**PCSI Syndemic Data Set**

**Overview and Purpose:** The Program Collaboration and Service Integration (PCSI) Syndemic Project matches HIV, STD, TB, communicable diseases, diabetes surveillance data and vital statistics mortality data in order to understand syndemics of disease. The goal is to better understand the characteristics of NYC residents with co-infections, visualize the geographic distribution of disease, disseminate data on co-infections, increase data sharing across DOHMH programs, and facilitate the delivery of integrated services.

**Methodology:** The PCSI Syndemic Project includes a retrospective match of prevalent cases of HIV, chronic hepatitis B (HBV) and chronic hepatitis C (HCV) diagnosed before and alive as of 2000 and incident cases of HIV, HBV, HCV, tuberculosis, sexually transmitted diseases (chlamydia, gonorrhea, and syphilis) during 2000–2013; sexually transmitted disease incidence data (chlamydia, gonorrhea, and syphilis) during 2000–2013; vital statistics mortality data during 2000–2013; and hemoglobin A1C reports during 2006-2013. An additional 30 communicable diseases were also matched. A deterministic method was used by the Bureau of HIV Data Support Unit to link records across multiple disease registries. Matches were established using 14 keys comprised of combinations of first name, last name, date of birth, and social security number, with the higher numbered keys having more flexibility. Following the match, discordance in static demographic variables between disease datasets was addressed and geospatial data were coded, based on address at report. All records were de-identified and assigned a unique ID.

**Response and Cooperation Rates:** N/A

**Strengths and Limitations:** Co-infection status depends on whether or not a person matched to other disease registries. Therefore, potential for over- or under-matching may affect the validity of our results. These data do not account for persons who may have re-located outside of New York City during the study period. Due to the observational nature of this research and limited demographic and risk data, the potential for unmeasured or mis-measured factors to bias our results exists.

**Priority Research Questions:**
- Ecological/geographic analyses to identify and validate population-level variables (in addition to poverty) that could be used to examine co-infection.
- Understand disease burden and clinical outcomes among persons with multiple co-infections to identify needs for integrated services opportunities for provider training.

**Publications:**