**Diabetes Mellitus: Is Cannabis Smoking Protective?**

**Results from Meta-Analysis with Eight Independent Replication Samples**

**Omayma Alshaarawy, PhD & James C. Anthony, PhD**

**Electronic Appendix**

**Study population**

1. *The National Health and Nutrition Examination Survey (NHANES) 2005-2012*

Each NHANES replication is designed to yield nationally representative sample survey estimates for the US non-institutionalized civilian population through multistage area probability sampling. In 2005-2012, selection was based on counties, blocks, households and individuals within households, and included oversampling of certain groups in order to provide stable estimates of these groups. Data were collected from home interviews, questionnaire, physical examinations and laboratory tests of biological samples. The data have been released for public use in 2-year increments since 1999. In the current study, NHANES replication estimates can be derived by combining survey cycles as follows: 2005-6; 2007-8; 2009-10; 2011-12. The within-cycle analysis weights take into account specific subgroups that are over-sampled as well as post-stratification adjustments. Pooled across years, the NHANES samples have included 12666 individuals aged 20–59 years, all of whom answered cannabis smoking questions (1120 cases of diabetes mellitus).

1. *The National Surveys on Drug use and Health (*NSDUH) 2005-2012

Similar to NHANES, the NSDUH is designed to be nationally representative for the US non-institutionalized civilian population, and in 205-2012, a multi-stage area probability sampling approach was used. For comparability with NHANES, four NSDUH replication estimates can be derived by combining survey cycles as follows: 2005-6; 2007-8; 2009-10; 2011-12. Here also, analysis weights take into account specific subgroups that are over-sampled in order to increase precision of NSDUH estimates, as well as post-stratification adjustments. Pooled across years, the NSDUH samples have included 242250 adults aged 20-64 years, all of whom answered cannabis smoking questions (9553 cases of diabetes mellitus). Compared with subjects who were included in the final study sample, those who were excluded owing to missing covariate data were not different on cannabis smoking status, diabetes status, or other demographic, background, or lifestyle characteristics studied here.

**Response variable**

The key response variable is diabetes mellitus, which we presume to be type 2 DM in the CS-DM associations, based on NHANES findings, including post-CS onset for 93% of the DM cases (i.e., post-adolescent). We note this assumption in our coverage of study limitations, including the possibility that type 1 DM cases (and some type 2 DM cases) might have been told by their physicians to avoid cannabis smoking, and this might produce a spurious non-causal CS-DM association.

 In NSDUH, participants were asked if they “ever been told by a doctor or health professional that they have diabetes mellitus”. Participants were also asked about recent physician diagnosis through the question “In the past 12 months have you ever been told by a doctor or health professional that they have diabetes mellitus?”

In NHANES, participants were asked if they were ever told that they had diabetes. NHANES also has questions on current use of insulin and/or oral hypoglycemic medicines. In NHANES, diabetes mellitus can be assessed by measures of blood glycohemoglobin, fasting plasma glucose, and serum insulin. Glycohemoglobin measures were available for the full sample while measures of fasting plasma glucose and serum insulin were measured in the morning examination session only after a 9 hour fasting. A detailed description of the laboratory methods used can be found on the NHANES website (<http://wwwn.cdc.gov/nchs/nhanes/search/datapage.aspx?Component=Laboratory>).

For NHANES 2005-2006, glucose measurements were performed on the Roche/Hitachi 911 (Roche Diagnostics, 9115 Hague Road, Indianapolis, IN 46250.) using the hexokinase assay and for NHANES 2007-12 glucose measurements were performed on the Roche Modular P chemistry analyzer (Roche Diagnostics, 9115 Hague Road, Indianapolis, IN 46250.) using the hexokinase assay. The insulin measurement was performed 2005-2009 using the Mercodia sandwich ELISA assay and was switched in late 2009 to a Roche chemiluminescent immunoassay performed on the Elecsys 2010 analyzer. As recommended by the NHANES study and to trend insulin from 2011-2012 to match previous NHANES cycles, the following “backward” fractional polynomial regression was applied to 2011-2012 insulin values in the current analyses “Insulin (Mercodia-equivalent) = 0.6295 + [1.0770\*Insulin (Roche)] – [0.0008566\*Insulin (Roche)\*\*2]”. Glycohemoglobin measurements were performed on the A1c 2.2 Plus Glycohemoglobin Analyzer (Tosoh Medics, Inc., 347 Oyster Pt. Blvd., Suite 201, So. San Francisco, Ca 94080) in NHANES 2005-06. For NHANES 2007-12 glycohemoglobin measurements were performed on the A1c G7 HPLC Glycohemoglobin Analyzer (Tosoh Medics, Inc., 347 Oyster Pt. Blvd., Suite 201, So. San Francisco, Ca 94080).

**Exposure of interest**

Cannabis smoking was assessed for both NHANES and NSDUH via self-report responses to standardized questions during an Audio Computer Assisted Self-Interview (ACASI) drug use module. For NHANES, this assessment occurs within a private room of the mobile examination center at the time of the physical examination. For NSDUH, the assessment typically is in a private location within or sometimes near the participant's dwelling unit.

The ACASI approach is intended to promote accuracy and completeness of reporting on sensitive topics such as the lifetime history of cannabis smoking and how many days cannabis was used in the 30 days just prior to assessment. On this basis, participants can be classified as never smokers; former smokers (smoked cannabis at least once in lifetime but not in the 30 days prior to the interview); and recently active smokers (smoked cannabis at least once in the 30 days prior to the interview).

**Multiple Logistic Regression and Discrete time survival analysis**

The standard multiple logistic regression approach has been used to estimate the strength of association linking CS with DM, taking into account analysis weights, and with Taylor series linearization to estimate variances for complex sample survey designs of the type applied in NHANES and NSDUH. This primary analysis/estimation step leaves uncertainty about temporal sequencing: diabetes might precede cannabis smoking (as in the instance of type 1 DM cases, who might have been counseled to avoid cannabis smoking). Uncertainty can be removed somewhat via a post-estimation exploratory analysis that makes use of NHANES data on age of diabetes diagnosis and age of first cannabis smoking, which NHANES gathered for a subset of survey participants. These NHANES observed age-of onset values were used to construct discrete time survival analysis models (DTSA). In the first post-estimation exploratory DTSA, time to diabetes diagnosis has been modeled as function of the time-varying covariate term for cannabis smoking (coded 0 before cannabis onset age and then coded 1), using a now-standard DTSA approach for cross-sectional data.

For this approach, the dataset is reconstructed in a person-time format with the number of records equal to the participant's age at NHANES examination (right-censored), or equal to the age of first diabetes diagnosis. For example, a participant whose age of onset of diabetes was 40 had 40 records in the data set. Another participant who was interviewed at the age of 30 but never had developed diabetes had 30 records in the data set. Each record contained a binary dependent variable (1 if diabetes occurred and 0 if no diabetes), as well as a time indicator variable (t=1, 2, 3….). Additional variables included covariates such as sex, ethnic self-identification, education, income-poverty ratio, alcohol drinking as well as time-varying covariates to reflect onset of cannabis smoking and onset of tobacco cigarette smoking (‘0’ until first use, ‘1’ thereafter). Analysis weights and Taylor series methods were used for this estimation task as well.

In the second post-estimation DTSA, time to CS cessation was modeled as a function of age at DM diagnosis, with DM diagnosis-age coded as a time-varying covariate (‘0’ in the records for each case up to the age of DM diagnosis, thereafter ‘1’). As described in the text, this DTSA analysis suggested that DM diagnosis did not have a major influence on cessation of cannabis smoking or on whether a cannabis smoker persisted to the point of being a recently active CS on the date of NHANES assessment.

**Supplementary Figures and Tables**

These appendix figures and tables provide details mentioned in the text but not presented there due to limits on the numbers of figures and tables.

into account specific subgroups that are over-sampled in order to increase precision of NSDUH 5

estimates, as well as post-stratification adjustments. Pooled across years, the NSDUH samples

**eFigure 1**: Flow chart of the study population

 Data from the National Health and Nutrition Examination Survey and the United States National Surveys on Drug Use and Health, 2005–2012

NSDUH 2005-06

20-64 years of age (n=60199)

Missing information on key study variables n=870

NSDUH 2005-06 (n=59329)

Missing information on key study variables n=854

NSDUH 2007-08

20-64 years of age (n=60895)

NSDUH 2007-08 (n=60041)

NSDUH 2009-10

20-64 years of age (n=62101)

Missing information on key study variables n=828

NSDUH 2009-10 (n=61273)

NSDUH 2011-12

20-64 years of age (n=62390)

Missing information on key study variables n=783

NSDUH 2011-12 (n=61607)

NHANES 2011-12 (n=3040)

NHANES 2005-06

20-59 years of age (n=3285)

Missing information on key study variables n=369

NHANES 2005-06 (n=2916)

NHANES 2007-08

20-59 years of age (n=3662)

Missing information on key study variables n=417

NHANES 2007-08 (n=3245)

Missing information on key study variables n=589

NHANES 2009-10 (n=3465)

NHANES 2011-12

20-59 years of age (n=3632)

Missing information on key study variables n=592

NHANES 2009-10

20-59 years of age (n=4054)

|  |
| --- |
| **eTable 1**: Unweighted numbers of cannabis smoking and diabetes mellitus: Data from the National Health and Nutrition Examination Survey and the United States National Surveys on Drug Use and Health, 2005–2012 |
| Survey  | Sample size (n of diabetes cases)  | Never cannabis smokers (n of diabetes cases)  | Former cannabis smokers(n of diabetes cases)  | Recently active cannabis smokers(n of diabetes cases)  |
| *NHANES 2005-12*  | 12666 (1120)  | 5808 (612)  | 5252 (433) | 1606 (75)  |
| 2005-06  | 2916 (199)  | 1300 (112)  | 1291 (75)  | 325 (12)  |
| 2007-08  | 3245 (323)  | 1440 (179)  | 1416 (129)  | 389 (15)  |
| 2009-10  | 3465 (293)  | 1647 (155)  | 1354 (113)  | 464 (25)  |
| 2011-12  | 3040 (305)  | 1421 (166)  | 1191 (116)  | 428 (23)  |
| *NSDUH 2005-12*  | 242250 (9553)  | 112573 (5068)  | 103115 (4034)  | 26562 (451)  |
| 2005-06  | 59329 (2113)  | 27726 (1169)  | 25594 (851)  | 6009 (93)  |
| 2007-08  | 60041 (2343)  | 28266(1283)  | 25744 (957)  | 6031(103)  |
| 2009-10  | 61273 (2507)  | 28319 (1300)  | 25796 (1086)  | 7158 (121)  |
| 2011-12  | 61607 (2590)  | 28262 (1316)  | 25981 (1140)  | 7364 (134)  |



**Percentage (%)**

**eFigure2**: Distribution of individual age differences obtained via subtraction of the age of onset of cannabis smoking and age of onset of diabetes mellitus

Data for the US based on the National Health And Nutrition Examination Survey (2005-2012)

|  |
| --- |
| **eTable 2**: Post-estimation exploration of subgroup variation in the size of the association that links occurrence of diabetes mellitus with recently active cannabis smoking for major subgroups defined in terms of covariates for this study’s multiple logistic regression modelsa. Data from the National Health and Nutrition Examination Survey and the United States National Surveys on Drug Use and Health, 2005–2012. |
| **Subgroups** | NHANES covariate adjusted odds ratio (95%CI)b | NSDUH covariate adjusted odds ratio (95%CI)c | Covariate adjusted meta-analytic summary odds ratio (95%CI)d |
| ***Age***  |  |  |  |
| 20-49 | 0.6 (0.4, 0.8) | 0.6 (0.5, 0.7) | 0.6 (0.5, 0.8) |
| 50+ | 0.5 (0.3, 0.9) | 0.7 (0.5, 1.0) | 0.5 (0.3, 0.9) |
| ***Sex*** |  |  |  |
| Male | 0.5 (0.3, 0.8) | 0.6 (0.5, 0.8) | 0.6 (0.5, 0.7) |
| Female | 0.9 (0.6, 1.4) | 0.9 (0.7, 1.2) | 0.9 (0.7, 1.1) |
| ***Ethnic self-identification*** |  |  |  |
| Non-Hispanic White | 0.6 (0.4, 1.0) | 0.7 (0.6, 0.9) | 0.7 (0.6, 0.8) |
| Non-Hispanic Black | 0.6 (0.3, 1.0) | 0.7 (0.5, 1.0) | 0.6 (0.4, 1.0) |
| All others  | 0.7 (0.3, 1.7) | 0.9 (0.5, 1.4) | 0.7 (0.4, 1.1) |
| ***Cigarette smoking*** |  |  |  |
| Never  | 0.9 (0.5, 1.7) | 0.8 (0.4, 1.5) | 0.8 (0.5, 1.2) |
| Former  | 0.6 (0.3, 1.3) | 0.6 (0.4, 0.9) | 0.4 (0.4, 0.8) |
| Recently active  | 0.6 (0.4, 1.0) | 0.8 (0.7, 1.1) | 0.8 (0.6, 1.0) |
| ***Past year alcohol drinking***  |  |  |  |
| No  | 0.7 (0.3, 1.4) | 0.7 (0.5, 1.0) | 0.7 (0.3, 1.4) |
| Yes | 0.6 (0.4, 0.9) | 0.7 (0.6, 0.9) | 0.7 (0.5, 0.8) |
| ***BMI(Kg/m2)*** |  |  |  |
| <25 | 0.4 (0.1, 1.2) | ---------------- | 0.4 (0.1, 1.1) |
| 25-29.9 | 0.8 (0.4, 1.6) | --------------- | 0.9 (0.5, 1.6) |
| ≥30 | 0.8 (0.5, 1.2) | --------------- | 0.9 (0.6, 1.3) |
| a Each row presents an odds ratio estimate of the CS-DM association for the indicated subgroup. These are covariate-adjusted estimates from the multiple logistic regression models. To illustrate, the estimated OR for age is based on a model that excludes the variable for age (which is the basis for forming the stratified subgroups) but otherwise includes all of the other previously mentioned covariates: sex (male and female), ethnic self-identification (non-Hispanic Whites, non-Hispanic Blacks, Hispanics, and all others), education (less than high school, high school, and above high school), income-poverty ratio (<1 and ≥1), past-year alcohol drinking (never user, used before but not in the 12 months prior to the interview, and used in the 12 months prior to the interview) and tobacco cigarette smoking (never, former, non-daily smoker, and daily smoker). Likewise, the OR estimates for males and for females drop sex as a covariate, but include all of the other previously listed covariates. All of the subgroup-specific estimates were formed in this same fashion.b Estimates are from NHANES based on the merged years 2005-2012.c Estimates are from NSDUH based on the merged years 2005-2012.d Estimates are based on a meta-analysis summary approach treating each of the eight independent samples as a separate study. |

|  |
| --- |
| **eTable 3**: Association of cannabis smoking and markers of glucose metabolism among participants who did not report physician diagnosis of diabetes using linear regression models. Data from the National Health and Nutrition Examination Survey and the United States National Surveys on Drug Use and Health, 2005–2012. |
| **Markers of glucose metabolism** | Age-adjusted estimated difference (95% CI) | Covariatea-adjusted estimated difference (95% CI) |
| **HbA1c (%)** |  |  |
| *Never smokers* |  0.00 (referent) |  0.00 (referent) |
| *Former cannabis smokers* | -0.08 (-0.10, -0.05) | -0.05 (-0.08, -0.03) |
| *Current cannabis smokers* | -0.04 (-0.08, -0.01) | -0.07 (-0.11,-0.04) |
|  |  |  |
| **Fasting glucose (mmol/L)** |  |  |
| *Never smokers* |  0.00 (referent) |  0.00 (referent) |
| *Former cannabis smokers* | -0.08 (-0.14, -0.03) | -0.05 (-0.12, 0.00) |
| *Current cannabis smokers* | -0.01 (-0.08, 0.06) | -0.05 (-0.20, 0.09) |
|  |  |  |
| **Serum insulin (pmol/L)** |  |   |
| *Never smokers* |  0.00 (referent) |  0.00 (referent) |
| *Former cannabis smokers* | -8.56 (-14.08, -3.22) | -7.08 (-12.08, -2.07) |
| *Current cannabis smokers* | -17.52 (-22.85, -12.20) | -15.89 (-22.42, -9.36) |
|  |  |  |
| **Homeostatic Model Assessment-Insulin Resistance (HOMA-IR)** |  |  |
| *Never smokers* |  0.00 (referent) |  0.00 (referent) |
| *Former cannabis smokers* | -0.38 (-0.59, -0.17) | -0.31 (-0.55, -0.06) |
| *Current cannabis smokers* | -0.72 (-0.96, -0.49) | -0.64 (-0.93, -0.34) |
| a Covariate adjustments for age (years), sex (male and female), ethnic self-identification (non-Hispanic Whites, non-Hispanic Blacks, Hispanics, and all others), education (less than high school, high school, and above high school), income-poverty ratio (<1 and ≥1), past-year alcohol drinking (never user, used before but not in the 12 months prior to the interview, and used in the 12 months prior to the interview) and tobacco cigarette smoking (never, former, non-daily, and daily smoker), using linear regression model |

b HOMA-IR was calculated as ((Fasting insulin\*Fasting glucose/135).

|  |
| --- |
| **eTable 4**: Replication-specific estimates of odds ratios that quantify associations linking occurrence of diabetes mellitus with cannabis smoking status stratified in relation to past 30 days cannabis smoking frequency: Data for the United States based on eight independent replications from the National Health and Nutrition Examination Survey (NHANES) and the National Surveys on Drug Use and Health (NSDUH), 2005-2012. |
|  | Covariate-adjusteda odds ratio of diabetes mellitus (95% CI) |
| Independent replication sample | Never smokers | Former smokers  | Recently active smokers (1-7 days per month) | Recently active smokers (>7 days/month) |
| NHANES 2005-06 | 1 (referent) | 0.8 (0.5, 1.3) | 0.5 (0.3, 1.1) | 0.7 (0.3, 1.8) |
| NSDUH 2005-06 | 1 (referent) | 0.9 (0.8, 1.1) | 0.6 (0.3, 1.0) | 0.8 (0.5, 1.2) |
| NHANES 2007-08 | 1 (referent) | 0.6 (0.4, 1.0) | 0.4 (0.2, 0.9) | 0.3 (0.1, 1.0) |
| NSDUH 2007-08 | 1 (referent) | 0.9 (0.8, 1.1) | 0.6 (0.4, 1.0) | 0.5 (0.3, 0.7) |
| NHANES 2009-10 | 1 (referent) | 1.2 (0.9, 1.6) | 1.0 (0.4, 2.5) | 0.9 (0.3, 2.3) |
| NSDUH 2009-10 | 1 (referent) | 1.0 (0.8, 1.1) | 0.6 (0.3, 1.1) | 0.9 (0.5, 1.4) |
| NHANES 2011-12 | 1 (referent) | 1.0 (0.5, 1.7) | 0.7 (0.3, 1.5) | 0.8 (0.3, 1.9) |
| NSDUH 2011-12 | 1 (referent) | 1.1 (0.9, 1.2) | 1.0 (0.6, 1.5) | 0.9 (0.6, 1.3) |
|  |  |  |  |  |
| Meta-analytic summary odds ratiob | 1 (referent) | 0.9 (0.9, 1.0) | 0.7 (0.5, 0.8) | 0.7 (0.6, 0.9) |
| Heterogeneity test statistic(*p* value) | ------ | 9.3 (0.23) | 5.43 ( 0.61) | 7.34 (0.40) |
| a NHANES and NSDUH included standardized questions on days of cannabis smoking in the most recent CS month, which made this analysis possible. Covariate adjustments for age (years), sex (male and female), ethnic self-identification (non-Hispanic Whites, non-Hispanic Blacks, Hispanics, and all others), education (less than high school, high school, and above high school), income-poverty ratio (<1 and ≥1), past-year alcohol drinking (never user, used before but not in the 12 months prior to the interview, and used in the 12 months prior to the interview) and tobacco cigarette smoking (never, former, non-daily, and daily smoker), using the multiple logistic regression model.b The meta-analysis summary estimate makes an allowance for between-replication variability in the effect estimate. Some readers may wish to see 95% confidence intervals based on the fixed effects meta-analysis approach, given that there is no appreciable heterogeneity across replication samples (as disclosed in large p-values for the heterogeneity test statistics). For each meta-analysis odds ratio summary, the 'fixed effects' 95% CI are as follows: (1) former smokers (95% CI = 0.9, 1.0); for lower frequency recent CS (95% CI = 0.6 0.8); for higher frequency recent CS (95% CI = 0.6; 0.9). |