Title: Developing a Predictive Model for HIV Clinical Care Disengagement at Duke

Principal Investigator: Nwora Lance Okeke MD, MPH

Rationale:
Missed HIV clinic visits are independently associated with all cause mortality in persons living with HIV\(^1\). As a result, developing innovative tools to help identify persons likely to disengage is critical to maintaining persons in lifesaving HIV care. Although a few risk stratification methods have been developed by our group and others to address HIV care disengagement, many of these tools are not generalizable, and difficult to readily implement in the local context\(^2,3\). The electronic health record gives clinicians and researchers access to numerous variables that report on relevant factors associated with access to care in a quantifiable, longitudinal and reproducible manner. Moreover, the availability of these variables on the electronic health record allows for easy incorporation into risk scores that can be instantly queried at the clinical interface, for use by clinicians. Predictive analytic methods have bolstered our ability to reduce multiple EHR variables on multiple observations to clinically useful parsimonious models\(^4\). In HIV treatment and prevention, this approach has been deployed to for identification of persons at risk for HIV infection, enabling early initiation of HIV pre-exposure prophylaxis (PrEP)\(^5,6\). We hypothesize that a similar approach can be utilized to identify HIV-positive persons in care who at increased risk for falling out of HIV care.

Aims:
- To assess a set of EHR-ready candidate variables relevant to HIV care access for their independent association with HIV care disengagement (defined as a failure to attend HIV clinics for 12 months or more after a documented period in HIV care).
- To develop a predictive model for HIV care disengagement based on the aforementioned variables
- To validate the model on locally-available contemporary clinical data

Objectives of Internship:
- To gain knowledge on HIV care engagement and retention as well as barriers to HIV care
- To familiarize student with longitudinal clinical data and the electronic health record
- To provide practical experience in clinical data science methodology including predictive analytics

Scope of Work:
Datasets for relevant variables will be provided to the prospective student for analysis. The investigator and the student will ensure that the provided data is adequately curated and “fit for use”\(^7\). The student will then help develop a predictive model using least absolute shrinkage and selection operator (LASSO)- based approach\(^5,6\). The entrainment dataset will be Duke HIV Clinic data from 2014-2016. If the student’s advisor believes that other methods should be employed, this will be discussed and decided upon with the investigator. Once the model is developed and trained, it will be validated on Duke HIV Clinic data from 2016-2018.
REFERENCES:


