and age composition of participants versus non-participants. Patients ≥5 years old with known race and gender and at least one healthcare encounter between 2021 and 2024 were included. Interventional trial enrollment was identified by a "research flag" indicating current or past participation in an interventional study within an Epic system contributing data to Cosmos. Race was categorized as American Indian, Asian, Black, Native Hawaiian, or White. Age-adjusted relative representation (RR) ratios were used to compare participation, with RR >1 indicating over-representation and RR RESULTS/ANTICIPATED RESULTS: Of 130,455,189 patients meeting eligibility criteria, 0.52% (673,425) of patients were active or inactive in an interventional clinical trial. Results are shown in the figure below. The poorest representation was from Asian and NH/PI persons. Representation was most similar to the patient population for whites and AI/AN persons. Black males participated less and women, more than predicted by patient composition. Older patients participated more frequently than younger (age, mean (SD), y, 53 (22) vs. 46 (23); p DISCUSSION/SIGNIFICANCE OF IMPACT: This is the first study we know of describing interventional trial participation in the USA across millions and millions of patients. Further research is needed to clarify whether these differences are due to the nature of the studies themselves (e.g., OB/GYN trials including only women, etc.) versus disparities in recruitment or otherwise.

381

Host-bacterial immune responses to ventilator-associated pneumonia in COVID-19 patients

Cecilia Chung[1], Yaa Kyerematen[2], Kendrew Wong[2], Miao Chang[2], Rajbir Singh[2], Colin Mccormick[2], Anna Czachor[2], Clea Barnett[2], Yonghua Li[2], Tsay Juh-Chieh[3], Leopoldo N. Segal[2] and Benjamin G. Wu[3]

[1]New York University Clinical and Translational Science Institute (NYU Grossman School of Medicine); [2]New York University Grossman School of Medicine – NYUGSoM and [3]New York Harbor Healthcare System

OBJECTIVES/GOALS: Ventilator-associated pneumonia (VAP) is an infection caused by bacteria, viruses, or fungi during mechanical ventilation. We analyzed a cohort of COVID-19 patients admitted to the intensive care unit with respiratory failure with different VAP outcomes. We hypothesize that the multiomics data can help predict VAP development within this cohort. METHODS/STUDY POPULATION: We recruited participants from a cohort on a NYU IRB protocol (i22–00616), who had COVID19 respiratory failure, admitted to ICU, and required invasive mechanical ventilation (n = 245). We collected and analyzed research specimens (bronchoalveolar lavage [BAL, n = 213], tracheal aspirates [n = 246], background [n = 18]) and clinical cultures (sputum and BAL) for 245 participants. A panel of experts adjudicated VAP within the cohort, resulting in 92 VAP diagnoses. We annotated metatranscriptome (Illumina NovaSeq) using a Kraken/Bracken database, and KEGG for functional annotation of transcriptome data (Illumina HiSeq). We used edgeR (v.4.0.16) to analyze differential expression of metatranscriptome and transcriptome data. RESULTS/ANTICIPATED RESULTS: We diagnosed VAP in n = 92 (38%) participants. We found significant differences in days of overall hospital stay (p DISCUSSION/SIGNIFICANCE OF IMPACT: VAP is a serious complication of mechanical ventilation, and oral commensals alter the lung microbiome and host immunity. We identified a transcriptome-metatranscriptome signature that identifies those at VAP risk. VAP was associate with both pro- and anti-inflammatory gene expression resulting in increased risk for lower airway infection.

382

Online graphical interface for bulk to single-cell transcriptomics

Manoj Kandpal and Hong Hur

Rockefeller University

OBJECTIVES/GOALS: We aim to develop an intuitive interface to understand the possible relationships between data from different RNA-Seq technologies. It can help novice users and educators to understand, analyze, explain, and visualize such datasets from diverse platforms, all without the need for additional software installations or strong programming expertise. METHODS/STUDY POPULATION: An online interactive interface is developed, integrating robust algorithms for three distinct types of analyses: DESeq2 for bulk RNA-seq, CIBERSORTx for deconvolution, and Seurat for single-cell analysis, with plans to include more algorithms. It allows a demo mode for training using the sample datasets and option for tailored analysis using user's partially processed data. The interface provides capability to process bulk RNA-seq data from raw counts or a differential gene list. Further, deconvolution analysis for bulk RNA-seq data can be done using raw counts and single-cell data analysis can be performed using processed sequence reads, organized into three key files: barcodes, matrix, and features. Users also have an option to download the results as a zipped file, for samples from human and mouse studies. RESULTS/ANTICIPATED RESULTS: Users with an active internet connection can access the interface from any major web browser. They can adjust parameters – such as genome type, cutoff thresholds, and batch effect correction – according to their specific needs. Bulk RNA-seq results are presented in the form of volcano plots, heatmaps, clusters, gene expression across samples, DEGs, and enrichment plots from KEGG and GO analyses. Deconvolution analysis can be performed using either the "LM22" signature matrix (for human leukocyte cell types) or Derm22 (for skin-specific cell types). The single-cell workflow provides results including quality control metrics, UMAP clustering, gene expression plots/tables, and cluster annotation using CellTypist. Comprehensive details on methods and tutorials are available in the GitHub repository. DISCUSSION/SIGNIFICANCE OF IMPACT: Although multiple workflows are available to process bulk and single cell RNA-Seq data along with deconvolution methods to bridge the gap between the two, this is the first online interface to provide the capability to explore and analyze data from all three approaches in one place, without requiring strong computational expertise or resources.

383

A CTS team approach to leveraging EHR data for predicting necrotizing enterocolitis in NICU*

Keliy Fordham and Yao Hou

University of Florida CTSI

OBJECTIVES/GOALS: This research aims to harness electronic health records (EHR) combined with machine learning (ML) to predict necrotizing enterocolitis (NEC) in preterm infants using data up to their first 14 days of life. We aim to provide interpretable results for clinical decisions that can reduce infant mortality rates and complications from NEC. METHODS/STUDY POPULATION: Through a retrospective cohort study using data from the

https://doi.org/10.1017/cts.2024.1006 Published online by Cambridge University Press

University of Florida Integrated Data Repository and One Florida, we will develop machine learning models suitable for sequential data to predict NEC. Our inclusion criteria include very low birth weight (VLBW; <1500g) infants born <32 weeks gestation and EHR data availability from the first 14 days of life. We will include infants with NEC and infants without NEC to train our ML model. Exclusion criteria include infants diagnosed with spontaneous intestinal perforation and severe congenital anomalies/defects requiring surgery. RESULTS/ANTICIPATED RESULTS: We anticipate that our model will provide an accurate and dynamic prediction for the risk of NEC in neonates using data up to the first 14 days of life. Our model will be interpretable to identify key risk factors and can integrate real-world clinical insights to increase early detection and improve patient outcomes. DISCUSSION/SIGNIFICANCE OF IMPACT: The development of a model to predict NEC could be used in neonatal intensive care guidelines and protocols and could ultimately decrease mortality, reduce complications, improve the overall quality of care, and lower healthcare costs associated with NEC.

384

A systematic approach to understanding nursing documentation tasks

Victoria Tiase, Varun R. Selvam, Julio Facelli and Katherine A. Sward
University of Utah

OBJECTIVES/GOALS: Healthcare organizations must track electronic health record (EHR) activity at the user level, including logons, accessed records, and viewed or entered documentation. There is little standardization in EHR audit logs and nurse workload has not been explored using these data. In this project, we characterized nurse actions from EHR audit logs. METHODS/STUDY POPULATION: We performed an analysis of EHR audit log data collected from 8,149 nurses over 5 years at University of Utah Health. We preprocessed nursing-centric EHR audit logs from the Epic EHR by cleaning and preparing the data for analysis. We calculated basic statistics for the variables labeled user_id (nurse) and metric_id (action). We reviewed the actions used by nurses and categories the actions as navigation, view, and entry. To capture the clinical context of the actions, two nurses categorized each action. A third nurse resolved any discrepancies. RESULTS/ANTICIPATED RESULTS: We found that of the 4,419 available metrics, nurses used 1,461 unique metrics during the timeframe. The actions most used by nurses were 1) report with patient data viewed, 2) inpatient system list, and 3) storyboard viewed. Most of the metrics were categorized as navigation. The number of nurses interacting with the EHR increased each year and on average, we found that 1000 unique metrics were used by each nurse user in a 24-hour period. The expected outcome is a set of actions that can be mapped to higher level nursing interventions and in the future contribute to models for nursing workload measurement. DISCUSSION/SIGNIFICANCE OF IMPACT: We found great value in using EHR audit logs to provide insights into nursing actions. Information gleaned can benefit organizations that are crafting interventions to decrease workload. Ultimately, the goal is to ensure that nurses have an appropriate workload allowing for safe and high-quality patient care while maintaining their well-being.

385

Impact of skin pigmentation on genomic alterations and mutation load

Rojina Nekoonam, Aravind K. Bandari, Delahny Deivendran, Bishal Tandukar, Harsh Sharma and Alan Hunter Shain
University of California San Francisco

OBJECTIVES/GOALS: Despite the documented link between ultraviolet (UV) exposure and skin cancer, the potential protective role of pigmentation in preventing skin carcinogenesis is not fully understood. This project will test the hypothesis that individuals with light skin tones are more susceptible to skin cancer due to a higher accumulation of somatic mutations from UV exposure. METHODS/STUDY POPULATION: Skin biopsies were collected from various anatomical sites of seven donors aged 70 to 80 years with dark skin tones and used to generate clones of normal skin cells. The clones were then subjected to whole-exome and RNA sequencing. We developed a workflow that allows us to accurately detect somatic mutations in clonal expansions of individual cells with high specificity and sensitivity. For comparison, we also analyzed mutations in skin cells obtained from individuals with light skin tones, who served as our control group. In the future, we plan to employ admixture analysis to clarify the impact of race on somatic mutations by examining ancestral genetic contributions. RESULTS/ANTICIPATED RESULTS: Our preliminary analysis of the existing data supports our hypothesis. The average mutation burden in individuals with light skin tone was found to be 3.778 Mut/Mb, whereas in those with dark skin tones, was only 0.818 Mut/Mb. These findings align with the prevailing hypothesis that melanin acts as a protective factor against skin cancer, as it appears to reduce the mutagenic effects of UV radiation. This suggests that individuals with darker skin may have a lower cumulative mutational load, potentially contributing to their reduced risk of developing skin cancers such as melanoma. Similarly, our results showed variations in mutational burdens across different anatomical sites, which seemed to be influenced by differing levels of UV exposure, with the highest burdens observed in areas with the greatest exposure. DISCUSSION/SIGNIFICANCE OF IMPACT: Understanding how skin cancer operates in darker-skinned individuals is imperative for tailoring effective screening and targeted therapies to meet their specific needs. In the long run, elucidating the mechanisms underlying skin cancer development in this demographic will help refine screening protocols and prevention recommendations.

386

Developing an assessment tool for NIH data management and sharing plans to understand current data practices and needs

Michelle Yee, Alisa Surkis and Fred LaPolla
NYU Langone

OBJECTIVES/GOALS: NIH requires researchers submit Data Management and Sharing (DMS) Plans with their grant applications. Librarians developed an assessment tool for the plans and completed a pilot assessment in order to leverage the plans and understand current institutional research data management and sharing.

https://doi.org/10.1017/cts.2024.1006 Published online by Cambridge University Press