g., value-based payment modifiers)¹⁶ and efficient data retrieval (e.g., accountable care organizations and their information exchange).¹⁷ Furthermore, many sections of the ACA include patient-centered initiatives, which promote the use of patient-centered outcomes in clinical decision-making.¹⁸ Under the health care reform, accurate quality measurement is essential and should be patient centered.

Patient-centered care reflects a patient's overall health care experience and assesses the net effects of disease and treatment (e.g., disease-related quality of life, urinary incontinence, and overall health status) rather than physiological endpoints (e.g., laboratory values and disease-specific survival).¹⁹ Patient-centered endpoints are complex, require

documentation of patient communications, and have not been routinely collected by health care providers. Patient-centered outcomes are not routinely captured as structured or coded data and therefore do not exist in administrative billing or claims data.²⁰ Therefore, current patient-centered outcome reports must rely on size-limiting patient surveys (which contain ascertainment bias), prospective studies (which are not readily available and also contain ascertainment bias), or manual chart reviews (which are time limiting).

The purpose of this study was to test the feasibility of using data mining algorithms to identify patient-centered outcomes in routinely collected electronic health records (EHRs); we use postprostatectomy urinary incontinence as a use case. Our methods apply techniques from the fields of data mining and information extraction. They are distinguished from previous studies that combine structured and unstructured EHR data by their focus on patient-centered outcome detection.

Methods

To identify patient-centered outcomes in EHRs, a robust workflow for deriving these data from routinely used EHRs is essential. Many diseases, such as prostate cancer, have important patient-centered outcomes that are not reliably recorded as coded data. We have to extract these data from the free text existing in the EHRs (e.g., clinicians' reports, narrative text) using data mining algorithms, such as Natural Language Processing (NLP). NLP techniques automatically identify structured text or "knowledge" from free text using controlled vocabularies (e.g., ontologies or user-developed dictionaries) and grammatical rules.^{21,22} Often patterns and labels within the narrative text are queried to identify common phrases, such as regular expressions for high blood pressure.²³ These data mining algorithms are becoming common to identify diseases and



cohorts of patients as health care moves to the digital age.^{24,25} We have developed such a system to identify clinicians' reporting of postoperative urinary incontinence in patients who were diagnosed with localized prostate cancer and who underwent prostatectomy. The information on the patient-centered outcome were derived from the coded data (e.g., ICD 9 codes) as well as the narrative text portions of EHRs—including clinical progress notes, referral notes, procedure reports, and postoperative reports from patients receiving care at the academic center.

Data Set and Study Population

We obtained data from a large, tertiary academic medical center that provides inpatient-, outpatient-, and primary care. During the time of our analysis, the center used the Epic (Epic Systems, Verona Wisconsin) EHR system. The access to de identified EHR data was obtained through an innovative research data warehouse that facilitates research.²⁶ This translational research platform allows the capture of both structured data (e.g., ICD-9-CM codes, laboratory values, etc.) as well as unstructured data (e.g., clinicians' narrative text, preoperative notes, etc.) on all patients receiving care at the institute.

We identified patients in our research platform with localized prostate cancer based on ICD-9-CM code 185. (Figure 1). Patients were categorized into prostatectomy surgical groups according to ICD 9 procedure codes: open prostatectomy, ICD 9 60.5 and CPT 55845; robotic prostatectomy, ICD 9 60.5 plus 17.42 and CPT 55866; laparoscopic prostatectomy, 60.5 plus 54.21; and other prostatectomies, which included CPT codes that were not distinguishable between robotic and laparoscopic procedures, e.g., CPT 55840. In our data mining analysis, we exclude patients without a clinical note and without a follow-up visit within

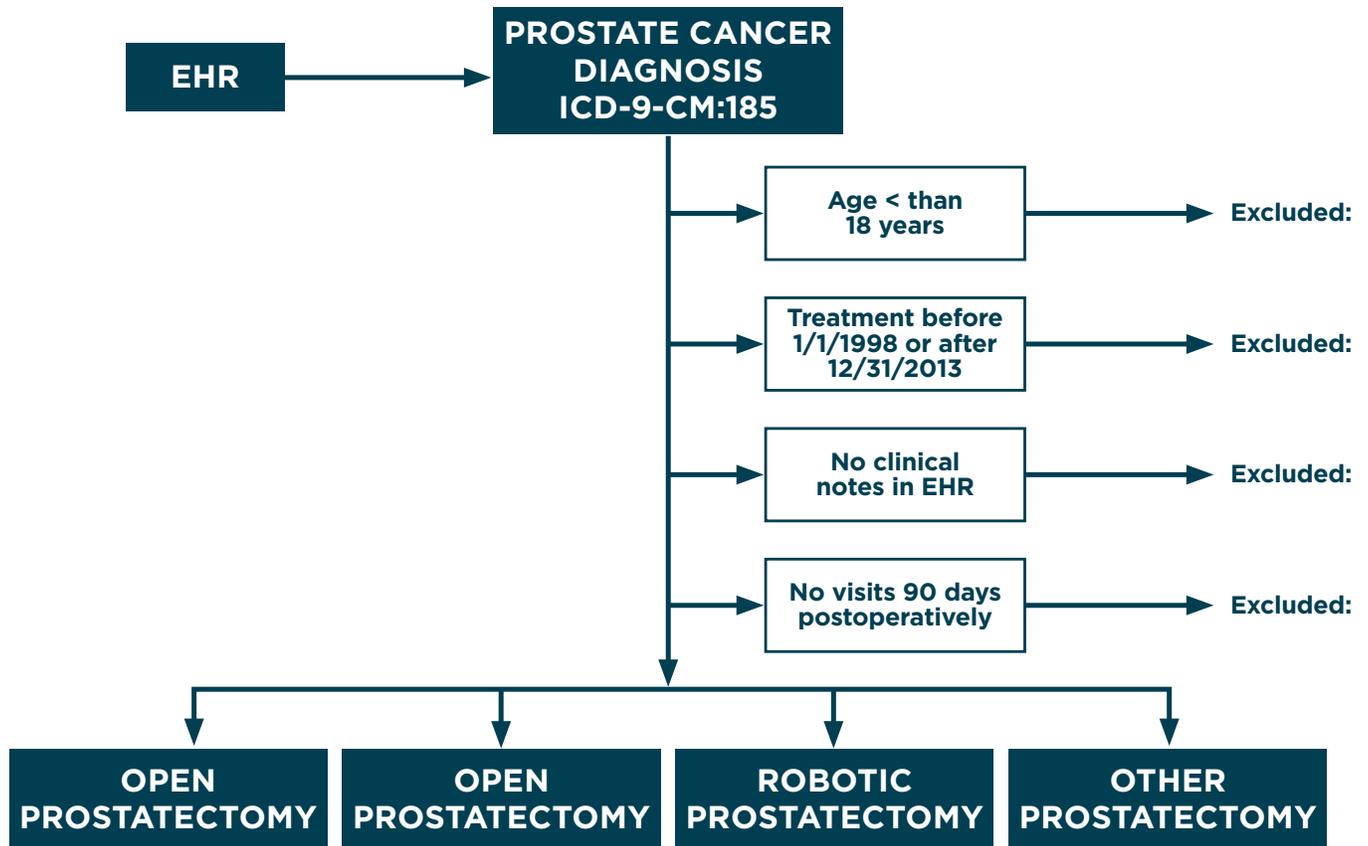
90 days postoperatively because they have no text notes to process for postoperative urinary incontinence.

Data Mining Workflow

Our data mining workflow used de-identified data from the institute's translational research data warehouse.²⁶ Urinary incontinence was identified using both structured data (ICD 9 CM: 788.30) and unstructured free text clinical notes (e.g., "urinary incontinence" or "urinary leakage"). To analyze free text, we used the NCBO Annotator to process our clinical notes.²⁷ The NCBO Annotator is a minimalist system that relies on a large dictionary of terms, their mappings to Unified Medical Language System (UMLS) concepts, and the NegEx negation detection system (a part of the ConText system)²⁸ to find mentions of biomedical concepts in clinical text and establish their negation status.^{27,29}

We customized our approach for identifying cases of urinary incontinence documented in free text using an approach that has been previously applied to develop task-specific extractors.³⁰ With the aim of improving sensitivity, we enhanced the annotator's terminology to include additional terms relevant to urinary incontinence (e.g., "wears adult diapers"). In addition, we extended the basic set of rules provided by NegEx to consider additional contextual information such as the following: hypothetical terms, e.g., "at risk for" (urinary incontinence); historical terms, e.g., "past history of" (urinary incontinence); and discussion terms, e.g., "discussed complications such as" (urinary incontinence).²⁹ After our workflow rules were applied, we defined "positive urinary incontinence mentions" as "those indicating documentation of a positive urinary incontinence case at the time of documentation" and all other types (e.g., negative, historical, or hypothetical mentions) of urinary incontinence as negative.

Figure 1. Cohort Selection Flowchart from Electronic Health Records



Our classifications were based on clinical information extracted from patient progress notes, consultations, referral reports, and postoperative notes; and on other types of unstructured free-text clinical notes available in the EHRs. We did not attempt to quantify the level of incontinence, we only identify if a patients' clinician reported any level of urinary incontinence or if they used an ICD 9 code for urinary incontinence. Our entire data-mining framework, which detects patient-centered outcomes from both structured and unstructured EHR data, can be executed on 1.8 million patient records (approximately 21 million clinical notes) in less than 24 hours on standard server hardware.

We performed a manual chart review on a subset of records to test the accuracy of the data mining workflow. For this review, 200 randomly selected entries were selected for review. A single reviewer was provided with a snippet of text surrounding the term of interest, urinary incontinence. The reviewer was blinded—the positive or negative determination of urinary incontinence from the workflow was not revealed. The reviewer marked each instance as positive or negative for urinary incontinence. Each instance corresponded to a single patient encounter. These results were used to calculate the positive predicted value and sensitivity of the workflow, standard performance tests for data mining algorithms.



The human subjects research review board of the participating institution approved this study.

Results

From 1998 to 2007, the inclusion of text notes in our EHR increased steadily. In 2008 our EPIC system was fully installed. Approximately half of all patient encounters contained some clinical note between 2008 and 2013. Patient demographics are presented in Table 1. Among the full cohort, 1485 patients had a text note in their EHR records.

The comparison of urinary incontinence recorded in patients' records is presented in Table 2. Of the 5,349 prostate cancer patients who were identified in our EHR, only 4 patient encounters had an ICD 9 CM code for urinary incontinence, yet 450 patients had urinary incontinence documented in the free

text note. Furthermore, in the free text note, 1,035 patients had documentation saying that the patient did not currently have urinary incontinence. For instances of urinary incontinence text mentions, our workflow had the following accuracy scores: positive predictive value 0.73 and sensitivity 0.84.

We display a number of patients with a text mention of urinary incontinence by postoperative follow-up in days (Figure 2). The number of patients seen postoperatively with a recording of a urinary incontinence assessment was 130, 177, and 417 for 30-, 60-, and 90 days, respectively. In this graph, patients may have multiple visits. As urinary incontinence can improve postoperatively, it is important to show that this patient-centered outcome is being assessed and documented beyond the first 30-day postoperative visit.

Table 1. Characteristics of Patients Receiving Prostatectomy for Localized Prostate Cancer, 1998–2013

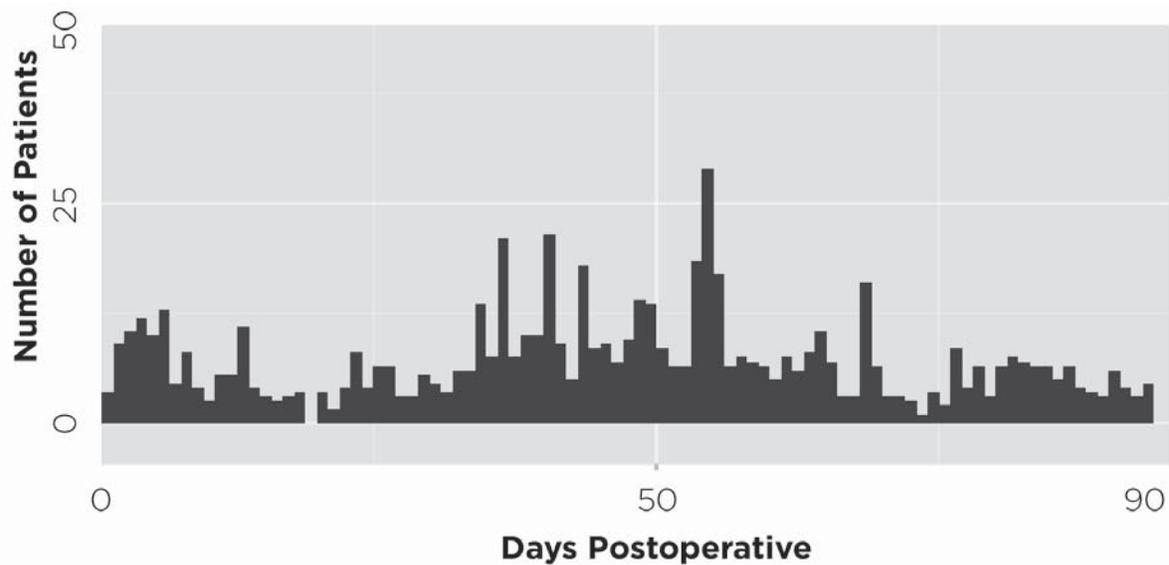
CHARACTERISTIC	TOTAL COHORT n = 5,353	COHORT WITH EHR NOTES n = 1,485
Age, mean(SD)	65.49 (0.14)	64.83 (0.26)
RACE, n(%)		
White	3,910 (73.04%)	995 (67.00%)
Black	130 (2.43%)	51 (3.43%)
Asian	362 (6.76%)	118 (7.95%)
Other	951 (17.77%)	321 (21.62%)
Hispanic, n(%)	216 (4.04%)	79 (5.32%)
SURGERY, n(%)		
Open	2,345 (43.81%)	564 (37.98%)
Robotic	385 (7.19%)	124 (8.35%)
Laparoscopic	414 (7.73%)	303 (20.40%)
Other	2,209 (41.27)	493 (33.27%)

Table 2. Postoperative Assessment of Urinary Incontinence Stratified by Structured Versus Unstructured Data Within the EHR

TYPE OF EHR INFORMATION	POSITIVE DOCUMENTATION OF URINARY INCONTINENCE	NEGATIVE DOCUMENTATION OF URINARY INCONTINENCE*	ABSENCE OF DOCUMENTATION OF URINARY INCONTINENCE
Text	450	1035	3868
ICD-9	4	n/a	5349

Note: *Negative Documentation refers to patients reporting that they are not suffering from urinary incontinence

Figure 2. Number of patients with a Mention of *Urinary Incontinence* in the EHR by Days from Surgery



Limitations

Note that we only report on what clinicians are documenting in the EHRs. These patient-centered outcomes reported by clinicians may vary from those reported by the patient. However, our data indicate that patient-centered outcomes, such as urinary incontinence, are documented in clinicians' text significantly more than they are recorded as coded data. Future studies should focus on the agreement between patient-reported and clinician-reported outcomes.

Discussion

Quality measurement is a means to monitor health care delivery and set benchmarks for timely, evidence-based care. With a disease such as prostate cancer, where survival is excellent, patient-centered outcomes might be among the best quality measures of health care delivered. In this study we found that urinary incontinence, an important patient-centered outcome following prostate cancer treatment, was reported almost exclusively in the free text of EHRs and was rarely coded as an ICD



9 diagnosis code. Here we tested the feasibility of efficiently and accurately extracting this patient-centered outcome from EHRs using standard data-mining techniques. This report provides evidence that patient-centered outcomes are recorded in EHRs and that these data can be efficiently and accurately extracted.

The widespread implementation of EHRs offers opportunities to support patient-centered care and quality improvement efforts.³¹ EHRs host a comprehensive set of care processes and outcomes, including outcomes other than physiological endpoints. EHRs capture clinicians' narrative text, images, and progress notes together with structured data elements. Over 80 percent of EHR data are captured as unstructured text, and here resides the rich, narrative text.^{32,33} The narrative text may contain information on patients' preferences, concerns, and often on patient-centered outcomes. However, the narrative text is stored as unstructured data (free text) that is difficult to assess using traditional measurement methods, which focus on structured data such as ICD-9-CM codes. Recent studies have used structured data within EHRs for quality improvement efforts³⁴⁻³⁶ and others have applied text-processing methods to sections of EHR (clinical notes, discharge notes, and pathology reports) for quality assessment.^{22,37-39} We extend these methods to include patient-centered outcomes.

Mining existing structured and unstructured EHR data for patient-centered outcomes has several immediate benefits and efficiencies. First, we have shown that longitudinal narrative data for these patient-centered outcomes are in the EHR. These data exist mainly in the narrative text and not in the structured data, so EHR studies must look beyond coded data. Indeed, our research found that urinary incontinence, one of the most reported outcomes with known effects on health-related quality of life

following prostate cancer treatment,⁴⁰ was almost exclusively reported in EHR free text. Second, studies derived from EHR data do not inherently contain ascertainment bias, as do many survey-based and prospective studies.⁴¹ EHR data exist across populations, care settings, and socioeconomic status, thus eliminating many of these known biases. Third, data-mining algorithms now allow for efficient processing, and for retrieval of data. It is clear that this is significantly advantageous over manual chart review, as previously noted.⁴²

Extracting and analyzing patient-centered outcome data in a precise and timely manner is the first step in creating treatment pathways that reflect the patients' individual risk values. Using prostatectomy as an example, if robotic surgery has a 20 percent relative risk of urinary incontinence and a 30 percent relative risk of erectile dysfunction and open prostatectomy has a 30 percent relative risk of urinary incontinence and a 20 percent relative risk of erectile dysfunction, patients can make informed treatment decisions based on their personal values of these different risks, which is a highlight of patient-centered care.⁴³ The patient's perspective of risk can be incorporated into the treatment pathway only if we have valid and accurate rates of these important patient-centered outcomes—for which evidence is currently limited.¹³

To move to a value-based care system, we must expand our measures of quality beyond simple coded data and include a comprehensive set of health care outcomes. As prostate cancer has excellent survival, patient-centered outcomes should be reflected in the quality measures used to assess the disease treatment. Our data indicate that important patient-centered outcomes, such as urinary incontinence, are being captured in EHRs as free text. This highlights the long-standing importance of accurate clinician documentation.

Development of generalizable benchmarks and accurate and complete assessment of these outcomes are essential to move practice into the patient-centered realm of health care.

Acknowledgements

Research reported in this presentation was supported by the National Cancer Institute of the National Institutes of Health under Award Number R01CA183962. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

- Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. *CA: a cancer journal for clinicians*. 2014;64(1):9-29.
- Penson DF, McLerran D, Feng Z, et al. 5-year urinary and sexual outcomes after radical prostatectomy: results from the prostate cancer outcomes study. *The Journal of urology*. 2005;173(5):1701-1705.
- Litwin MS, Hays RD, Fink A, et al. Quality-of-life outcomes in men treated for localized prostate cancer. *JAMA : the journal of the American Medical Association*. 1995;273(2):129-135.
- Stanford JL, Feng Z, Hamilton AS, et al. Urinary and sexual function after radical prostatectomy for clinically localized prostate cancer: the Prostate Cancer Outcomes Study. *JAMA : the journal of the American Medical Association*. 2000;283(3):354-360.
- Spencer BA, Steinberg M, Malin J, Adams J, Litwin MS. Quality-of-care indicators for early-stage prostate cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2003;21(10):1928-1936.
- Miller DC, Sanda MG, Dunn RL, et al. Long-term outcomes among localized prostate cancer survivors: health-related quality-of-life changes after radical prostatectomy, external radiation, and brachytherapy. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2005;23(12):2772-2780.
- Ellison LM, Trock BJ, Poe NR, Partin AW. The effect of hospital volume on cancer control after radical prostatectomy. *The Journal of urology*. 2005;173(6):2094-2098.
- Sanda MG, Dunn RL, Michalski J, et al. Quality of life and satisfaction with outcome among prostate-cancer survivors. *N Engl J Med*. 2008;358(12):1250-1261.
- Resnick MJ, Koyama T, Fan KH, et al. Long-term functional outcomes after treatment for localized prostate cancer. *N Engl J Med*. 2013;368(5):436-445.
- Federman DG, Pitkin P, Carbone V, Concato J, Kravetz JD. Screening for prostate cancer: are digital rectal examinations being performed? *Hospital practice*. 2014;42(2):103-107.
- Yu HY, Hevelone ND, Lipsitz SR, Kowalczyk KJ, Hu JC. Use, costs and comparative effectiveness of robotic assisted, laparoscopic and open urological surgery. *The Journal of urology*. 2012;187(4):1392-1398.
- Yu HY, Hevelone ND, Lipsitz SR, Kowalczyk KJ, Nguyen PL, Hu JC. Hospital volume, utilization, costs and outcomes of robot-assisted laparoscopic radical prostatectomy. *The Journal of urology*. 2012;187(5):1632-1637.
- Wilt TJ, MacDonald R, Rutks I, Shamliyan TA, Taylor BC, Kane RL. Systematic review: comparative effectiveness and harms of treatments for clinically localized prostate cancer. *Annals of internal medicine*. 2008;148(6):435-448.
- Thompson I, Thrasher JB, Aus G, et al. Guideline for the management of clinically localized prostate cancer: 2007 update. *The Journal of urology*. 2007;177(6):2106-2131.
- U.S. Department of Health & Human Services. The Patient Protection and Affordable Care Act. Washington, D.C.: U.S. Government;; 2010.
- VanLare JM, Blum JD, Conway PH. Linking performance with payment: implementing the Physician Value-Based Payment Modifier. *JAMA*. 2012;308(20):2089-2090.
- Kuperman GJ. Health-information exchange: why are we doing it, and what are we doing? *Journal of the American Medical Informatics Association : JAMIA*. 2011;18(5):678-682.
- Selby JV, Beal AC, Frank L. The Patient-Centered Outcomes Research Institute (PCORI) national priorities for research and initial research agenda. *JAMA : the journal of the American Medical Association*. 2012;307(15):1583-1584.
- Epstein RM, Fiscella K, Lesser CS, Stange KC. Why the nation needs a policy push on patient-centered health care. *Health Aff (Millwood)*. 2010;29(8):1489-1495.
- Capurro D, van Eaton E, Black R, Tarczy-Hornoch P. Availability of Structured and Unstructured Clinical Data for Comparative Effectiveness Research and Quality Improvement: A Multisite Assessment. *EGEMS*. 2014;2(1).
- D'Avolio LW, Litwin MS, Rogers SO, Jr., Bui AA. Facilitating Clinical Outcomes Assessment through the automated identification of quality measures for prostate cancer surgery. *Journal of the American Medical Informatics Association : JAMIA*. 2008;15(3):341-348.
- Harkema H, Chapman WW, Saul M, Dellon ES, Schoen RE, Mehrotra A. Developing a natural language processing application for measuring the quality of colonoscopy procedures. *Journal of the American Medical Informatics Association : JAMIA*. 2011;18 Suppl 1:i150-156.
- Turchin A, Kolatkar NS, Grant RW, Makhni EC, Pendergrass ML, Einbinder JS. Using regular expressions to abstract blood pressure and treatment intensification information from the text of physician notes. *Journal of the American Medical Informatics Association : JAMIA*. 2006;13(6):691-695.
- Bates DW, Saria S, Ohno-Machado L, Shah A, Escobar G. Big data in health care: using analytics to identify and manage high-risk and high-cost patients. *Health Aff (Millwood)*. 2014;33(7):1123-1131.



25. Kho AN, Hayes MG, Rasmussen-Torvik L, et al. Use of diverse electronic medical record systems to identify genetic risk for type 2 diabetes within a genome-wide association study. *Journal of the American Medical Informatics Association : JAMIA*. 2012;19(2):212-218.
26. Lowe HJ, Ferris TA, Hernandez PM, Weber SC. STRIDE--An integrated standards-based translational research informatics platform. *AMIA ... Annual Symposium proceedings / AMIA Symposium*. AMIA Symposium. 2009;2009:391-395.
27. Shah NH, Bhatia N, Jonquet C, Rubin D, Chiang AP, Musen MA. Comparison of concept recognizers for building the Open Biomedical Annotator. *BMC bioinformatics*. 2009;10 Suppl 9:S14.
28. Harkema H, Dowling JN, Thornblade T, Chapman WW. ConText: an algorithm for determining negation, experimenter, and temporal status from clinical reports. *Journal of biomedical informatics*. 2009;42(5):839-851.
29. Chapman WW, Bridewell W, Hanbury P, Cooper GF, Buchanan BG. A simple algorithm for identifying negated findings and diseases in discharge summaries. *Journal of biomedical informatics*. 2001;34(5):301-310.
30. Tamang S, Patel MI, Blayney DW, et al. Detecting unplanned care from clinician notes in electronic health records. *Journal of oncology practice / American Society of Clinical Oncology*. 2015;11(3):e313-319.
31. Snyder CF, Jensen RE, Segal JB, Wu AW. Patient-reported outcomes (PROs): putting the patient perspective in patient-centered outcomes research. *Med Care*. 2013;51(8 Suppl 3):S73-79.
32. Grimes S. Unstructured data and the 80 percent rule. 2008; <http://breakthroughanalysis.com/2008/08/01/unstructured-data-and-the-80-percent-rule/>. Accessed September 6, 2015.
33. Murdoch TB, Detsky AS. The inevitable application of big data to health care. *JAMA*. 2013;309(13):1351-1352.
34. Persell SD, Kho AN, Thompson JA, Baker DW. Improving hypertension quality measurement using electronic health records. *Medical care*. 2009;47(4):388-394.
35. Tang PC, Ralston M, Arrigotti MF, Qureshi L, Graham J. Comparison of methodologies for calculating quality measures based on administrative data versus clinical data from an electronic health record system: implications for performance measures. *Journal of the American Medical Informatics Association : JAMIA*. 2007;14(1):10-15.
36. Steidl M, Zimmern P. Data for free--can an electronic medical record provide outcome data for incontinence/prolapse repair procedures? *The Journal of urology*. 2013;189(1):194-199.
37. Murff HJ, FitzHenry F, Matheny ME, et al. Automated identification of postoperative complications within an electronic medical record using natural language processing. *JAMA : the journal of the American Medical Association*. 2011;306(8):848-855.
38. Chiang JH, Lin JW, Yang CW. Automated evaluation of electronic discharge notes to assess quality of care for cardiovascular diseases using Medical Language Extraction and Encoding System (MedLEE). *Journal of the American Medical Informatics Association : JAMIA*. 2010;17(3):245-252.
39. Anderson C. EHR measure aims at halting heart disease. *PhysBizTech*. 2013.
40. Person DF, Litwin MS. Quality of life after treatment for prostate cancer. *Curr Urol Rep*. 2003;4(3):185-195.
41. Goldberg RJ, McManus DD, Allison J. Greater knowledge and appreciation of commonly-used research study designs. *Am J Med*. 2013;126(2):169 e161-168.
42. Parvizi J, Miller AG, Gandhi K. Multimodal pain management after total joint arthroplasty. *The Journal of bone and joint surgery. American volume*. 2011;93(11):1075-1084.
43. Barry MJ, Edgman-Levitan S. Shared decision making--pinnacle of patient-centered care. *N Engl J Med*. 2012;366(9):780-781.