Supplementary Online Content


eMethods

eReferences

eTable. Sensitivity Analysis for Marijuana Joint-Year Estimation Assumptions

eFigure 1. Associations Between Continuous Smoothed Exposure to Current and Lifetime Tobacco and Marijuana and Pulmonary Function With No Recoding

eFigure 2. Air Pollution Data From the Environmental Protection Agency, With Imputation

This supplementary material has been provided by the authors to give readers additional information about their work.
eMethods

Joint-Year Calculation Assumptions

In order to calculate joint-years of exposure to marijuana, we used repeated self-reports of current marijuana use at each examination, past exposure to marijuana use, and the intervals between exams. Current marijuana use was assessed at each examination using two questions: “During the last 30 days, on how many days did you use marijuana?”, and “…when you used marijuana, what was the usual number of joints or pipe fulls you smoked per day?” We made the following assumptions when interpreting responses to these queries: 1) participants reporting current marijuana exposure in the last 30 days but “0” joints/episode actually smoked part of a joint, so we assigned a value of “0.5” joints/episode; and 2) participants indicating greater than 5 joints/episode (although not predominant, we observed responses of 10, 15, 20, 30 and above) interpreted this question incorrectly as querying the total number of joints smoked in the last 30 days. The cutoff of greater than 5 joints/episode represents our best guess about where participant errors occurred in responding to this question, and is consistent with the distribution of responses about marijuana use reported in a White House report on drug use published in 2001. However, some responses lower than 6 are also likely to be errors, while some responses higher than 6 are likely to be correct.

To assure that our results were not sensitive to this assumption, we performed the following sensitivity analyses. We varied the cutoff for recoding of putative participant errors between 2 and 10 joints/episode, ran a simulation where very high values for lifetime exposure to marijuana were “trimmed” (values > 99th percentile were set to the 99th percentile value), and ran a “no recoding” simulation without trimming or recoding (except for recoding of 0 to 0.5, as in assumption 1, above).

The Appendix Table shows that p-values for overall association and for non-linearity for all smoking exposure variables (as presented in Figure 2) were nearly unchanged across scenarios. The Appendix Figure, which is identical in format and interpretation to Figure 2 in the manuscript, illustrates the shapes of smoking-pulmonary function associations in the “no recoding” scenario. No important differences are evident, though the different scale of the x-axes for the lifetime exposure plots should be noted. One important estimate, the slope for FEV1 per additional joint-year of exposure at >10 joint-years, was slightly smaller in magnitude but became statistically significant in the “no recoding” scenario: -1.4 ml/joint-year (-2.5
-0.3), p=.014 (compared with p=.079 in the base case, see Table 3). Other slope and net association estimates for each alternate scenario are available by request.

Air Pollution Imputation

In order to adjust for changes in air quality over time at the different CARDIA sites, we included in fully adjusted models the average annual measurements of airborne particulate matter < 10 microns in size (PM10) and particulate matter <2.5 microns in size (PM2.5), which are considered major pollutants and are associated with pulmonary disease. Measurements were linked to each observation by exam date and study center. In addition to these direct adjustments, we also included center, year and center by year interaction terms as fixed effects in our adjusted models, which should address a variety of ambient/environmental indicators of air pollution above and beyond particulate matter.

We gathered our pollutant data from the Environmental Protection Agency, which included PM10 from 1990-2009 and PM2.5 from 1999-2009 for Birmingham-Hoover, AL (for our Birmingham center), Chicago-Naperville-Joliet, IL-IN-WI (for our Chicago center), Minneapolis-St. Paul-Bloomington, MN-WI (for our Minneapolis center), and San Francisco-Oakland-Fremont, CA (for our Oakland center). Using linear regression, we modeled the time trend in PM10 within each city using a 4-knot cubic spline, and extrapolated the trend back to 1985 to imputed missing values of PM10 in 1985-1989. We then fit another linear regression model for PM2.5 that used the time trend spline as well as the imputed PM10 values, with interactions by clinic included. We used this model to impute missing values of PM2.5 from 1985-1999. Data points and models are illustrated in Appendix Figure 2. Imputed values for PM10 and PM2.5 for each clinic-exam date were included in regression models in the main adjusted pulmonary function analyses.

Marijuana and Tobacco Exposure Splines

As described in the Methods, we used cubic splines to explore continuous non-linear associations between smoking exposure variables and pulmonary function. Spline knots, chosen based on the distribution of data for each exposure variable, were included at 5, 10, 15, and 20 for tobacco cigarettes per day and for marijuana episodes per month, at 2, 5, 10 and 20 for pack-years of tobacco, and at 1, 2, 4 and 10 for joint-years of marijuana. Our results were not sensitive to alternate reasonable choices for the placement of these spline knots.
eReferences


<table>
<thead>
<tr>
<th>Assumption</th>
<th>Overall</th>
<th>Non-linearity</th>
<th>Overall</th>
<th>Non-linearity</th>
<th>Overall</th>
<th>Non-linearity</th>
<th>Overall</th>
<th>Non-linearity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joints per episode response is correct, and no trimming (NO RECODING, see Appendix Figure 1)</td>
<td>FEV1: &lt;.001, FVC: .004</td>
<td>FEV1: &lt;.001, FVC: .01</td>
<td>FEV1: .064, FVC: &lt;.001</td>
<td>FEV1: &lt;.001, FVC: &lt;.001</td>
<td>FEV1: &lt;.001, FVC: &lt;.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joints per episode &gt; 2 recoded assuming this indicates participant error*, and no trimming</td>
<td>FEV1: &lt;.001, FVC: .004</td>
<td>FEV1: &lt;.001, FVC: .01</td>
<td>FEV1: .060, FVC: &lt;.001</td>
<td>FEV1: &lt;.001, FVC: &lt;.001</td>
<td>FEV1: &lt;.001, FVC: &lt;.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joints per episode &gt; 5 recoded assuming this indicates participant error*, and no trimming (BASE CASE†, see Figure 2)</td>
<td>FEV1: &lt;.001, FVC: .003</td>
<td>FEV1: &lt;.001, FVC: .01</td>
<td>FEV1: .059, FVC: &lt;.001</td>
<td>FEV1: &lt;.001, FVC: &lt;.001</td>
<td>FEV1: &lt;.001, FVC: &lt;.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joints per episode &gt; 9 recoded assuming this indicates participant error*, and no trimming</td>
<td>FEV1: &lt;.001, FVC: .003</td>
<td>FEV1: &lt;.001, FVC: .01</td>
<td>FEV1: .060, FVC: &lt;.001</td>
<td>FEV1: &lt;.001, FVC: &lt;.001</td>
<td>FEV1: &lt;.001, FVC: &lt;.001</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### eTable (Continued)

<table>
<thead>
<tr>
<th>Joints per episode &gt; 5 recoded assuming this indicates participant error*, with trimming of values over the 99th percentile‡</th>
<th>Overall</th>
<th>&lt;.001</th>
<th>.003</th>
<th>&lt;.001</th>
<th>&lt;.001</th>
<th>.063</th>
<th>&lt;.001</th>
<th>&lt;.001</th>
<th>&lt;.001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-linearity</td>
<td>.29</td>
<td>.73</td>
<td>.99</td>
<td>.88</td>
<td>.026</td>
<td>.034</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

* - Assuming this response indicates that the participant interpreted this question as "How many joints did you smoke in the last 30 days?", we recoded joints per episode by dividing by episodes per month.

† - This was our base-case assumption. P-values in this row match those presented in Table 3.

‡ - Values > 99th percentile were set to the 99th percentile value
eFigure 1. Associations Between Continuous Smoothed Exposure to Current and Lifetime Tobacco and Marijuana and Pulmonary Function With No Recoding

The figure on this and the subsequent page is identical to Figure 2 in the print manuscript (see Legend for Figure 2), except that no recoding of the joints per episode response variable was undertaken, and represents the NO RECODING model presented in the first row of the Appendix Table. Note the different scale of the x-axis for the lifetime exposure plots.
eFigure 2. Air Pollution Data From the Environmental Protection Agency, With Imputation

Values were imputed from 1985-1989 for PM10, and from 1985-1999 for PM2.5 (using PM10 values). Clinics were numbered alphabetically (1: Birmingham; 2: Chicago; 3: Minneapolis; 4: Oakland). PM10 – Particulate matter, 10 microns or less in size; PM2.5 – Particulate matter, 2.5 microns or less in size.