DISCUSSION/SIGNIFICANCE: Given the implementation of DPP at-scale there is an urgent need to understand the patient and systems-level factors that are associated with referring individuals in the DPP. By detecting characteristics of health systems and patients that warrant special attention, we can improve equitable access to evidence-based diabetes prevention.

56

Using machine learning to predict 30-day readmission and reoperation following resection of supratentorial high-grade gliomas: A national analysis of 9,418 patients.

Abdul Karim Ghaith[1], Marc Ghanem[3], Cameron Zamanian[1,2], Antonio A. Bon Nieves[1,2], Archis Bhandarkar[1,2], Karim Nathani[1,2], Mohamad Bydon[1,2], Alfredo Quiones-Hinojosa[4]
[1]Mayo Clinic Neuro-Informatics Laboratory, Mayo Clinic, Rochester, MN, USA [2]Department of Neurological Surgery, Mayo Clinic, Rochester, MN, USA [3]Gilbert and Rose-Marie Chagoury School of Medicine, Lebanese American University [4]Department of Neurological Surgery, Mayo Clinic, Jacksonville, FL, USA

OBJECTIVES/GOALS: High-grade gliomas (HGG) are among the rarest, most aggressive tumors in neurosurgical practice. We aimed to identify the clinical predictors for 30-day readmission and reoperation following HGGs surgery using the NSQIP database and seek to create web-based applications predicting each outcome. METHODS/STUDY POPULATION: We conducted a retrospective, multicenter cohort analysis of patients who underwent resection of supratentorial HGG between January 1, 2016, and December 31, 2020, using the NSQIP database. Demographics and comorbidities were extracted. The primary outcomes were 30-day unplanned readmission and reoperation. A stratified 80:20 split of the available data was carried out. Supervised machine learning algorithms were trained to predict 30-day outcomes. RESULTS/ANTICIPATED RESULTS: A total of 9,418 patients were included in our cohort. The rate of unplanned readmission within 30 days of surgery was 14.9%.Weight, chronic steroid use, pre-operative BUN, and WBC count were associated with a higher risk of readmission. The rate of early unplanned reoperation was 5.47%. Increased weight, higher operative time, and a longer period between hospital admission and the operation were linked to increased risk of early reoperation. Our Random Forest algorithm showed the highest predictive performance for early readmission (AUC = 0.967), while the XG Boost algorithm showed the highest predictive performance for early reoperation (AUC = 0.985).Web-based tools for both outcomes were deployed: https://glioma-readmission.herokuapp.com/, https://glioma-reoperation.herokuapp.com/. DISCUSSION/SIGNIFICANCE: A high fraction of documented early unplanned readmission and reoperation were considered preventable and related to surgery. Machine learning allows better prediction of resected HGG's prognosis based on findings from baseline methods leading to more personalized patient care.

57

Utility of Digital Phenotyping in Big Data to Answer Clinical Questions: Puberty as a Transdisciplinary Science Case Example

David J. Schnabel, Jr., Lorah D. Dorn
Pennsylvania State University

OBJECTIVES/GOALS: A disease-agnostic translational science framework for data mining is proposed for use across disciplines to: Answer clinical questions, justify future clinical research recruitment, and explore under-represented populations. As a case example, male puberty demonstrates utility of the framework. METHODS/STUDY POPULATION: As a case example using the generalizable framework, the following interdisciplinary question was asked: Does early pubertal timing increase the risk of developing type II diabetes (T2d) in boys? A digital phenotype of males < 18 years old was created in the TriNetX Diamond Network utilizing Boolean operator data queries. TriNetX contains patient electronic health record information (ICD-10 diagnoses, anthropometric measures). A case control analysis leveraging patient counts from various digital phenotypes allowed for outcome (T2d) comparison of boys diagnosed with precocious puberty (E30.1, ICD code for early pubertal timing) to those without, while controlling for body mass index (BMI). RESULTS/ANTICIPATED RESULTS: Subjects (N=12,996,132) displayed the following digital phenotype: Male, < 18 years old, without ever having a BMI documented >85th percentile. Boys diagnosed with precocious puberty (E30.1) were 6.89 times more likely to develop T2d when aged 14-18 years old than those without (OR 6.89, 95%CI: 5.17-9.19, p. DISCUSSION/SIGNIFICANCE: Boys are under-represented in the early pubertal timing literature, justifying future human subjects research on male puberty. This case example demonstrates a broader disease-agnostic framework which can be adapted across disciplines. Opportunities may include public health digital phenotyping.

58

Utilizing VA Data to Define Long COVID and Identify Patients at Risk

Peter L. Elkin[1,2,4,5], Skyler Resendez[1,2], H. Sebastian Ruiz[1], Wilmon McCray[1,2], Steven H. Brown[3], Jonathan Nebeker[3], Diane Montella[3]
[1]Department of Biomedical Informatics, Jacobs School of Medicine and Biomedical Sciences, University at Buffalo [2]Department of Veterans Affairs, VA Western New York Healthcare System and VA Research Service 3Office of Health Informatics, Department of Veterans Affairs 4Department of Internal Medicine, Jacobs School of Medicine and Biomedical Sciences, University at Buffalo 5Faculty of Engineering, University of Southern Denmark

OBJECTIVES/GOALS: To determine the signs, symptoms, and diagnoses that are significantly upregulated in cases of long COVID while identifying risk factors and demographics that increase one's likelihood of developing long COVID. METHODS/STUDY POPULATION: This is a retrospective, big data science study. Data from Veterans Affairs (VA) medical centers across the United States between the start of 2020 and the end of 2022 were utilized. Our cohort consists of 316,782 individuals with positive COVID-19 tests recorded in the VA EHR with a history of ICD10-CM diagnosis codes in the record for case-control comparison. We looked at all new diagnoses that were not present in the six months before COVID diagnosis but were present in the time period from one month after COVID through seven months after. We determined which were significantly enriched and calculated odds ratios for each, organized by long COVID subtypes by medical specialty / affected organ system. Demographic analyses were also performed for long COVID patients and patients without any new long COVID ICD10-CM codes. RESULTS/ANTICIPATED RESULTS: This profile shows disorders that are highly upregulated in the post-COVID population and provides strong evidence for a broad definition of long COVID. By breaking this into subtypes by medical specialty, we define cardiac long COVID, neurological long COVID, pulmonary long COVID, and eight others. The long COVID cohort

https://doi.org/10.1017/cts.2023.145 Published online by Cambridge University Press