

CVD than those with a protein score below the median score. DISCUSSION/SIGNIFICANCE OF FINDINGS: A protein score developed improved discrimination of PWH with CVD and those without, and helped identify PWH with high risk for developing CVD. If validated, this score and/or the individual proteins could be used in addition with established factors to identify CVD at-risk individuals who might benefit from aggressive risk-reduction.

Team Science

10227

A Framework for Bringing Secondary Analysis of EHR Data to Geographically Dispersed Clinician Scientists

James McClay, Jerrod Anzalone, Carol Geary and Ying Zhang
University of Nebraska Medical Center

ABSTRACT IMPACT: The described framework will enable other sites with a well-defined apparatus for enabling the secondary analysis of EHR data for research through education, team science, and resource consolidation. OBJECTIVES/GOALS: EHR's potential to improve healthcare outcomes extends far beyond the clinic. This vast repository of clinical insights has dramatic potential for biomedical research. To enhance accessibility for busy clinicians and underserved populations, we describe a framework for interfacing with EHR locally and through national network participation. METHODS/STUDY POPULATION: The Institutional Development Award (IDeA) program, which began in 1993, broadens NIH funding's geographic distribution for biomedical research. Included in this is the IDeA Networks for Clinical and Translational Research, which focuses on enhancing clinical and translational science across a network of IDeA-states with traditionally underserved communities and rural providers. A prior survey of the needs and capabilities of IDeA-CTR centers identified the need for improved research support. Based on our annual member survey we developed a process for supporting distributed research projects across the GP-CTR. NIH also recently made a funding announcement for the IDeA-CTR community identifying EHR research as a major priority in responding to the COVID-19 pandemic. RESULTS/ANTICIPATED RESULTS: Results from site interviews and member surveys show a clear need for dedicated resources to navigate the process of EHR-derived research. Most described a different set of requirements for increasing accessibility to EHR for research and a strong desire to participate in research networks. Local investigators cited a lack of tools, educational materials, and accessibility. Initial efforts demonstrate strong research questions but limited technical, statistical, and terminological capabilities to succeed. In response, a pipeline for team science and promotion of projects from local phenotypes to national studies. We created a facilitator training program to expand the number of facilitators (n=22), quarterly training for investigators (n=104), and ongoing efforts to advance COVID-19 research. DISCUSSION/SIGNIFICANCE OF FINDINGS: As evidenced in the expanding number of EHR-based research networks there is a need for a system to promote project development and best practices. The proposed model promotes education, resource sharing, and team formation to advance clinical questions from the idea stage toward national research network participation.

74123

A Learning Health Systems approach using health record data to construct patient frailty scores and predict safety events

Alex Bokov, Sara Espinoza, Chandana Tripathy and Kathleen Stevens
UT Health San Antonio

ABSTRACT IMPACT: Laying the groundwork for better predictive algorithms to inform clinical decisions and planning. OBJECTIVES/GOALS: Frailty scores predict poor patient outcomes. Validated against highly relevant outcomes, such scores can be used to inform clinical and resource utilization decisions. We generated and validated an electronic Frailty Index (EFI) from real-world EHR data using the Rockwood deficit-accumulation framework to predict patient safety events. METHODS/STUDY POPULATION: To assure that the research approach reflected perspectives of multiple stakeholders, our multidisciplinary group included an implementation scientist, a geriatrician, an internist, and an informatician. From our large academic health center, we accessed EHR data for 14,844 patients randomly sampled from the data warehouse underlying our ACT/SHRINE node. The per-visit EFI scores were calculated using EHR codes in a rolling 2-year time window. EFI was used as the predictor variable in the analytic design. The primary outcomes were preventable patient-safety events derived from ICD-10 codes including hospital-acquired infections, non-operative hospital-acquired trauma, and cardiac complications. Cox proportional hazard models were used to estimate risk for each outcome. RESULTS/ANTICIPATED RESULTS: We found statistically significant associations of EFI with clinically meaningful outcomes from EHR data. For most outcomes, we found significant correlation with EFI and c-statistics indicating good calibration of the models. The EFI was a strong predictor of clinically relevant outcomes without relying on any data other than diagnoses, vital signs, and laboratory results from the EHR. In contrast to previous studies, we treated EFI as a time-varying predictor with multiple follow-ups per patient, which is more realistic than relying on one static time-point. We used a representative sample of the adult patient population rather than limiting it to older individuals and found EFI to be a useful metric even at relatively young ages. DISCUSSION/SIGNIFICANCE OF FINDINGS: The EFI predicted safety events in adult patients using only routine, structured EHR data and can offer a low-effort, scalable method of risk assessment, valuable to clinical decisions. The capability to harness EHR data and rapidly generate clinical knowledge can be transformative for complex care and contributes to Learning Health Systems.

Translational Science, Policy, & Health Outcomes Science

21063

A Review of Novel Uses of REDCap in Clinical and Translational Science

Aaryn Toles, MS¹, Barbara Tafuto, MLS, PhD² and Doreen Lechner, PhD²

¹Rutgers Robert Wood Johnson Medical School and ²Rutgers School of Health Professions

ABSTRACT IMPACT: This review will encourage further development of novel uses of REDCap for the benefit of the research

community. OBJECTIVES/GOALS: REDCap is a clinical research data collection platform that is primarily used as intended. However, little is known about its more novel uses, specifically in clinical decision support in patient care and in clinical research management. Thus, the purpose of this review is to examine peer reviewed literature identifying and describing such novel uses. METHODS/STUDY POPULATION: A systematic search was conducted in both PubMed and Google Scholar using the equation ((REDCap) OR ('Research Electronic Data Capture')) AND ((Clinical Trial Management) OR (Clinical Research)). Articles were screened by title, then abstract, and then were reviewed in full if they met inclusion criteria. Articles were included if they had potential relevance to the topic of REDCap or if they mentioned activities related to fields of clinical and translational science including operational support in areas such as clinical research management. Articles were excluded if they focused on common clinical research activities relating to data collection software such as survey administration, database building or data collection for clinical trials, registries, and cohort studies. RESULTS/ANTICIPATED RESULTS: The initial search yielded 390 results, of which 40 underwent an abstract review; only 8 of these underwent full text review. Of these, 5 discussed uses of REDCap in the context of operational support in clinical research management; 3 were related to clinical decision support in patient care. For the 5 articles focused on operational support in clinical research management, topics include e-consenting procedures, collection and storage of protected health information (PHI), patient recruitment and tracking stakeholder engagement. The 3 articles about clinical decision support discuss REDCap tools for generating risk predictions for post-surgical clinical outcomes, generating recommendations and STI test orders, and increasing efficiency in hand-offs to enhance care of surgical oncology patients. DISCUSSION/SIGNIFICANCE OF FINDINGS: Considering that only a small percentage of peer reviewed research reports out on novel uses of REDCap, there is a need for the REDCap consortium to do further work to fulfill its mission to adopt, innovate, and suggest novel uses of REDCap, thus expanding the understanding of its functionalities and therefore its utility in the research community.

41250

Machine Learning to Identify Predictors of Iatrogenic Injury Using Empirical Bayes Estimates: A Cohort Study of Pressure Injury Prevention

William V. Padula¹, David G. Armstrong and Patricia M. Davidson
¹University of Southern California, ²Keck Medicine of USC and
³Johns Hopkins School of Nursing

ABSTRACT IMPACT: A machine learning approach using electronic health records can combine descriptive, population-level factors of pressure injury outcomes. OBJECTIVES/GOALS: Pressure injuries cause 60,000 deaths and cost \$26 billion annually in the US, but prevention is laborious. We used clinical data to develop a machine learning algorithm for predicting pressure injury risk and prescribe the timing of intervention to help clinicians balance competing priorities. METHODS/STUDY POPULATION: We obtained 94,745 electronic health records with 7,000 predictors to calibrate a predictive algorithm of pressure injury risk. Machine learning was used to mine features predicting changes in pressure injury risk; random forests outperformed neural networks, boosting and bagging in feature selection. These features were fit to multilevel ordered logistic regression to create an algorithm that generated empirical Bayes estimates informing a decision-rule for follow-up based on individual risk trajectories over time. We

used cross-validation to verify predictive validity, and constrained optimization to select a best-fit algorithm that reduced the time required to trigger patient follow-up. RESULTS/ANTICIPATED RESULTS: The algorithm significantly improved prediction of pressure injury risk ($p < 0.001$) with an area under the ROC curve of 0.60 compared to the Braden Scale, a traditional clinician instrument of pressure injury risk. At a specificity of 0.50, the model achieved a sensitivity of 0.63 within 2.5 patient-days. Machine learning identified categorical increases in risk when patients were prescribed vasopressors ($OR = 16.4$, $p < 0.001$), beta-blockers ($OR = 4.8$, $p < 0.001$), erythropoietin stimulating agents ($OR = 3.0$, $p < 0.001$), or were ordered a urinalysis screen ($OR = 9.1$, $p < 0.001$), lipid panel ($OR = 5.7$, $p < 0.001$) or pre-albumin panel ($OR = 2.0$, $p < 0.001$). DISCUSSION/SIGNIFICANCE OF FINDINGS: This algorithm could help hospitals conserve resources within a critical period of patient vulnerability for pressure injury not reimbursed by Medicare. Savings generated by this approach could justify investment in machine learning to develop electronic warning systems for many iatrogenic injuries.

67409

Quantifying Unmeasured Confounding in Relationship between Treatment Intensity and Outcomes among Older Patients with Hodgkin Lymphoma (HL) using Surveillance, Epidemiology and End Results (SEER)-Medicare Data

Angie Mae Rodday¹, Theresa Hahn², Peter K. Lindenauer² and Susan K. Parsons¹

¹Tufts Medical Center, ²Roswell Park Comprehensive Cancer Center and ³University of Massachusetts Medical School - Baystate

ABSTRACT IMPACT: E-values can help quantify the amount of unmeasured confounding necessary to fully explain away a relationship between treatment and outcomes in observational data. OBJECTIVES/GOALS: Older patients with HL have worse outcomes than younger patients, which may reflect treatment choice (e.g., fewer chemotherapy cycles). We studied the relationship between treatment intensity and 3-year overall survival (OS) in SEER-Medicare. We calculated an E-value to quantify the unmeasured confounding needed to explain away any relationship. METHODS/STUDY POPULATION: This retrospective cohort study of SEER-Medicare data from 1999-2016 included 1131 patients diagnosed with advanced stage HL at age ≥ 65 years. Treatment was categorized as: (1) full chemotherapy regimens ('full regimen', $n = 689$); (2) partial chemotherapy regimen ('partial regimen', $n = 175$); (3) single chemotherapy agent or radiotherapy ('single agent/RT', $n = 102$), or (4) no treatment ($n = 165$). A multivariable Cox regression model estimated the relationship between treatment and 3-year OS, adjusting for disease and patient factors. An E-value was computed to quantify the minimum strength of association that an unmeasured confounder would need to have with both the treatment and OS to completely explain away a significant association between treatment and OS based on the multivariable model. RESULTS/ANTICIPATED RESULTS: Results from the multivariable model found higher hazards of death for partial regimens ($HR = 1.81$, 95% $CI = 1.43, 2.29$), single agent/RT ($HR = 1.74$, 95% $CI = 1.30, 2.34$), or no treatment ($HR = 1.98$, 95% $CI = 1.56, 2.552$) compared to full regimens. We calculated an E-value for single agent/RT because it has the smallest HR of the treatment levels. The observed HR of 1.74 could be explained away by an unmeasured confounder that was associated with both treatment and OS with a HR of 2.29, above and beyond the measured confounders; the 95% CI could be moved to include the null by an unmeasured confounder that was associated with both the treatment and OS with a HR of 1.69. Of the