PhenX Toolkit

Data Collection Worksheet

> **Please Note:** The Data Collection Worksheet (DCW) is a tool to aid integration of a PhenX protocol into a study. The PhenX DCW is not designed to be a data collection instrument. Investigators will need to decide the best way to collect data for the PhenX protocol in their study. Variables captured in the DCW, along with variable names and unique PhenX variable identifiers, are included in the PhenX Data Dictionary (DD) files.

Density/Proximity of Tobacco Retailers to Known Residence

1. The Young-Wolff et al. protocol obtained addresses of licensed tobacco retailers from a state agency (board of equalization) and geocoded retailer and participant addresses (mapping rate=96.3%).
2. To compute tobacco retailer density, defined in the Young-Wolff et al. protocol as retailers per acre, the ArcGIS Spatial Join tool was used to calculate the count of tobacco retailers within 500-meter and 1-kilometer service areas around each participant address, then divided the count by the land area for each buffer. Land area was extracted from Census 2010 data.
3. To compute proximity, the ArcGIS Closest Facility tool was used to calculate the roadway distance (in meters) from each participant to the nearest tobacco retailer. Distance was positively skewed and the top 1% of observations were capped at the value for the 99th percentile.

Protocol source: https://www.phenxtoolkit.org/protocols/view/741202