electronic data warehouse (EDW) and then matched to records in CHND. With severity of HIE, gender, and confirmed seizures, each marker's association with LOS was calculated using multivariable Cox proportional hazards regression equations. These analyses were stratified by mortality. Candidate markers were vital signs, pulse oximetry, creatinine, acidosis (pH), international normalized ratio (INR), and supplemental oxygen (FiO_2). RESULTS/ANTICIPATED RESULTS: There were 66 eligible infants (38 males) and 1741 patient-days identified; Severe HIE (48%) and mortality (n = 21, 32%) were common. Overall, the median length of stay (mLOS) was 20.5 days (25th–75th centile: 10–31 days), although shorter for nonsurvivors [nonsurvivors mLOS = 8 days (5, 20); survivors mLOS = 24 days (14, 31), $p < 0.001$]. Median birthweight and gestational age were 3.3 kg and 39.4 weeks' gestation, respectively. In survivors (n = 45, 1290 days), regression analyses demonstrated that none of the selected parameters were associated with LOS. Among nonsurvivors (n = 21, 451 days), diastolic blood pressure changes [hazard ratio (HR) = 0.93, 95% confidence interval (CI) = 0.88, 0.97, $p = 0.04$] was related to longer time of survival; conversely, temperature (HR = 2.0, 95% CI = 1.24, 3.26, $p = 0.005$) was related to shorter survival. Creatinine, pH, INR, FiO_2, or other vital signs were unrelated to time-to-death in nonsurvivors. DISCUSSION/SIGNIFICANCE OF IMPACT: In a pilot study of neonatal HIE, changes in physiologic values were related to duration of survival in nonsurvivors, while neither physiologic nor laboratory values were related to survivors' mLOS. These results both exemplify novel uses for disease-specific, exposure-outcome relationships using EDWs and incorporates required functionalities of required software patches to extract, clean, and report from clinical information captured in electronic health records. We anticipate that text mining with techniques such as natural language processing will augment associations and/or predictions of short-term outcomes.

High-throughput phenotyping and the increased risk of OSA in Rosacia patients

Peter Elkin, Sarah Mullin, Sanjay Sethi, Shyamashree Sinha and Animesh Sinha
University at Buffalo, State University of New York, Buffalo, NY, USA

OBJECTIVES/SPECIFIC AIMS: To create a new semantically correct high-throughput phenotyping (HTP) platform. To demonstrate the utility of the HTP platform for observational research and can allow clinical investigators to perform studies in 5 minutes. To demonstrate the improved accuracy of observational research using this platform when compared with traditional observational research methods. To demonstrate that patients who have Roseacea are at increased risk of having obstructive sleep apnea (OSA). METHODS/STUDY POPULATION: This population is a set of 212,343 patients in the outpatient setting cared for in the Buffalo area over a 6-year period. All records for these patients were included in the study. Structured data was imported into an OMOP (OHDSI) database and all of the notes and reports were parsed by our HTP system which produces SNOMED CT codes. Each code is designated as a positive, negative or uncertain assertion and compositional expressions are automatically generated. We store the codified data 750,000,000 codes in Berkley DB, a NOSQL database, and we keep the compositional graphs in both Neo4J and in GraphDB (a triple store). Labs are coded in LOINC and drugs using RxNorm. We have developed a Web interface in .Net named BMI Search, which allows real-time query by subject matter experts. We analyzed the accuracy of structured Versus unstructured data by identifiying NVAF cases with ICD9 codes and then looked for any additional cases based on the SNOMED CT encodings of the clinical record. This was validated by 2 clinical human review of a set of 300 randomly selected cases. Separately we ran a study to determine the relative risk of OSA with and without Rosacea using the data set described above. We compared the rates using a Pearson χ^2 test. RESULTS/ANTICIPATED RESULTS: We are able to parse 7,000,000 records in an hour and a half on 1 node with 4 CPUs. This yielded 750,000,000 SNOMED CT codes. The HTP data set yielded 1849 cases using ICD9 codes and another 873 using the HTP-NLU data, leading to a final data set of 2722 cases from our population of 212,343 patients. In total, 580 patients had Rosacea;5443 patients had OSA without Rosacea and 51 patients had OSA with Rosacea. Patients with Rosaca had an 8.8% risk of OSA whereas patients without Rosacia only had a 2.6% risk of OSA. This was highly statistically significant with a $p < 0.0001$ (Pearson χ^2 test). The number needed to test was only 12. DISCUSSION/SIGNIFICANCE OF IMPACT: HTP can change how we do observational research and can lead to more accurate and more prolific investigation. This rapid turn around is part of what is necessary for both precision medicine and to create a learning health system. Patients with Rosacea are at increased risk of and should be screened for OSA.

Characterization of resistant hypertension in a statewide electronic health record-based database (OneFlorida)

Caitrin W. McDonough, William R. Hogan, Betsy Shenkman and Rhonda M. Cooper-DeHoff

OBJECTIVES/SPECIFIC AIMS: Our objective is to create a Resistant Hypertension (RHTN) computable phenotype from electronic health record (EHR)-based data, and to determine the characteristics associated with RHTN within a large, diverse, EHR-based database. METHODS/STUDY POPULATION: The OneFlorida Clinical Research Consortium includes 10 unique health care systems providing care for approximately half of the state (48%, ~10 million). OneFlorida houses a Data Trust which contains longitudinal EHR data and claims data from these providers in a common format, the PCORnet common data model v3.0. For the current project, data from 5 health care systems were considered. All of the adult hypertension (HTN) patients with a HTN diagnosis from an outpatient encounter were extracted from the OneFlorida Data Trust. Additional data such as demographics, prescribing, and vitals information were also extracted. The RHTN computable phenotype was created by constructing a drug exposure variable that took into consideration the number of antihypertensive medications an individual was prescribed at any point in time over the course of the OneFlorida dataset. RHTN was defined as any blood pressure requiring four or more antihypertensive drugs, or uncontrolled blood pressure ($\geq 140/90$) on 3 antihypertensive drugs. RHTN cases had to meet the definition criteria twice during the data period, at least 30 days apart. All data extraction, computation phenotype coding, and statistical analyses were conducted using SQL or SAS. RESULTS/ANTICIPATED RESULTS: Our preliminary results show that there were n = 342,026 adults with a HTN diagnosis from an outpatient visit in the data set. After the RHTN computable phenotype was constructed, n = 11,670 RHTN cases were identified from the n = 130,901 HTN individuals with all of the required variables in the data set (8.9% RHTN prevalence). In all, 55% of RHTN cases were Black or African American, compared with the total HTN population (25% Black/African American). RHTN cases also had more prescriptions for loop diuretics, centrally acting agents, α-blockers, and vasodilators compared with the total HTN population. Not surprisingly, the RHTN cases had 26% of the antihypertensive prescriptions in the data set, and the RHTN cases had fewer blood pressure readings that were in control (only 49.4% of readings <140/90). DISCUSSION/SIGNIFICANCE OF IMPACT: Overall, our preliminary data shows that it is possible to create the very complicated computable phenotype of RHTN within the OneFlorida Data Trust. We found that the RHTN prevalence in OneFlorida is 8.9% which is consistent with previous studies from NHANES. Although promising, these results require further validation of the computable phenotype and replication in other similar data sets in order to ascertain their true meaning. Once validated, the experience gained from this computable phenotype can be applied to many other phenotypes.

Identifying causative mutations in Treacher Collins syndrome using iobio

Alistair N. Ward, Matt Velinder, Chase Miller, Tony Di Sera, Yi Qiao, Dave Viskochil and Gabor Marth
The University of Utah School of Medicine, Salt Lake City, UT, USA

OBJECTIVES/SPECIFIC AIMS: The objective of the study was 2-fold; to identify potentially deleterious alleles in a child with Treacher Collins syndrome, and; to demonstrate the value of the iobio analysis platform for intuitively and rapidly analyzing genomic data. METHODS/STUDY POPULATION: We used the iobio suite of web-based applications to analyze quality metrics for the sequencing data and called variants for the proband and his parents. We then visually interrogated variants in genes potentially associated with the syndrome in real-time, using the intuitive gene.iobio application. We sought high impact variants that demonstrated a predicted impact on the protein function, and were simultaneously at low allele frequency in the general human population. Variants were also compared against the ClinVar database of known mutations to identify variants that have already been associated with this, or related syndromes in the literature or clinical studies. Finally, the gene.iobio tool allows users to interrogate the primary sequencing data to ensure that no variants had been missed by the primary variant calling pipeline. This analysis pipeline was performed using intuitive web-based apps in real time, and consequently represents a system that is available to users that traditionally are excluded from these analyses. RESULTS/ANTICIPATED RESULTS: The iobio suite was used to rapidly assess data quality and interrogate genetic variants for a child with

Treacher Collins syndrome. A compound heterozygote consisting of 2 missense alleles in the TCOF1 gene was identified as a compelling pathogenic allele, necessitating further functional investigation. The study helped validate the use of the intuitive iobio tools in such analyses, strengthening the case for greater involvement of medical professionals in data analysis. DISCUSSION/SIGNIFICANCE OF IMPACT: The performed analyses demonstrated that the whole genome sequencing data for the family being studied was of a very high quality, although 1 gene demonstrated a local region of almost zero coverage. This ensured that study conclusions can be presented with confidence. A variant associated with Treacher Collins syndrome 1 in ClinVar was uncovered in the TCOF1 gene, however, given it's benign rating, this variant was not considered further. The most interesting candidate was a compound heterozygote, consisting of 2 missense mutations, also in the TCOF1 gene. These mutations occurred with allele frequencies of 22% and 8% in the general population, and additional molecular and functional studies are currently being pursued.

2286

HOME Cell 2.0. Extending i2b2 to support community health outcome monitoring and evaluation via web-accessible software

William G. Adams, Michael Mendis, Shiby Thomas, David Center and Sara Curran

OBJECTIVES/SPECIFIC AIMS: The primary objective of this effort is to develop and distribute an easy to use i2b2 component that is capable of evaluating diverse complex relationships for a wide variety of exposures and outcomes over time. In this manner we are able to leverage the unique design of the i2b2 database to support health services research, comparative effectiveness, and quality improvement using a single tool. Furthermore, our novel database redesign has the potential to provide user-friendly access to individual and group CHC data for CER. METHODS/STUDY POPULATION: For this project we used software experts, clinical informatics specialists, and the existing i2b2 open-source software to convert our legacy HOME Cell into a web-client version. The tool will be used to study health outcomes within a network of Boston based Community Health Centers and the largest safety-net hospital in New England, Boston Medical Center. RESULTS/ANTICIPATED RESULTS: The new web-client HOME Cell will allow i2b2 users to model virtually any exposure (including therapeutic interventions such as medications or tests) in i2b2 against any outcome accounting for complex temporal relationships and other factors. In addition we plan to use our new Community Health Center views to enhance our community engagement activities by allowing direct access to their data for our partners. DISCUSSION/SIGNIFICANCE OF IMPACT: Our project addresses multiple national priorities related to data sharing, clinical research informatics, and comparative effectiveness. The web-client version of the HOME Cell substantially improves our community's access to HOME Cell functionality and is a novel, sharable resource for use within the CTSA/NCATS community. Our approach provides a new way to perform large-scale collaborative research without the need to actually move patient-level data and has demonstrated that CER, health services research, and quality measurement can share a common framework. In addition, and as demonstrated in our earlier pilot work, the HOME Cell also has the potential to support large-scale multivariate analyses in a distributed manner that does not require sharing of patient-level data. We believe our approach has great promise for supporting the reuse of clinical data for rapid, transparent, health outcome assessments on a national scale. Our efforts support multiple strategic goals including: (1) support for building national clinical and translational research capacity by enhancing a broadly adopted informatics tool (i2b2); (2) enhanced consortium-wide collaborations by offering a tool that can be easily shared within the CTSA network to support multi-institutional collaboration; and (3) improving the health of our communities by offering a tool that has the potential to provide new insights into health care processes and outcomes that could drive innovation and improvement activities.

2289

Will the Veteran Affairs (VA) electronic medical records (EMR) database reveal a signal that angiotensin II inhibiting medications ameliorate depression?

David D. Maron, Marc Blackman, Richard Amdur, Thomas Mellman and Kathryn Sandberg

OBJECTIVES/SPECIFIC AIMS: Angiotensin type 1 receptor blockers (ARBs) and angiotensin-converting enzyme inhibitors (ACEIs) are frequently prescribed for hypertension and associated cardiovascular and renal complications. In animal models, these drugs also reduce anxiety and depression. OBJECTIVE—to determine if Veteran Affairs (VA) clinical pharmacy data indicate a protective effect of ARBs and/or ACEIs for major depression. METHODS/STUDY POPULATION: Pharmacy records from nationwide VA electronic medical records (EMR) were extracted for patients prescribed ARBs, ACEIs, α-blockers, β-blockers, calcium channel blockers, or diuretics (n = 4,081,359). Patients were excluded if: they had not received medications for 6 months with >70% coverage; were diagnosed with substance/alcohol abuse, dementia, psychosis, schizophrenia, or prescribed insulin. The study population was categorized as "ARB/ACEI" (A/A) or "Never ARB/ACEI" (NA/A). Using the Greedy Matching Algorithm, subjects were matched on a 1:1 ratio for sex and race over a 5 year age range resulting in 2 equal groups of n = 1,350,236 each. Subjects were older (M = 71.6, SD = 12) and mostly men (97%). RESULTS/ANTICIPATED RESULTS: In the A/A Versus NA/A, respectively, the incidence of anti-depressant use was greater during (9.9% vs. 8.9%) but was lower after (11.8% vs. 12.2%) the study period. PHQ-2 scores (Mean ± SD) were statistically lower, albeit similar, during (0.79 ± 1.56 vs. 0.85 ± 1.63) and after (1.00 ± 1.73 vs. 1.07 ± 1.79) the study period. DISCUSSION/SIGNIFICANCE OF IMPACT: These preliminary data suggest that inhibiting angiotensin II action does not provide a protective effect on major depression when compared with other classes of antihypertensive drugs. This study illustrates how "Big Data" may inform the design, or obviate the need, for large-scale randomized-controlled trials.

2293

Passive intracranial EEG-based localization of the central sulcus during sleep

Rafeed Alkawadri

Yale School of Medicine, Guilford, CT, USA

OBJECTIVES/SPECIFIC AIMS: To investigate the performance of a metric for passive localization of central sulcus. METHODS/STUDY POPULATION: We studied 7 patients with intractable epilepsy undergoing intra-cranial EEG (icEEG) monitoring at Yale, in whom central sulcus (CS) localization was obtained by standard methods. Our method takes advantage of inherent properties of the primary motor cortex (MC), which exhibits enhanced icEEG band-power and coherence across the CS. For each contact x we calculated the z-score of a composite power and synchrony value $\log_{10}(px) \times x$; cx, where px is sum of the root mean square of the icEEG in the high gamma band (80–115 Hz) for contact x over the 6–10 minutes of NREM sleep studied, and cx is the mean magnitude squared coherence in the same band using a 500-ms Hamming window between contact x and all other contacts. z-score values lower than threshold (th) were set to 0. Finally, we calculated a metric $m = z/d$, where d is the mean Euclidian distance of each contact from contacts with z scores greater than 0. The last step was implemented to emphasize local network activity. RESULTS/ANTICIPATED RESULTS: We report the results of a pilot study to test the performance of CS with a new operator independent method for passive identification of CS with intractable epilepsy undergoing icEEG monitoring at Yale, in whom CS localization was obtained by standard methods. The sensorimotor (SM) cortex exhibited higher EEG-gamma power compared with non-SM cortex ($p < 0.0002$). There was no significant difference between the motor/premotor and sensory cortex ($p < 0.47$). CS was successfully localized in all patients with thresholds between 0.4 and 0.6. In 2 patients, knowledge of anatomy was needed to distinguish the MC from adjacent epileptic foci. The primary hand and leg motor areas exhibited the highest metric values consistently followed by the tongue motor area. Higher threshold values were very specific (94%) for the anterior bank of the CS but not sensitive. Intermediate threshold values achieved a reasonable trade-off (0.4: 89% specific and 70% sensitive). DISCUSSION/SIGNIFICANCE OF IMPACT: We present and successfully implement a rapid procedure for task-free and stimulation free localization of the central sulcus during sleep based on intrinsic electrophysiological properties of the primary motor strip which exhibits increased power and enhanced local connectivity. We successfully localized the central sulcus in all patients. When implementing appropriate thresholds, our proposed metric M is very specific for the anterior lip of the central sulcus which may make it ideal to identify this important anatomical landmark. Our method is sensitive for epileptogenic regions as well, therefore basic knowledge about central sulcus anatomy may be needed in cases where there is an epileptogenic lesion in the vicinity of the central sulcus. Our method makes a few a priori assumptions: The regions around the central sulcus are adequately sampled and the occipital or parieto-occipital regions are not included in the analysis. In order for the method to function properly, nonsensori-MC should be sampled adequately as well. In the future, normative data could be generated for the composite product of connectivity × power which may replace within-patient z-scoring. Our method is rapid and can be implemented on short segments of