


Electronic Health Record Phenotypes for Identifying Patients with Late-Stage Disease: a Method for Research and Clinical Application



Natalie C. Ernecoff, PhD, MPH¹ , Kathryn L. Wessell, BS², Laura C. Hanson, MD, MPH^{2,3}, Adam M. Lee, MBA⁴, Christopher M. Shea, PhD, MA, MPA⁵, Stacie B. Dusetzina, PhD⁶, Morris Weinberger, PhD, MS⁵, and Antonia V. Bennett, PhD⁵

¹Division of General Internal Medicine, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA; ²Sheps Center for Health Services Research, University of North Carolina-Chapel Hill, Chapel Hill, NC, USA; ³Division of Geriatric Medicine & Palliative Care Program, University of North Carolina-Chapel Hill, Chapel Hill, NC, USA; ⁴North Carolina Translational and Clinical Sciences Institute, University of North Carolina-Chapel Hill, Chapel Hill, NC, USA; ⁵Department of Health Policy and Management, Gillings School of Global Public Health, University of North Carolina-Chapel Hill, Chapel Hill, NC, USA; ⁶Department of Health Policy, Vanderbilt University Medical Center, Nashville, TN, USA.

BACKGROUND: Systematic identification of patients allows researchers and clinicians to test new models of care delivery. EHR phenotypes—structured algorithms based on clinical indicators from EHRs—can aid in such identification.

OBJECTIVE: To develop EHR phenotypes to identify decedents with stage 4 solid-tumor cancer or stage 4–5 chronic kidney disease (CKD).

DESIGN: We developed two EHR phenotypes. Each phenotype included International Classification of Diseases (ICD)-9 and ICD-10 codes. We used natural language processing (NLP) to further specify stage 4 cancer, and lab values for CKD.

SUBJECTS: Decedents with cancer or CKD who had been admitted to an academic medical center in the last 6 months of life and died August 26, 2017–December 31, 2017.

MAIN MEASURE: We calculated positive predictive values (PPV), false discovery rates (FDR), false negative rates (FNR), and sensitivity. Phenotypes were validated by a comparison with manual chart review. We also compared the EHR phenotype results to those admitted to the oncology and nephrology inpatient services.

KEY RESULTS: The EHR phenotypes identified 271 decedents with cancer, of whom 186 had stage 4 disease; of 192 decedents with CKD, 89 had stage 4–5 disease. The EHR phenotype for stage 4 cancer had a PPV of 68.6%, FDR of 31.4%, FNR of 0.5%, and 99.5% sensitivity. The EHR phenotype for stage 4–5 CKD had a PPV of 46.4%, FDR of 53.7%, FNR of 0.0%, and 100% sensitivity.

CONCLUSIONS: EHR phenotypes efficiently identified patients who died with late-stage cancer or CKD. Future EHR phenotypes can prioritize specificity over sensitivity, and incorporate stratification of high- and low-palliative care need. EHR phenotypes are a promising method for identifying patients for research and clinical purposes, including equitable distribution of specialty palliative care.

J Gen Intern Med 34(12):2818–23

DOI: 10.1007/s11606-019-05219-9

© Society of General Internal Medicine 2019

INTRODUCTION

Palliative care services improve quality of life, reduce intensive treatments, and in some settings extend survival for patients with late-stage serious illness.¹ Experts are calling for additional research to maximize uptake of palliative care by expanding and implementing existing interventions to improve access to specialty palliative care.^{2–4} Patients with late-stage cancer and chronic kidney disease (CKD) have high symptom burden and require complex decision making, particularly around initiation and discontinuation of treatments such as radiation therapy, chemotherapy, dialysis, and transitions to hospice. Generally, specialty palliative care is referral-based, and access is dependent on the referring provider and availability of a palliative care clinic or inpatient service.⁵ An important step for improving access is systematically identifying patients with late-stage disease who could benefit from referral to palliative care.⁶ However, unlike vital signs or lab results, disease stage is not consistently reported in structured data fields in electronic health records (EHRs). Patients with late-stage disease are thus difficult to distinguish for purposes of palliative care research or for targeted referral to clinical services.

EHR phenotypes can be used to identify populations of patients with specific conditions or service needs. EHR phenotypes are customizable algorithms that query EHR data, typically for International Classification of Disease (ICD)-9 or ICD-10 codes, lab values, or demographic or clinical characteristics. In prior work, our research team found that an EHR phenotype was an efficient mechanism to identify hospitalized patients with late-stage dementia for enrollment in a palliative care

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11606-019-05219-9>) contains supplementary material, which is available to authorized users.

Received July 9, 2019

Accepted July 12, 2019

Published online August 8, 2019

clinical trial based on ICD codes for dementia or cognitive impairment, age, and inpatient status.^{7, 8} Traditionally, EHR phenotypes have been built and applied to identify genotype and biologic data.^{9, 10} Although EHR phenotypes typically rely on quantitative data fields, natural language processing (NLP) offers an additional avenue for defining and refining EHR phenotypes. NLP is a method by which one can search text (e.g., notes in the EHR) to find information based on free text, including key words and synonyms, that would not otherwise be identifiable from structured data fields. NLP is a method by which one can search text (e.g., notes in the EHR) to find information based on free text that would not otherwise be identifiable from structured data fields.¹¹

To date, informatics approaches to palliative care have rarely been applied due to the difficulty in generating effective EHR phenotypes.¹² In general, EHR phenotypes can be difficult to develop when based solely on structured data (e.g., ICD codes) due to cost of data and software requirements, though those barriers are becoming easier to overcome. For palliative care and many disciplines, effective interpretation of data relies largely on unstructured data, compared with fields that can rely on structured data (e.g., genotypes, lab values). Incorporating NLP into EHR phenotypes offers a novel approach to systematically and accurately identify patients who may benefit from palliative care services and may be able to make EHR phenotypes more accurate by encompassing unstructured data. For example, ICD codes can be used to indicate diagnostic testing rather than actual diagnoses; in these instances, properly calibrated NLP has potential to improve specificity. While many clinical interventions and research projects depend on referral and manual screening of census or clinic lists, EHR phenotypes (including NLP) may reduce time and effort for staff. Effective EHR phenotypes can support more thorough identification of patient populations in both clinical and research settings.

As part of a retrospective study to assess receipt of palliative care in the last 6 months of life among decedents with serious illness (cancer and CKD), we sought to use structured queries and NLP (in cancer) to develop two separate EHR phenotypes to identify decedents with¹ stage 4 solid-tumor cancer or² stage 4–5 CKD. We then tested the EHR phenotype on a cohort of decedents who had been admitted to the hospital at least once in their last 6 months of life to determine its ability to identify decedents with late-stage disease.

METHODS

EHR phenotype development requires several steps, as noted here and described in detail below. These include (1) defining inclusion and exclusion criteria for each disease cohort; (2) testing and refining criteria by comparing them with a

clinically defined “gold standard” validation group; and (3) analyzing each EHR phenotype’s effectiveness in appropriately including and excluding decedents with late-stage disease.

Setting and Cohort Eligibility

The EHR phenotypes for this study were developed and tested at the University of North Carolina (UNC) Medical Center, with 929 adult beds with nearly 30,000 non-obstetric admissions in FY17. The research dataset consisted of EHR records for all decedents who had the following¹: a date of death between August 26, 2017, and December 31, 2017; and² an acute, non-planned hospitalization lasting at least 24 h during the 6 months preceding death. Decedents were excluded if they were prisoners at the time of admission, were < 18 years of age, or had been at an outside hospital for > 24 h before transfer to UNC Medical Center. All study procedures were approved by the University of North Carolina Institutional Review Board.

Development and Refinement of the EHR Phenotype

To construct the EHR phenotype, we used data from the EHR linked by name and medical record number (MRN) to the state death certificate database. We used Informatics for Integrating Biology and Bedside (i2b2, Boston, MA) software to search structured data fields to return a de-identified count of decedents who meet the criteria.¹³ We linked the de-identified cohort from i2b2 to names and MRN from the Carolina Data Warehouse (IBM Business Objects, Armonk, NY), which stores both UNC Medical Center EHR data and North Carolina state death certificate data, to build a list of decedents who met the inclusion criteria. A list was generated to a secure research server with the identifying information of those decedents who met the parameters.

Two researchers (NCE, KKW) with extensive experience in chart review and EHR-based research were trained on the specific eligibility criteria. Upon receiving output from the EHR phenotypes, one reviewer then screened each medical record for evidence of stage 4 cancer or stage 4–5 CKD, respectively, by reviewing structured data in the chart and all physician or advance practice provider notes. Decedents with unclear disease classification were assessed by the second researcher, and the two reviewers deliberated to reach a consensus judgment. A physician was available for adjudication (LCH).

Stage 4 cancer EHR phenotype: We identified 1081 ICD-10 codes and 14 categories of cancer-related ICD-9 codes to identify any malignant solid tumor (all stages); we also included 69 ICD-10 codes and 4 ICD-9 categories of codes that specifically designated metastases or stage 4 disease. This first iteration of the phenotype identified 272 decedents, 186 of whom were confirmed by chart review to have stage 4 solid tumor disease. (See [Appendix](#) for the complete list of ICD codes.) For the

second iteration, we refined the EHR phenotype by incorporating NLP techniques using EMERSE (Electronic Medical Record Search Engine; Ann Arbor, MI), a software that searches text- and note-based fields within the EHR.¹⁴ Because the NLP was intended to refine the EHR phenotypes by making them more specific (rather than more sensitive), the cohort of decedents identified by the first structured iteration of the EHR phenotype were uploaded into EMERSE. We used key word searches by adding the search terms “stage 4,” “stage iv,” “stg 4,” “metastatic,” and “mets” in all EHR notes to eliminate decedents from the cohort who did not also have one of the search terms. “Mets” and “stg 4” did not appear in any true positive results without at least one other search term, so we eliminated that term from the final search criteria. The final search criteria yielded 271 decedents, 186 (68.6%) of whom were confirmed to have stage 4 solid tumor cancers.

Stage 4–5 CKD EHR phenotype: We started by identifying 80 ICD-10 codes and 3 ICD-9 codes associated with stage 4–5 CKD; this first iteration identified 259 decedents, 89 of whom were confirmed to have late-stage disease. For the second iteration, we incorporated lab values for glomerular filtration rate (GFR) associated with stage 4–5 CKD. We started by broadly requiring that a record indicates at least one value for which $GFR < 30$ ml/min, which is associated with stage 4 ($GFR 15–29$ ml/min) and stage 5 ($GFR < 15$ ml/min) CKD.¹⁵ At that point, the phenotype identified 192 decedents, 89 of whom were confirmed to have late-stage disease.

We further refined the GFR values in a third iteration by requiring at least 2 values below 30 ml/min within 90 days, which is consistent with the time-based clinical definition for CKD.¹⁵ At that point, the date range yielded 171 decedents, 86 (50.3%) of whom were confirmed to have late-stage disease.

Validation Cohort

To validate the EHR phenotypes, we compared the lists of decedents compiled by the EHR phenotypes with a gold standard of lists of decedents admitted to inpatient services led by subspecialty physicians with expertise for the respective disease states. For cancer, we used the list of all decedents who had been admitted to services led by medical or gynecologic oncologists from September 3, 2017, through December 31, 2017 ($n = 122$). Similarly, we used the list of decedents with CKD admitted to the inpatient service led by nephrologists from August 26, 2017, through December 31, 2017 ($n = 21$). We selected primary inpatient services as the comparison groups because they have the greatest expertise in diagnoses and staging these diseases. Further, clinical research often identifies patients via a disease group’s primary clinic or admitting service. We assessed whether an EHR phenotype could potentially increase the count of patients assessed for eligibility for research studies, a common barrier to recruitment in

clinical research. The same trained researchers reviewed all medical records from the services led by respective subspecialty physicians to identify the number of decedents with each late-stage disease, respectively, $n = 74$ decedents with stage 4 cancer and $n = 14$ decedents with stage 4–5 CKD. Decedent counts for all cohorts are displayed in Table 1.

Positive Predictive Value, False Discovery Rate, and Sensitivity

All calculations were conducted separately for stage 4 cancer and stage 4–5 CKD.

Positive predictive value (PPV) was defined as the percent of cases identified by the EHR phenotype that did have late-stage disease as confirmed on chart review. The false discovery rate (FDR) was defined as the percent of cases identified by the EHR phenotype that did not have late-stage disease based on chart review. The false negative rate (FNR) was defined as the percent of case who were on the gold standard lists of decedents admitted to inpatient services led by subspecialty experts (for cancer, medical or gynecologic oncologists; for CKD, nephrologists) who were not identified by the EHR phenotype, relative to decedents with late-stage disease identified by the EHR phenotype confirmed by chart review.

We calculated the sensitivity, defined as the ratio of those identified by the EHR phenotype who had late-stage disease as confirmed on chart review to the total from both the EHR phenotype and gold standard who had late-stage disease during the same timeframe. To examine the EHR phenotypes’ added value, we also describe the number of cases confirmed to be late stage on chart review that were identified by the EHR phenotypes and not admitted to the primary services used for the gold standard.

Table 1 Counts of Decedents Identified by the EHR Phenotypes and from the Validation Cohort

	Stage 4 cancer	Stage 4–5 chronic kidney disease (CKD)
Total no. in first iteration	272	259
No. in first iteration with late-stage diagnosis	186 (68.4% of total in iteration)	89 (34.4% of total in iteration)
Total no. in second iteration (cancer NLP; CKD 1xGFR < 30)	271	192
No. in second iteration with late-stage diagnosis	186 (68.6% of total in iteration)	89 (46.4% of total in iteration)
Total no. in third iteration (2xGFR < 30 within 90 days)		171
No. in third iteration with late-stage diagnosis		86 (50.3% of total in iteration)
Total no. in validation cohort	122	21
No. in validation cohort with late-stage diagnosis	74 (60.7% of validation cohort)	14 (66.7% of validation cohort)

GFR glomerular filtration rate, *laboratory value*

RESULTS

Positive Predictive Value, False Discovery Rate, False Negative Rate, and Sensitivity of the EHR Phenotypes

Cancer EHR Phenotype. The first iteration of the EHR phenotype for cancer had a PPV of 68.4% (186/272). After adding the NLP components to the EHR phenotype for cancer, the second iteration had virtually identical PPV (68.6%; 186/271). Likewise, the FDR of both iterations was very similar, respectively, 31.6% (85/272) and 31.4% (85/271). The EHR phenotype for cancer had a sensitivity of 99.5% for both iterations, meaning it identified nearly all late-stage cancer decedents in the EHR. Both iterations of the EHR phenotype for cancer had almost no false negatives, with a FNR for the more specific (i.e., less sensitive) final phenotype of 0.5% (1/187) when compared with the gold standard of stage 4 cancer confirmed by chart review on the subspecialty service lists. Conversely, the final phenotype had a sensitivity of 99.5% (186/187), identifying nearly all decedents with stage 4 cancer on the gold standard list.

CKD EHR Phenotype. The first iteration of the EHR phenotype for CKD had a PPV of 34.4% (89/259). The second (requiring at least one GFR < 30 within 90 days) and third iterations (2xGFR < 30 within 90 days) improved the PPV to 46.4% (89/192) and 50.3% (86/171), respectively. Likewise, the FDR for the three iterations were 65.6% (170/259), 53.7% (103/192), and 49.7% (85/171), respectively. All iterations of the EHR phenotype for CKD had a FNR of 0.0% (final iteration 0/86) compared with the gold standard of stage

4–5 CKD confirmed by chart review on the subspecialty service lists. Likewise, all three iterations had a sensitivity of 100% (final iteration 86/86), identifying all decedents with stage 4–5 CKD on the gold standard list.

Calculations and final PPV, FDR, FNR, and sensitivity for each EHR phenotype are presented in Table 2.

EHR Phenotypes' Added Value Beyond the Gold Standard

As an additional evaluation of the EHR phenotypes' capacity for identifying the two serious-illness populations efficiently and completely, we examined how many decedents were hospitalized but cared for by non-subspecialty services. For stage 4 cancer, the EHR phenotype identified 112 additional decedents who were not admitted to a service led by an oncology subspecialist (60.2% of the total population of persons with stage 4 cancer who were identified by the EHR phenotype). Likewise, the final EHR phenotype for stage 4–5 CKD identified 72 decedents who were not admitted to a hospital service led by a nephrology subspecialist (83.7% of persons with stage 4–5 CKD who were identified by the EHR phenotype). Thus, the EHR phenotype identified many more decedents with late-stage illness than a search based on the specialty designation of attending physicians.

DISCUSSION

We developed and tested separate EHR phenotypes for stage 4 cancer and stage 4–5 CKD that efficiently identified decedents with serious illness who were hospitalized in their last

Table 2 Results of the Final Computable Phenotypes for, Respectively, Stage 4 Cancer and Stage 4–5 Chronic Kidney Disease (CKD)

Formula	Calculation of formula components with project-specific data	Project-specific result	Rate (%)
Late-stage cancer			
Positive Predictive Value (PPV) = True Positive (TP) / (True Positive (TP) + False Positive (FP))	TP = EHR phenotype Stage 4 CA FP = patients from EHR phenotype without stage 4 CA	186 85	68.6
False Discovery Rate (FDR) = False Positive (FP) / (True Positive (TP) + False Positive (FP))	FP = Patients from EHR phenotype without stage 4 CA	85	31.4
False Negative Rate = False Negative (FN) / (False Negative(FN) + True Positive (TP))	TP = EHR phenotype stage 4 CA FN = Stage 4 CA on a primary oncology service AND not identified by EHR phenotype	186 1	0.5
Sensitivity = True Positive (TP) / (True Positive (TP) + False Negative (FN))	TP = EHR phenotype stage 4 CA FN = Stage 4 CA on a primary oncology service AND not identified by EHR phenotype	186 1	99.5
Late-stage CKD			
Positive Predictive Value (PPV) = True Positive (TP) / (True Positive (TP) + False Positive (FP))	TP = EHR phenotype stage 4–5 CKD FP = Patients from EHR phenotype without stage 4–5 CKD	86 85	50.3
False Discovery Rate (FDR) = False Positive (FP) / (True Positive (TP) + False Positive (FP))	TP = EHR phenotype Stage 4–5 CKD FP = Patients from EHR phenotype without stage 4–5 CKD	86 85	49.7
False Negative Rate = False Negative (FN) / (False Negative(FN) + True Positive (TP))	FN = stage 4–5 CKD on a primary nephrology service AND not identified by EHR phenotype	0	0
Sensitivity = True Positive (TP) / (True Positive (TP) + False Negative (FN))	TP = EHR phenotype stage 4–5 CKD FN = stages 4–5 CKD on a primary oncology service AND not identified by EHR phenotype	86 86 0	100

6 months of life. The EHR phenotypes successfully identified nearly all decedents with late-stage disease cared for by relevant subspecialists and a large number of decedents with late-stage disease who were cared for by other attending physicians. The ability of EHR phenotypes to identify patients across the hospital is promising for both improving systematic clinical palliative care delivery and identifying patients for research. The added benefit of the EHR phenotype is that it becomes feasible to identify patients with cancer, for example, across the entire hospital. Although patients with cancer may very reasonably be admitted to other services across the hospital, they would indeed not have been included in a primary data collection study that requires manual screening, which would often in practice be limited to the oncology service. These phenotypes had relatively few false positives and can thus be used to expedite manual screening. For this reason, we prioritized sensitivity over specificity to still get a (nearly) complete pool of decedents eligible for inclusion in the study cohort while making manual review feasible as it would not have been if researchers had to screen the entire inpatient census. For comparison, one NLP system was designed to identify primary diagnosis that had an accuracy of 82% when excluding cases with “insufficient data.”¹⁶

Specialty palliative care in most hospital settings is a consult service, meaning the primary clinical service to which a patient is admitted must contact the palliative care team to see an individual patient.¹⁷ This consulting model often results in inconsistent referral patterns.¹⁸ Because EHR phenotypes have the potential to identify patients who could benefit from symptom management or discussion of goals of care,^{7, 19} current referral-based models of palliative care delivery can be improved by incorporating automated EHR phenotypes across a hospital. Currently, automated best practice alerts (BPAs) are increasingly being implemented in EHRs to flag patients who may be appropriate for screening or services and to enhance adherence to best practices.^{20–22} EHR phenotypes such as those in this study can be incorporated into BPAs to identify patients in real time who may benefit from healthcare that is specific to the needs of late-stage disease, including, but not limited to, palliative care.

Systematically identifying patients with late-stage disease is also relevant for descriptive and interventional research. The use of EHR phenotypes can expedite screening processes to identify patients who are potentially eligible for research protocols.⁸ Further, EHR phenotypes can systematize identification of patients in research using secondary data sources, potentially improving validity and generalizability of research findings.²³ Our study illustrates the utility of EHR phenotypes, as we identified twice as many decedents with late-stage disease by not limiting searches to the primary clinical services, indicating that patients with late-stage serious illness were not exclusively clustered with subspecialty attending physicians or hospital units.²⁴ Therefore, although much of research recruitment and enrollment currently is focused at the unit level, EHR phenotypes offer enhanced utility when a patient population is diffused throughout a healthcare system.

Limitations

Our study has several limitations. First, our gold standard list of decedents excludes those with late-stage cancer or CKD who may be admitted to services that do not correspond with those disease groups; further, primary subspecialty services corresponding with cancer and CKD, respectively, may be more likely to document ICD codes and diagnoses more systematically, limiting the observed effect of NLP. As discussed above, these limitations extend to current clinical and research practice, so the comparison used in this study reflects real-world patient identification. Second, this research was conducted in one academic medical center, which limits generalizability due to differences in syntax that may be used in notes across institutions. Third, this research utilized cohorts of decedents to test and validate the EHR phenotypes; components of the phenotypes, including ICD codes, may be retrospectively applied to patient records upon death. Although we waited the designated lag time for state death records to be completed, we may not have captured all deaths.²⁵ Phenotypes should be tested prospectively in order to better understand their utility for living populations.

Future Research

Future research can build more finely titrated EHR phenotypes for patients with late-stage disease, incorporating evidence for high palliative care needs such as symptom distress, frequent hospitalization, or lack of advance care planning.²⁶ Refined EHR phenotypes will enhance automation in both clinical and research settings.²⁷ Some components of these definitions are already in development, particularly around algorithms to predict prognosis as patients approach the end of life.^{28–30} Notably, in this cohort, ICD-10 codes were used with high efficiency for the cohort with cancer, so the NLP did not improve specificity to a high degree. Similar approaches may be applicable to other disease states, particularly when structured data fields (e.g., laboratory values, ICD codes) do not facilitate traditional EHR phenotypes. EHR phenotypes can also be adapted based on the specific clinical or research need, to prioritize specificity over sensitivity. Implementation is also limited to healthcare settings' access to structured EHR data that is able to be queried and a software (such as i2b2) that can run the EHR phenotypes. Additionally, future EHR phenotypes can be designed to prioritize specificity over sensitivity for research and clinical projects that target only patients with late-stage disease.

CONCLUSION

We built two EHR phenotypes that provide an efficient mechanism for identifying a cohort of late-stage decedents near the end of life in the inpatient setting. Such phenotypes can support both clinical practice and research by systematically identifying patients who may benefit from palliative care.

Corresponding Author: Natalie C. Ernecoff, PhD, MPH; Division of General Internal Medicine, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA (e-mail: ernecoff.natalie@pitt.edu).

Funding Information This work was supported by the North Carolina Translational and Clinical Sciences Institute (NC TraCS), Clinical and Translational Science Awards (CTSA) grant no. UL1TR002489.

Compliance with Ethical Standards:

Conflict of Interest: The authors declare that they do not have a conflict of interest.

REFERENCES

- Kavaliaratos D, Corbelli J, Zhang D, Dionne-Odom JN, Ernecoff NC, Hanmer J, et al. Association Between Palliative Care and Patient and Caregiver Outcomes. *JAMA*. 2016 Nov 22;316(20):2104.
- Institute of Medicine (US). Dying in America: Improving Quality and Honoring Individual Preferences Near the End of Life : Health and Medicine Division. 2015.
- O'hare AM, Song M-K, Tamura MK, Moss AH. Priorities in Palliative Care Research II Research Priorities for Palliative Care for Older Adults with Advanced Chronic Kidney Disease.
- O'Hare AM, Armistead N, Schrag WLF, Diamond L, Moss AH. Patient-Centered Care: An Opportunity to Accomplish the "Three Aims" of the National Quality Strategy in the Medicare ESRD Program. *Clin J Am Soc Nephrol*. 2014 Dec 5;9(12):2189-94.
- Quill TE, Abernethy AP. Generalist plus Specialist Palliative Care — Creating a More Sustainable Model. *N Engl J Med*. 2013 Mar 28;368(13):1173-5.
- Kelley AS, Bollens-Lund E. Identifying the Population with Serious Illness: The "Denominator" Challenge. *J Palliat Med*. 2018 Mar 1;21(S2):S-7-S-16.
- Hanson LC, Gabriel S, Ernecoff NC, Wessell KL, Mitchell SL, Sachs G, et al. Triggered Palliative Care in Advanced Dementia: A Pilot Study. 2017.
- Ernecoff NC, Wessell KL, Gabriel S, Carey TS, Hanson LC. A Novel Screening Method to Identify Late-Stage Dementia Patients for Palliative Care Research and Practice. *J Pain Symptom Manage*. 2018;
- Xu J, Rasmussen L V., Shaw PL, Jiang G, Kiefer RC, Mo H, et al. Review and evaluation of electronic health records-driven phenotype algorithm authoring tools for clinical and translational research. *J Am Med Informatics Assoc*. 2015 Jul 29;22(6):ocv070
- Conway M, Berg RL, Carrell D, Denny JC, Kho AN, Kullo IJ, et al. Analyzing the Heterogeneity and Complexity of Electronic Health Record Oriented Phenotyping Algorithms.
- Goebel JR, Ahluwalia SC, Chong K, Shreve ST, Goldzweig CL, Austin C, et al. Developing an Informatics Tool To Advance Supportive Care: The Veterans Health Care Administration Palliative Care National Clinical Template. *J Palliat Med*. 2014 Mar;17(3):266-73.
- Lilley EJ, Lindvall AC, Lillemoe KD, Tulskey JA, Wiener DC, Cooper Z, et al. Measuring Processes of Care in Palliative Surgery A Novel Approach Using Natural Language Processing. 2017;XX(Xx):1-3.
- I2b2. Informatics for Integrating Biology and Bedside. Boston, MA;
- Hanauer DA, Mei Q, Law J, Khanna R, Zheng K. Supporting information retrieval from electronic health records: A report of University of Michigan's nine-year experience in developing and using the Electronic Medical Record Search Engine (EMERSE). *J Biomed Inform*. 2015 Jun;55:290-300.
- Levey AS, Eckardt K-U, Tsukamoto Y, Levin A, Coresh J, Rossert J, et al. Definition and classification of chronic kidney disease: A position statement from Kidney Disease: Improving Global Outcomes (KDIGO). Vol. 67. *Kidney International*. 2005.
- Zeng QT, Goryachev S, Weiss S, Sordo M, Murphy SN, Lazarus R. BMC Medical Informatics and Decision Making Extracting principal diagnosis, co-morbidity and smoking status for asthma research: evaluation of a natural language processing system. 2006;
- Kelley AS, Morrison RS. Palliative Care for the Seriously Ill. *Campion EW*, editor. *N Engl J Med*. 2015 Aug 20;373(8):747-55.
- Aldridge MD, Hasselaar J, Garralda E, van der Eerden M, Stevenson D, McKendrick K, et al. Education, implementation, and policy barriers to greater integration of palliative care: A literature review. *Palliat Med*. 2016 Mar 24;30(3):224-39.
- Hanson LC, Collichio F, Bernard SA, Wood WA, Milowsky M, Burgess E, et al. Integrating Palliative and Oncology Care for Patients with Advanced Cancer: A Quality Improvement Intervention. *J Palliat Med*. 2017 Dec 1;20(12):1366-71.
- Shiekh Sroujeh L, Monroy D, Warren E. Using Electronic Health Record (EHR) Best Practice Alert (BPA) to Improve RBC Transfusion Practices and Adherence to the Guidelines. *Chest*. 2016 Oct 1;150(4):599A.
- American Academy of Family Physicians. A. Stoeckle J, Diaz V, Valko G, Arenson C. Family practice management. Vol. 24, *Family Practice Management*. American Academy of Family Physicians; 2017. 30-33 p.
- Lofthus DM, Gadgil JY, De Lemos JA. Overriding Concerns: The Role of Electronic Medical Record-Based Best Practice Alerts in Reducing Unnecessary Laboratory Testing. 2014;
- Kukhareva P, Staes C, Noonan KW, Mueller HL, Warner P, Shields DE, et al. Single-reviewer electronic phenotyping validation in operational settings: Comparison of strategies and recommendations. *J Biomed Inform*. 2017 Feb;66:1-10.
- Drawz PE, Archdeacon P, McDonald CJ, Powe NR, Smith KA, Norton J, et al. Special Feature CKD as a Model for Improving Chronic Disease Care through Electronic Health Records. *Clin J Am Soc Nephrol*. 2015;10:1488-99.
- Navar AM, Peterson ED, Steen DL, Wojdyla DM, Sanchez RJ, Khan I, et al. Evaluation of Mortality Data from the Social Security Administration Death Master File for Clinical Research. *JAMA Cardiol*. 2019;
- Adelson K, Paris J, Horton JR, Hernandez-Tellez L, Ricks D, Morrison RS, et al. Standardized Criteria for Palliative Care Consultation on a Solid Tumor Oncology Service Reduces Downstream Health Care Use. *J Oncol Pract*. 2017 May 17;13(5):e431-40.
- Weissman DE, Meier DE. Special Report Identifying Patients in Need of a Palliative Care Assessment in the Hospital Setting A Consensus Report from the Center to Advance Palliative Care.
- Avati A, Jung K, Harman S, Downing L, Ng A, Shah NH. Improving palliative care with deep learning. In: 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE; 2017. p. 311-6.
- Sahni N, Simon G, Arora R. Development and Validation of Machine Learning Models for Prediction of 1-Year Mortality Utilizing Electronic Medical Record Data Available at the End of Hospitalization in Multicondition Patients: a Proof-of-Concept Study. *J Gen Intern Med*. 2018;33(6):921-9.
- Cardona-Morrell M, Hillman K. Development of a tool for defining and identifying the dying patient in hospital: Criteria for Screening and Triaging to Appropriate alternative care (CriSTAL).

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.