Supplementary Online Content


eTable. Completeness for Each Step of Data Preparation Process

eAppendix. Data Preparation Process

eFigure 1. Racial Composition Variable Selection

eFigure 2. Age, Sex, Time Trends in Prescription Prevalence of Opioids, Benzodiazepines, and Stimulants

This supplementary material has been provided by the authors to give readers additional information about their work.
**eTable. Completeness for Each Step of Data Preparation Process**

<table>
<thead>
<tr>
<th>Year</th>
<th>Observations</th>
<th>Non-Missing and Valid Age and Binary Sex (%)</th>
<th>Matches NDC Mapping (%)</th>
<th>Matches ZCTA ZIP Code Mapping (%)</th>
<th>Valid Patient ID (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>46,335,044</td>
<td>99.97571</td>
<td>98.75568</td>
<td>98.65795</td>
<td>100</td>
</tr>
<tr>
<td>2012</td>
<td>50,058,970</td>
<td>99.96371</td>
<td>98.30636</td>
<td>98.68315</td>
<td>100</td>
</tr>
<tr>
<td>2013</td>
<td>50,018,803</td>
<td>99.94743</td>
<td>97.43226</td>
<td>98.49504</td>
<td>100</td>
</tr>
<tr>
<td>2014</td>
<td>50,750,360</td>
<td>99.93883</td>
<td>97.02643</td>
<td>98.45746</td>
<td>100</td>
</tr>
<tr>
<td>2015</td>
<td>50,785,251</td>
<td>99.90991</td>
<td>97.692</td>
<td>98.34997</td>
<td>100</td>
</tr>
</tbody>
</table>

eTable 1. Completeness statistics from each step of the process explained above. Note that each value represents the percent of complete observations after the prior steps have been completed, so that the final percent of complete values is simply the multiplication of each value, used as a proportion.
eAppendix. Data Preparation Process

The following steps detail the basic data preparation process used in the analysis.

We start with all observations from the CURES database for 2011 through 2015. For each year we performed the following operations:

1) We determined observations that have a valid age between 0 and 109 years, and gender of “male” or “female.” We removed observations with missing values, and recorded percent missing (see below).

2) We determined observations that have an NDC code matching our mapping from NDC code to drug class (opioids, benzodiazepines, stimulants, and other). We recorded the percent with no match (see below). We merged on drug class mapping, dropping non-matching observations.

3) We removed prescriptions commonly given for medication assisted therapy, including methadone and buprenorphine.

4) We determined observations that have a valid Zip Code matching our Zip Code to ZCTA mapping. We recorded the percent with no match (see below). We merge on the mapping, dropping non-matching observations.

5) We recorded the percent with valid a patient ID, recording the percent missing, and dropping individuals missing valid patient ID (see below).

6) We summed counts of unique patients by age group, sex, and drug class. We combined them with state-level populations to calculate state-level prevalence.

7) We summed counts of unique patients by ZCTA, age group, sex, and drug class. We combined them with ZCTA-level populations to calculate ZCTA-level prevalence.

In order to calculate prevalence by quintiles of neighborhood race and income characteristics, we performed the following operations:

1) We extracted ZCTA-level racial composition and average per-person income data from files obtained from the American Community Survey 5-year files. Although the files represent characteristics for a 5-year period, we used the last year represented by each file to combine race and income with single year prevalence data calculated as described above. For example, we combined race and income data 2007 to 2011 with prevalence data from 2011.

2) We calculated quartiles of the percent of individuals identifying as Non-Hispanic White, and income, across all study years and ZCTAs. We calculated prevalence for each quintile, as well as all 25 quintile-quintile combinations of race and income.

Sources for Mappings and Shapefiles

Below are the sources for each mapping and shapefile used in the analysis.


In order to select a metric to represent racial composition, we considered the ZCTA-level association between each of the three drug classes (on the X axis) and four racial groups (on the Y axis): non-Hispanic White, Black, Latino, and Asian. Each box shows the regression coefficient of a bivariate regression with the prescription prevalence as the dependent variable, and the proportion of the racial group as the explanatory variable. The regression coefficients are standardized for differences in the magnitude of the prevalence of each drug class by dividing by the standard deviation of prevalence. A positive value, shown in purple, indicates a positive association. A negative value, shown in green, indicates a negative association. The $p$-value of each regression is shown in parentheses to highlight statistical significance. All 3 drug classes are significantly positively associated with the percent of white individuals. All drug classes showed a statistically significant negative relationship with the proportion of individuals identifying as Black, Latino, and Asian, with the exception of stimulants and Asian individuals, which had a non-significant relationship and a small magnitude. We therefore chose to use a simplistic representation of racial composition, the percent of individuals identifying as white, as it captures the mainly dichotomous nature of the racial differences in prescription prevalence.
eFigure 2. Age, Sex, Time Trends in Prescription Prevalence of Opioids, Benzodiazepines, and Stimulants

eFigure 2. Basic demographic and time trends for prescription prevalence. Opioid and benzodiazepine prescription prevalence increase steadily with age. Across all age groups opioid and benzodiazepine prescription rates are higher among women than men. Stimulants are the only drug class observed to be negatively associated with increasing age, with a strong peak among individuals aged 10-14, especially males. Little time trend is observed in any drug class.