the homeless individuals from HMSI using OpenEMPI software package, which is an open source implementation of an Enterprise Master Patient Index (EMPI). An entity model was generated based on the selective data elements from HMIS database, which were relevant for the patient identity management and healthcare service management. An automated script was implemented to extract data from HMIS and load it into OpenEMPI to build the MPI. Once the MPI is setup, the Emergency Department users were able to perform patient identity matching and confirm housing insecure or homeless status of their patients by querying the index using the web-based tool. We developed structured data elements to record homelessness information, which will allow us to measure the prevalence of this risk among patients. We are also exploring the possibility to integrate the systems the using the IHE PIX/PDQ profile, which provides ways for healthcare applications to query a patient information server for a patient based on user-defined search criteria, and retrieve a patient's information directly into the application. RESULTS/ANTICIPATED RESULTS: We implemented a MPI of homeless individuals, which would allow the emergency department users to perform patient identity matching of housing insecure or homeless patients, without undue privacy intrusions. We are confident that IHE PIX/PDQ profile is able to support the integration of healthcare and housing and homeless services systems and enable the data sharing in an efficient way. DISCUSSION/SIGNIFICANCE OF IMPACT: The project addressed the gap in the sharing of data about housing insecure or homeless persons between healthcare and housing and social services that will result in improvements in coordination of care, reduce the cycle time from recognition of risk to the referral to housing and services and improve health outcomes and residential stability. Successful completion of this integration project will give us a model that we can scale to many other communities.

CLINICAL EPIDEMIOLOGY

2027

Racial differences in leukemia prognosis: New epidemiologic analysis

Shuangge Ma, Yinjun Zhao and Yu Wang
Yale School of Medicine, New Haven, CT, USA

OBJECTIVES/SPECIFIC AIMS: Research on cancer difference is of significant scientific and practical value. For leukemia, the survival disadvantage of the Blacks has been suggested in multiple studies. However, the existing epidemiologic analysis has multiple technical limitations. The goal of this study is to more accurately quantify so as to better understand different sources of racial differences in leukemia survival. METHODS/STUDY POPULATION: A new statistical method, which is based on robust regression and resampling, is developed. Data are obtained from the SEER (Surveillance, Epidemiology, and End Results) database. Using the "classic" epidemiologic methods as well as the new method, analysis is conducted on the prognosis of 4 leukemia subtypes (ALL, CLL, AML, and CML) for 4 major racial groups (White, non-Hispanic White, Black, and Asian and Pacific Islander). RESULTS/ANTICIPATED RESULTS: After effectively removing differences caused by the observed clinicopathological and demographic factors, the survival disadvantage of the Blacks persists for the following patient groups: ALL and $age > 14$, CLL and $age > 14$, and ALL and $age \leq 14$. The quantitative results are significantly different from those from classic epidemiologic analysis. Such observed racial differences are more attributable to the unobserved risk factors and cancer disparity. DISCUSSION/SIGNIFICANCE OF IMPACT: This study provides a more effective and more direct quantification of racial difference in leukemia prognosis. The survival disadvantage of the Blacks which is observed for certain subtypes/age groups deserves further attention but should not be overstated. More data collection and analysis are needed to more accurately decipher racial differences in leukemia and other cancer types.

2039

Modifiable risk factors Versus age on developing high predicted cardiovascular disease risk in African Americans

Adam Bress, Lisandro D. Colantonio, John N. Booth, Tanya M. Spruill, Joseph Ravenell, Mark Butler, Amanda J. Shallcross, Samantha R. Seals, Kristi Reynolds, Gbenga Ogedegbe, Daichi Shimbo and Paul Muntner
Department of Population Health Sciences, University of Utah, Salt Lake City, UT, USA

OBJECTIVES/SPECIFIC AIMS: Clinical guidelines recommend using predicted atherosclerotic cardiovascular disease (ASCVD) risk to inform treatment decisions. The objective was to compare the contribution of changes in modifiable risk factors Versus aging to the development of high 10-year predicted ASCVD risk. METHODS/STUDY POPULATION: Prospective follow-up of the Jackson Heart Study, an exclusively African-American cohort, at visit 1 (2000–2004) and visit 3 (2009–2012). Analyses included 1115 African-American participants without a high 10-year predicted ASCVD risk (<7.5%), hypertension, diabetes, or ASCVD at visit 1. We used the Pooled Cohort equations to calculate the incidence of high (≥ 7.5%) 10-year predicted ASCVD risk at visit 3. We recalculated the percentage with a high 10-year predicted ASCVD risk at visit 3 assuming each risk factor [age, systolic blood pressure (SBP), antihypertensive medication use, diabetes, smoking, total and high-density lipoprotein cholesterol], one at a time, did not change from visit 1. RESULTS/ANTICIPATED RESULTS: The mean age at visit 1 was 45.2 ± 9.5 years. Overall, 30.9% (95% CI 28.3%–33.4%) of participants developed high 10-year predicted ASCVD risk. Aging accounted for 59.7% (95% CI 54.2%–65.1%) of the development of high 10-year predicted ASCVD risk compared with 32.8% (95% CI 27.0%–38.2%) for increases in SBP or antihypertensive medication initiation and 12.8% (95% CI 9.6%–16.5%) for incident diabetes. Among participants <50 years, the contribution of increases in SBP or antihypertensive medication initiation was similar to aging. DISCUSSION/SIGNIFICANCE OF IMPACT: Increases in SBP and antihypertensive medication initiation are major contributors to the development of high 10-year predicted ASCVD risk in African Americans, particularly among younger adults.

2041

Before hospice: Symptom burden, dementia, and social participation in the final years

Halima Amjad, David Roth, Jennifer Wolff, Esther Oh and Quincy Samus
Johns Hopkins University, Baltimore, MD, USA

OBJECTIVES/SPECIFIC AIMS: Traditional hospice focuses on symptoms and quality of life (QOL) at the very end of life. Clinical symptoms and QOL in the last 1–2 years of life are also important and may be affected by dementia. Our objective was to characterize how symptoms differ between people with and without dementia in the last years before death and whether symptoms impact social dimensions of QOL. METHODS/STUDY POPULATION: We studied 1270 community-dwelling participants who died between 2011 and 2015 in the National Health and Aging Trends Study, a nationally representative cohort of older adults. From the last interview before death, we examined sensory (vision; hearing), physical (pain; problems with breathing, chewing/swallowing, speaking, upper or lower extremity strength/movement, and balance/coordination), and psychiatric (depression; anxiety; insomnia) symptoms by dementia status. We examined associations between symptoms and participation restrictions (visiting family/friends, attending religious services, participating in clubs/activities, going out for enjoyment, and engaging in favorite activity). RESULTS/ANTICIPATED RESULTS: Low energy (69%), pain (59%), and lower extremity strength/movement problems (56%) were most common. People with dementia (37.3% of decedents) had higher prevalence of all symptoms ($p \leq 0.01$), except pain, breathing problems, and insomnia. Dementia and greater symptom burden were independently associated with greater odds of participation restrictions ($p < 0.05$). Problems speaking were significantly associated with limitations in all activities except for attending religious services. Balance/coordination, energy, and strength/movement problems were associated with limitations in 3 activities. DISCUSSION/SIGNIFICANCE OF IMPACT: Sensory, physical, and psychiatric symptoms are common in the year before death, with greater symptom prevalence in people with dementia. Both dementia and symptoms are associated with restrictions in participation. Older patients may benefit not only from earlier emphasis on palliative care but also programs and assistive devices that accommodate physical impairments.

2091

Evaluating the validity and utility of surrogate endpoints in clinical trials of chronic kidney disease (CKD)

Jian Ying, Andrew Redd and Tom Greene
The University of Utah School of Medicine, Salt Lake City, UT, USA

OBJECTIVES/SPECIFIC AIMS: The objective of this research is to determine under what conditions endpoints based on estimated glomerular filtration rate (eGFR) slope or on relatively small declines in eGFR provide valid and useful surrogate endpoints for pivotal clinical trials in chronic kidney disease (CKD) patients. METHODS/STUDY POPULATION: We consider 2 classes of surrogate endpoints. The first class includes endpoints defined by the average

rate of change in eGFR during defined portions of the follow-up period of the trial, following initiation of the randomized treatment interventions. The second class includes composite endpoints defined by the time from randomization until the occurrence of a designated decline in eGFR or kidney failure. The true clinical endpoint is considered to be the time from randomization until kidney failure, irrespective of the trajectory in eGFR measurements prior to kidney failure. We apply statistical simulation to determine conditions under which alternative endpoints within the 2 classes are (1) valid surrogate endpoints, in the sense of preserving a low probability of rejecting the null hypothesis of no treatment effect on the surrogate endpoint when there is no treatment effect on the clinical endpoints and are also (2) useful surrogate endpoints, in the sense of providing increased statistical power that allows significant reductions in sample size and/or duration of follow-up. Input parameters for the simulations include (a) characteristics of the joint distribution of the longitudinal eGFR measurements and the time to occurrence of renal failure, (b) characteristics of the short-term and long-term effects of the treatment, and (c) design parameters, including the duration of accrual and follow-up and the spacing of eGFR measurements during the follow-up period. We use joint analyses of 19 treatment comparisons across 13 previous clinical trials of CKD patients to guide the selection of input parameters for the simulations. We apply longitudinal mixed effects models for analysis of endpoints based on eGFR slope, and Cox regression for analyses of the composite time-to-event endpoints. RESULTS/ANTICIPATED RESULTS: We have previously shown that surrogate endpoints defined by eGFR declines of 30% or 40% can provide valid and useful alternative endpoints in CKD clinical trials for interventions that do not produce short-term effects on eGFR which differ from the longer-term effects of the interventions. Other factors influencing the validity and utility of these endpoints include the average baseline eGFR, the mean rate of change in eGFR, and the extent to which the size of the treatment effect depends on the patient's underling rate of eGFR decline. We will extend these results by presenting preliminary results describing conditions under which outcomes based on eGFR slope provide valid and useful alternatives to the clinical endpoint of time until occurrence of kidney failure. DISCUSSION/SIGNIFI-CANCE OF IMPACT: The statistical simulation strategy described in this research can be used during the design of clinical trials of chronic kidney disease to assist in the selection of endpoints that maximize savings in sample size and duration of follow-up while retaining a low risk of producing a false positive conclusion in the absence of a true effect of the treatment on the time until kidney failure.

2093

Utilization of an ICD-coded electronic health records (EHR) database to characterize the epidemiology of prosopagnosia

Christina Pressl, Caroline Jiang, Joel Correa da Rosa, Maximilian Friedrich, Winrich Freiwald and Jonathan Tobin

OBJECTIVES/SPECIFIC AIMS: We aim to examine the epidemiological characteristics of prosopagnosia by querying and analyzing a large deidentified clinical data set from 12 New York City-based hospitals and Federally Qualified Health Centers (FQHCs). The PCORI-funded New York City Clinical Data Research Network (NYC-CDRN) contains ~4.5 million deidentified ICD-coded electronic health records (EHRs) with comprehensive longitudinal information on demographics, patient visits, clinical conditions/diagnoses, laboratory and radiology results, medications, and clinical procedures. The NYC-CDRN will be expanded to include other data sources, including insurance claims, social determinant of health, patient reported outcomes, and patient generated data. The central hypothesis was that systematic mining of this database would reveal new epidemiological information about prosopagnosia. We developed a computable phenotype for prosopagnosia, using the International Classification of Diseases version 9 (ICD-9). The computable phenotype consisted of the diagnostic code for the condition under study, prosopagnosia (ICD-9 code 368.16), as well as the codes for known surrogate diagnoses. We expected to identify cases of acquired prosopagnosia, where the condition occurs only after brain damage, due to stroke, trauma, or meningitis for example, and cases of developmental prosopagnosia, where the condition is present from an early age, with no history of brain damage. The goals of this project were to provide new information about the condition's prevalence rate in the New York City area, which could be furthermore translated into wider geographical areas and to yield novel details about its antecedents and comorbid conditions. METHODS/STUDY POPULATION: To determine the presence of the diagnosis of interest, prosopagnosia, and common co-occurring conditions among a New York City-based study population, we investigated a large database in collaboration with the NYC-CDRN. At the time the large database was mined it contained ~4 million ICD-9 coded EHRs. We first created a search paradigm; applicable for screening the database that consists of ICD-9 coded

EHRs. We generated a list of ICD-9 codes indicative for the patients' difficulties with the perception of faces (368.16), which indicates the presence of the condition as part of the psychophysical visual disturbances complex, and this code identified 871 patients. Furthermore, we collected codes that indicate the presence of conditions that are known to be surrogate diagnoses of prosopagnosia. ICD-9 codes for surrogate diagnoses included for example, 854.* (coding for personal history of traumatic brain injury, n = 1,409), 434.01, 434.11, and 434.91 (coding for cerebral thrombosis, embolus and artery occlusion unspecified with cerebral infarction, n = 19,409), and 191.2 (coding for malignant neoplasm of the temporal lobe, n = 566). In October 2015, coding was changed to the new ICD-10 coding system. No additional patients were revealed from the data set when the cohort was searched for the presence of corresponding ICD-10 codes, as institutions are currently in transition from ICD-9 to ICD-10. Using this search query with the large database, we extracted novel information about the epidemiological and demographical distribution of prosopagnosia and furthermore, gained new knowledge about commonly associated diseases. The fact that it must be presumed that the majority of diagnoses of prosopagnosia have been made on the basis of patients' self-reports and clinicians' judgments represents a limiting factor in this study. We are currently exploring machine-learning strategies to identify potential false-negative cases among the patients with surrogate diagnoses. RESULTS/ANTICIPATED RESULTS: Investigations and application of our search query revealed a total number of n = 129,549 patients carrying either the diagnosis code for prosopagnosia or the codes for the known surrogate diagnoses. There were 871 patients who carried the ICD-9 code 368.16, indicating the potential presence of prosopagnosia among other visual disturbances. Remaining patients (n = 128,678) carried codes for known surrogate diagnoses, contained in the search query. Statistical analyses revealed elevated odds ratios for men (OR = 1.55, 95% CI: 1.36, 1.77, $p < 0.0001$), and for Black/African Americans Versus White individuals (OR = 2.09, 95% CI: 1.74, 2.51, $p < 0.0001$). DISCUSSION/SIGNIFICANCE OF IMPACT: Currently, the prevalence of prosopagnosia remains unknown. Face blind individuals are struggling to recognize their social contacts by their face only in every day life and are therefore prone to experience reduced quality of life. We searched the large NYC-based clinical database, containing more than 4.5 million deidentified ICD-coded health records, for cases of prosopagnosia to shed light into its prevalence and epidemiological characteristics. We furthermore, mined the database for cases carrying known surrogate diagnoses to explore the magnitude and characteristics of individuals potentially under increased risk. Our efforts address a great healthcare need, as they revealed new epidemiological knowledge of a vulnerable and understudied population. The results of this project reveal new insights into the epidemiological character-istics of prosopagnosia and its surrogate diagnoses, and demonstrate the feasibility of mining large clinical databases to identify rare clinical populations. Our results suggest the need for a more targeted diagnostic assessment of face perception abilities in populations under increased risk.

2111

Insulin resistance patterns over 25-years of adulthood and nonalcoholic fatty liver disease in middle age: The Coronary Artery Risk Development in Young Adults (CARDIA) Study

Lisa B. VanWagner, Michael Bancks, Hongyan Ning, Juned Siddique, Cora Lewis, John Jeffrey Carr, Miriam Vos, Elizabeth Speliotes, Norah Terrault, Mary E. Rinella, Norrina B. Allen and Donald Lloyd-Jones
Northwestern University, Evanston, IL, USA

OBJECTIVES/SPECIFIC AIMS: Nonalcoholic fatty liver disease (NAFLD) is the most common cause of liver disease in the United States and increases risk for cirrhosis and liver cancer. Identifying modifiable risk factors for NAFLD could allow better targeting of prevention programs. Insulin resistance (IR) plays a significant role in the development and progression of NAFLD. IR is also an important precursor to the development of type 2 diabetes (T2DM). However, the development and duration of IR during young adulthood and its association with NAFLD and T2DM in midlife is unclear. To test whether trajectories of IR using homeostatic model assessment (HOMA-IR) change throughout early adulthood are associated with risk of prevalent NAFLD and T2DM among persons with NAFLD in midlife independent of current or baseline HOMA-IR. METHODS/STUDY POPULATION: Participants from the CARDIA study, a prospective multicenter population-based biracial cohort of adults (baseline age 18–30 years), underwent HOMA-IR measurement (≥8 h fasting and not pregnant) at baseline (1985–1986) and follow-up exam years 7, 10, 15, 20, and 25. At Year 25 (Y25, 2010–2011), liver fat was assessed by noncontrast computed tomography (CT). NAFLD was defined as CT liver attenuation <51