Population Health Data System

**Objectives**

- To develop a flexible, de-identified person-level composite for the Maine Health Data Organization (MHDO) that can be used for research and analysis.
- To improve data quality and completeness by addressing data entry and collection errors.
- To capture multi-racial identities and provide flexibility in data group selections.

**Study Design**

The MHDO APCD includes medical, pharmacy, and dental claims data from various sources. Data collection practices vary across hospitals and payors, making it difficult to replicate or ensure data consistency.

**Principal Findings**

- **Race/Ethnicity Reporting Methods**
  - **Most Recent Race/Ethnicity**: Includes the last reported race/ethnicity for an individual, regardless of whether or not it is updated in subsequent data.
  - **Most Recent Race/Ethnicity (Data Source)**: Includes the most recent race/ethnicity recorded in a specific data source, even if it differs from other sources.

- **Consistency of Race/Ethnicity Data by Data Source**
  - **Across all systems and data sources**: 95%
  - **Within an individual's data across all data sources**: 99%

- **Data Set Comparisons**
  - **APCD vs. Hospital/Outpatient Encounter Data**: 98% consistency for race data.
  - **APCD vs. Claims Data (APCD)**: 98.5% consistency for race data.

- **Demographic Breakdown by Race Source**
  - **Minority Populations**: 20% of the population (excluding White).
  - **Race and Ethnicity Data Systems**: 95% completeness for race data.

**Limitations & Implications**

- **Data Collection Practices**
  - **Hospital and Outpatient Data**: Data is collected at the hospital and outpatient level, and consistency across different data sources is challenging.
  - **Claims Data (APCD)**: Data is collected at the claim level, and consistency is affected by data entry errors.

**Conclusions**

- Data quality and completeness are critical for improving access, costs, and outcomes in health care.
- Custom logic can be developed and applied to address specific research questions.
- Future work should focus on integrating additional data sources and improving data quality through enhanced data entry and submission practices.