**Prevalence Calculations Description**

This document describes the methodology used to obtain estimates of prevalence in children and adults at the county level. The diseases calculated for children were ADHD, Asthma, Depression, and Diabetes. For adults, Asthma, Cardiovascular diseases, Depression, and Diabetes were calculated. To obtain these calculations, patient claims data was used, and the disease of each patient was determined using the EDC (Episode Diagnostic Category) classification. Diseases are grouped into one of 533 EDCs [1], and certain diseases fall under multiple EDCs. In the data, the EDCs associated with each patient were given. Along with the EDCs associated with each patient, the data included the number of months during 2012 that each patient was on Medicaid. After mapping each disease to a specific EDC (or multiple EDCs), a Python script was used to determine which patients were receiving treatment for each given disease. The prevalence of each disease was calculated by dividing the total number of member months of patients treated for the given disease by the total number of member months of all patients on Medicaid for each county. This calculation is segmented by children and adults, and the prevalence of certain diseases vary depending on age. Children include all patients under the age of 18, and adults are classified as all patients over the age of 17.

The python script also checked to ensure counties with less than 11 patients receiving treatment for the given disease are grouped with other counties of less than 11 patients to follow the guidelines required of all data that is removed from the room. The counties that required this grouping were grouped using their Rural Urban Continuum Code (RUCC). The code is a single digit (1-9) classification that groups counties based on the population of their metro area or their proximity to an urban area [2]. Counties with less than 11 patients receiving treatment were grouped with other counties in the same RUCC classification that also needed to be grouped. If the entire group of counties within an RUCC classification still did not have at least 11 patients receiving treatment, counties in the next closest RUCC were grouped with it until at least 11 patients were in the group. The prevalence of the entire group was calculated and reported for all counties included in the group. This process was not necessary for certain diseases that were more common. After the analysis was complete, a file was created for children and adults for each disease. The files each contained the county FIPS code and the prevalence for that disease.

[1] “Identifying and Classifying Children With Chronic Conditions Using Administrative Data With the Clinical Risk Group Classification System.” John M. Neff, MD; Virginia L. Sharp, MA; John Muldoon, MHA; Jeff Graham, MD; Jean Popalisky, RN; James C. Gay, MD. <http://www.dhcs.ca.gov/services/ccs/Documents/IDCronicConditions.pdf>

# [2] “Rural-Urban Continuum Codes.” *United States Department of Agriculture Economic Research Service*.

# http://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx