

## Biostatistics, Epidemiology, and Research Design

1

### Analysis of dietary intake and biliary tract cancer risk in 638,860 Chinese and European adults in three cohort studies

Valerie Gunchick, Wei Zheng

<sup>1</sup>Division of Epidemiology, Vanderbilt University

**OBJECTIVES/GOALS:** Biliary tract cancer is an uncommon, aggressive malignancy. Incidence varies geographically and is highest in East Asia and South America and lowest in Western countries. Previous dietary risk evaluations have primarily been case-control studies. We evaluated associations of dietary intakes with BTC risk in three cohort studies in two countries. **METHODS/STUDY POPULATION:** We evaluated 638,860 adults in China and the United Kingdom enrolled in the Shanghai Women's Health Study (SWHS), Shanghai Men's Health Study (SMHS), or UK Biobank (UKB). Dietary intake information was obtained from study participants at baseline using food frequency questionnaires previously validated for these studies. Dietary intakes of major food groups and macronutrients were divided into low, middle, and high intake tertiles. Cox regression was used to estimate hazard ratios and 95% CIs for biliary tract cancer risk associated with major food groups in all three cohorts and macronutrients in the SWHS and SMHS. Participants were excluded if, at enrollment, they had a history of cancer, CHD, stroke, a total daily Kcal count below 500 or exceeding 3500, or developed cancer or died within one year after enrollment. **RESULTS/ANTICIPATED RESULTS:** The analyzed cohort includes 558,372 participants: 66,945 SWHS, 55,750 SMHS, and 435,677 UKB participants with median enrollment ages of 49, 52, and 57 years old and median follow-up times of 18.1, 12.3, and 10.3 years, respectively. The SWHS observed 205 eligible BTC cases, SMHS 97, and UKB 366. SWHS and SMHS participants were combined for dietary evaluation and the highest tertile of fruit intake showed an inverse association with biliary tract cancer risk when compared to the lowest tertile: HR 0.74 (95% CI, 0.55-0.99); p-trend 0.044. In UKB, the highest tertile of fish intake was associated with a reduced risk when compared to the lowest tertile: HR 0.76 (95% CI, 0.59-0.98); p-trend 0.034. **DISCUSSION/SIGNIFICANCE:** High dietary fruit intake was associated with a reduced risk of biliary tract cancer only in SWHS and SMHS. High fish intake was associated with a reduced risk of biliary tract cancer only in UKB. Our findings reflect geographic-based BTC risk variation which we will further explore in our next model accounting for population-specific risk factors.

2

### A CTS Team Approach to Topological Data Analysis of Electronic Health Records for Subtyping and Clinical Outcomes Prediction in Patients with COVID-19

Yara Skaf, Osama Dasa, Jason Cory Brunson, Thomas Pearson, Reinhard Laubenbacher

University of Florida

**OBJECTIVES/GOALS:** Analysis and modeling of large, complex clinical data remain challenging despite modern advances in

biomedical informatics. We aim to explore the potential of topological data analysis (TDA) to address such challenges in the context of COVID-19 outcomes using electronic health records (EHRs). **METHODS/STUDY POPULATION:** In this work, we develop TDA approaches to characterize subtypes and predict outcomes in patients with COVID-19 infection. First, data for >70,000 COVID-19 patients were extracted from the OneFlorida EHR database. Next, enhancements to the TDA algorithm Mapper were designed and implemented to adapt the technique to this type of data. Clinical variables, including patient demographics, vital signs, and lab values, were then used as input to conduct a population-level exploratory analysis with an emphasis on identifying phenotypic subtypes at increased risk of adverse outcomes such as major adverse cardiovascular events (MACE), mechanical ventilation, and death. **RESULTS/ANTICIPATED RESULTS:** Preliminary Mapper experiments have produced visual representations of the COVID-19 patient population that are well-suited to exploratory analysis. Such visualizations facilitate easy identification of phenotypic subnetworks that differ from the general population in terms of baseline variables or clinical outcomes. In this and subsequent work, we aim to fully characterize and quantify differences between these subnetworks to identify factors that may confer increased risk (or protection from) adverse outcomes. We also plan to validate and rigorously compare the efficacy of this TDA-based approach to common alternatives such as clustering, principal component analysis, and machine learning. **DISCUSSION/SIGNIFICANCE:** This work demonstrates the potential utility of TDA for the characterization of complex biomedical data. Mapper provides a novel means of exploring EHR data, which are otherwise difficult to visualize and can aid in identifying or characterizing patient subtypes in diseases such as COVID-19.

3

### Addressing Recovery Support, Social determinants of Health and Treatment Retention in Postpartum and Parenting women with Opioid Use Disorder

Doris Titus-Glover, Fadia Shaya, Chris Welsh, George Unick, Kelly Doran

University of Maryland, Baltimore

**OBJECTIVES/GOALS:** The purpose of this study is to determine critical recovery support factors (SDOH, postpartum and post-discharge continuity of care), to optimize continuity of recovery and to determine the best intervention among postpartum and parenting women for treatment retention. **METHODS/STUDY POPULATION:** Through a mixed methods approach, we will review retrospective hospital discharge data to identify hospital-based gaps in treatment. We will conduct key informant interviews with postpartum women, treatment providers and stakeholders to broaden understanding of critical recovery factors from lived experiences and test a parent-centered evidence-based intervention for a comprehensive and targeted approach to recovery. **RESULTS/**

\*Blue Ribbon Awardee

© The Author(s), 2023. The Association for Clinical and Translational Science. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives licence (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is unaltered and is properly cited. The written permission of Cambridge University Press must be obtained for commercial re-use or in order to create a derivative work.