



Biomedical Informatics Grand Rounds

Wednesday, October 23, 2024

3:00 pm – 4:00 pm

Cancer InFocus: An Efficient Solution for Gathering and Visualizing Cancer Surveillance Data

Todd Burus, Ph.D.

Data Scientist, Community Outreach and Engagement Office, Markey Cancer Center

Join Zoom Meeting <https://stonybrook.zoom.us/j/95617197636?pwd=KytzZ2pVRG9SZGpKZUtpNXJISjNjZz09>

Meeting ID: 956 1719 7636 Passcode: 924293

Bio:

Todd Burus, MAS, is a Data Scientist for the Community Outreach and Engagement office at Markey Cancer Center. He serves as the lead developer on Markey's Cancer InFocus project, working to improve cancer surveillance data gathering and dissemination processes for US cancer centers. He is also an Epidemiology and Biostatistics PhD candidate at the University of Kentucky with a focus on advanced statistical methods for cancer surveillance.

Abstract:

Background: Successfully developing strategies for addressing the cancer burden in a geographic area requires looking at quantitative data related to a number of different topics such as demographic and socioeconomic factors, educational attainment and risk factor prevalence, as well as cancer incidence and mortality rates. Most of this data is publicly available, but exists across numerous sources, at various geographic levels and in different file formats. This makes collecting, summarizing and visualizing it an inefficient and time-consuming process. **Methods:** We constructed programs in Python to access data from various publicly-available sources through application programming interfaces, automated data downloads and web scraping. This data was then manipulated into datasets at different geographic levels, and exported as an organized collection of files. The data is then transformed into interactive mapping applications using R Shiny. All code was structured to allow for automation of updates, and generalized for easy adaptation to any cancer center catchment area structured as a set of US counties. **Results:** This process resulted in a comprehensive software solution licensed under the name of Cancer InFocus. Cancer InFocus creates a quick, efficient and automatable mechanism for gathering much of the data necessary to monitor the cancer burden in any US geographic area of interest and translate it into simple applications for either internal or external distribution. Cancer InFocus is available through a no-cost licensing agreement with the University of Kentucky. The functionality of Cancer InFocus is maintained and expanded upon by the online community of users who have chosen to adopt this platform. **Impact:** The construction of this process for gathering and displaying publicly available data for a cancer center catchment area makes doing an initial assessment of the cancer burden more efficient, allowing greater time to be spent on developing strategic priorities and operationalizing insights. The use of open source tools to perform this task allows for its free dissemination to other institutions looking for a ready-made solution to monitor their catchment area.

Educational Objectives:

1. Learn how to integrate data from multiple sources (demographic, socioeconomic, and health data) to assess the cancer burden within specific geographic areas.
2. Understand how Python can be used to automate the data-gathering process through APIs, web scraping, and automated downloads.
3. Explore open-source solutions like Cancer InFocus, which facilitate data access and promote collaboration.
4. Learn how streamlined data workflows help institutions allocate more time to developing strategic priorities and actionable insights for cancer control.

Disclosure Statement:

The faculty and planners have no relevant financial relationship with ineligible companies, whose primary business is producing, marketing, selling, reselling, or distributing health care products used by or on patients.

Continuing Medical Education Credits:

The School of Medicine, State University of New York at Stony Brook, is accredited by the Accreditation Council for Continuing Medical Education to provide continuing medical education for physicians. The School of Medicine, State University of New York at Stony Brook designates this live activity for a maximum of **1 AMA PRA Category 1 Credits™**. Physicians should only claim credit commensurate with the extent of their participation in the activity.