University of Kentucky

From the Selected Works of Olga A. Vsevolozhskaya

October 1, 2014

Confidence Interval Estimation in R-DAS

Olga A. Vsevolozhskaya, *Michigan State University*

James C. Anthony, *Michigan State University*

Available at: <https://works.bepress.com/vsevolozhskaya/4/>

Copyright © 2014 Published by Elsevier Ireland Ltd.

This manuscript version is made available under the CC-BY-NC-ND 4.0 license

<http://creativecommons.org/licenses/by-nc-nd/4.0/>

PLEASE DO NOT QUOTE OR CITE UNTIL PUBLISHED

Version July 25, 2014

Drug and Alcohol Dependence

Full-length Report¹

Confidence Interval Estimation in R-DAS.

Olga A. Vsevolozhskaya

James C. Anthony*

Michigan State University

Word Counts:

Abstract: 244

Body of Text: 6591

* Corresponding author: James C. Anthony, Professor of Epidemiology & Biostatistics, College of Human Medicine, Michigan State University, West Fee Hall, 909 Fee Road, East Lansing, MI 48824-1030, United States. Tel.: +1 517 353 8623; Fax: +1 517 432 1130.

Email address: janthony@msu.edu

¹ As required by journal instructions to authors, this footnote directs attention to Appendix A, which we provide as a separate document for a DAD online supplement to this article.

Abstract

Roughly 25 years ago, the United States National Institute on Drug Abuse (US NIDA) initiated the creation of public use datasets for its National Household Survey on Drug Abuse, since re-named the National Survey on Drug Use and Health (NSDUH). The Substance Abuse and Mental Health Services Administration (SAMHSA), which assumed responsibility for the survey in 1992, has continued and expanded this effort to make the survey data available to researchers. During 2012, SAMHSA created a “Restricted-Use Data Analysis System” (R-DAS) to provide researchers with the capability to create tabulations using restricted NSDUH variables not otherwise available on the public-use files. This methods focused article is intended to help potential users of R-DAS-like online data analysis systems by (i) clarifying statistical issues involving approximation of confidence intervals (CI), (ii) providing a way to estimate CI when tabular output is suppressed with an ‘error message’ based on confidentiality restrictions, and (iii) showing how to make pairwise comparisons of estimates not otherwise allowed. For illustration purposes, some empirical estimates are presented on a topic of continuing of public health concern in the US namely, extra-medical use of pain relievers (generally opioids), where the drugs are being used to get high and otherwise outside the boundaries intended by prescribing clinicians. The R-DAS makes it possible to derive state-level estimates of male-female and age-related differences in incidence of extra-medical prescription pain reliever (EMPPR) use, not previously reported in peer-reviewed articles, and not available without research approaches described here.

Keywords: Epidemiology; opiates; opioids; incidence; United States; misuse; pain relievers; R-DAS

1 Introduction

1.1 Background

In the background of this methodological article is a principle of data sharing expressed in policies of United States agencies such as the National Institute on Drug Abuse and the Substance Abuse and Mental Health Services Administration (US; NIDA; SAMHSA) when government funds have been used to support completion of public health research projects, particularly those on the scale of epidemiological field surveys. On occasion, public use files (PUF) are shared for flexible end-user analyses of individual level 'microdata,' often made available as downloadable "flat file" raw data or as corresponding analysis-ready datasets for commercial software packages (e.g., .xls format for Microsoft Excel; .dta format for Stata). In some instances, agencies have commissioned software programming required for "online analysis systems" so that potential users can log in to complete basic online analyses of the unweighted or weighted data. The "Restricted-use Data Analysis System" (R- DAS) and the "Survey Documentation and Analysis" (SDA) for the National Surveys on Drug Use and Health (NSDUH) in the United States are two online analysis systems that are most pertinent here, although many of this article's methodological topics pertain to PUF data analyses from complex epidemiological sample surveys generally.

As might be imagined, there is a substantial potential for erroneous results when PUF analyses are conducted by investigators who lack familiarity with statistical methods for epidemiological field survey data and who might not have seen excellent textbooks on applied survey analysis (e.g., [Heeringa et al., 2010](#)). One frequent source of error involves statistical inference requiring use of analysis weights – for example, as required when the study design has more complexity than can be achieved by simple random sampling from a census roster of all study population members. As an example, analysis weights may be required because the estimated population parameters need to be adjusted for a potential violation of an "independent and identically distributed" (i.i.d.) assumption. In epidemiological field surveys, the i.i.d. assumption can be violated and analysis weights may be needed because of stratification and cluster sampling with unequal probabilities of selection and for other reasons. Several analysis weights may be encountered. For example, there may be a weight based on the inverse of the probability of selection at one or more stages of sampling (which we can call an 'IPS' weight or a 'base weight'). Sometimes there is a weight

intended to make adjustments for missing data. In many surveys conducted since the 1940s, there is a 'poststratification adjustment factor' (PSAF) derived from survey research ideas introduced by [Deming et al., 1940](#), who noted how survey estimates might be improved, after data collection, via post-survey stratifications and adjustments to marginal totals known from a more accurate source (e.g., to accounting for missing data, variations survey participation levels, and other aspects of survey design).

To illustrate the idea of an 'IPS' or 'base' weight, consider a hypothetical simple one-stage survey set up such that all census-derived community dwelling units (DU) are included, but only one DU occupant is sampled for assessment. In a DU with one occupant, the selection probability is 1.0; the weight is the inverse of $1.0 = 1.0$. In a DU with just two occupants, selection probability is 50%; IPS or base weight is $1/(0.5)$. For DU with three occupants, this weight is the inverse of $1/3$, and so on.

In contrast to what a research team might specify as IPS or base weights in advance of the study, the PSAF is created after survey fieldwork and a comparison of what is thought to be known about the study population (e.g., based on a 100% census sample). For example, the most recent decennial census of the non-institutionalized non-military US population age 12 years and older yields a male-female ratio of 151,781,326 to 156,964,212 ([United States Census Bureau, 2011](#)). Because not all eligible women participate in a survey (due to sample frame non-coverage or omission), when the survey-based number of females is compared to the known census count of women in a stratum of interest, the survey-based count often falls short of what the census states. Also, because sampled women often participate in surveys more than sampled men, the survey-based count of men can show a greater shortfall when compared to known census counts. A post-survey stratification analysis displays this type of survey error in its comparison of the survey-based estimates for each population stratum relative to the known census population count. A PSAF involves correction multiplier to bring survey counts up to census when a goal is to produce estimates for the total study population. Because this type of survey error often varies across cells in a multiway contingency table formed by census variables like sex, age, and race-ethnicity, many surveys create a PSAF as a multiplier for each cell of that multiway table.

Investigators unfamiliar with applied survey analysis principles may not appreciate that the PSAF multiplier can be made to serve multiple valuable functions, but it is not uniformly required

nor is it always useful. For example, the PSAF actually can produce misleading study estimates when the estimation task is focused on subgroups within a population, and may be based on a fallible assumption about within-subgroup homogeneity. In contrast, when the estimation task is to produce estimated counts for the entire population, the use of a PSAF can be essential. For these reasons, 'best practices' for an optimal PUF for data sharing will include creation of separate variables and codebook documentation for the survey's IPS weight or weights as they might vary across one or more stages of sampling and for any PSAF created by the survey team, as well as a summary analysis weight that combines the IPS weight with the PSA factor, along with any other adjustments as judged to be useful by the PUF creators.

At present, the SAMHSA R-DAS and SDA provide only summary analysis weights. More precisely, the R-DAS provides a set of summary analysis weights that can be chosen by the end-user based on decisions about which single summary estimate is most useful. For example, for a relatively rare drug using behavior (e.g., injecting methamphetamine), a single estimate based a summary analysis weight for 8- or 10-years of NSDUH data might be most useful. In contrast, recent summary estimates have been used to study US trends in incidence of crack-cocaine use, with a display of one trend line based on 1-year SDA summary analysis weights, and with another trend line based on 2-year R-DAS summary analysis weights ([Parker and Anthony, 2014](#)).

It also is noteworthy that the SDA gives investigators access to downloadable unweighted datasets and the corresponding summary analysis weight, allowing direct comparison of unweighted and weighted estimates. In contrast, the R-DAS is strictly for online tabular analyses; it yields weighted analyses with no straightforward way to know the actual unweighted number of study participants in the analysis sample. In circumstances of this type, one might ask why the R-DAS was created, and why investigators would choose to use R-DAS versus SDA. The answer rests on a consideration of survey years, as well as an issue of accessibility to NSDUH study participants and NSDUH variables.

The SDA datasets are from national field surveys of drug use conducted since the late 1970s, whereas the R-DAS are restricted to national surveys conducted only since 2002. For both SDA and R-DAS, confidentiality and privacy protections required removal of some participants from the survey datasets, but the participants removed from the SDA datasets are not always the same individuals who have been removed from the R-DAS datasets.

With respect to variables available for analysis, the R-DAS and the SDA portals give access to different selections. To illustrate, investigators interested in monthly or seasonal variations in newly incident drug use must turn to the SDA, whereas investigators interested in state or sub-state geographic variations must turn to the R-DAS. R-DAS includes variables for each US state, but no month/season variable; SDA includes month variables, but includes no state variable.

SDA estimates from each year can be combined in meta-analysis fashion, with each year’s survey treated as an independent replication (DeAndrea et al., 2013). In some contexts, the SDA meta-analysis summary estimate has been found to perform as well as or better than an alternative approach that involves pooling data across years (DeAndrea et al., 2013). Similarly, a series of R-DAS 2-year estimates can be summarized in meta-analytic fashion (Mohammed and Anthony, 2014), and the R-DAS 8-year and 10-year summary analysis weights are available when the goal is to optimize estimates based across-year pooling of R-DAS data (e.g., see Seedall and Anthony, 2013).

Indubitably valuable, the R-DAS remains a “work in progress” for which there is no extensive online documentation. Unlike SDA, with considerable online help support (Online Help for Analysis Programs-SDA 3.5) and a NSDUH statistical inference report that provides details about a NSDUH statistical inference procedures (Aldworth et al., 2011), the FAQs page for R-DAS is more constrained (Help with the Restricted-use Data Analysis System (R-DAS)). For instance, R-DAS documentation does not state whether the 95% default confidence intervals for estimated percentages are produced using normal approximations to the binomial distribution (i.e., yielding standard symmetric CIs), or by converting a standard error and the estimate of each percentage to a natural logarithm scale (i.e., the logit transformation). Additionally, if the unweighted number of observations in any cell in the R-DAS table falls below some unstated threshold value, R-DAS delivers this system response: “**Error** Your analysis has returned one or more errors: To preserve confidentiality, tables cannot be displayed when the number of observations in any cell in the table is too low.” That is, the output is suppressed due to confidentiality restrictions. Also, certain statistical procedures are not available with the current version of R-DAS. For example, R-DAS includes no approach for significance testing or derivation of p-values when the goal is statistical inference about pairwise comparisons of percentages.

Finally, via anonymous reviews of this article, we were encouraged to mention a new NSDUH

“Data Portal” and an arrangement that allows investigators to apply for access to microdata files with fewer restrictions on variable selection. We have not yet been approved for this privilege, and cannot speak from experience, but it is possible that the value of approaches described in this article will be greater for R-DAS users than for users of the new data portal.

1.2 This Paper’s Focus

With a methods focus, this paper seeks to fill some gaps in R-DAS documentation based on our own end-user experiences, and to provide a simple alternative to current R-DAS research approaches, with an illustration that sheds light on several facets of the epidemiology of extra-medical prescription pain reliever use in the US. Using R-DAS data to look state by state within the US, we estimate (1) male-female differences and (2) age-related differences in the risk of becoming a newly incident user of these drug compounds. This illustration clarifies state-by-state variations in incidence of becoming a newly incident extra-medical user of prescription pain relievers (mainly opioid drugs). Here, our measure of incidence has a numerator based on the estimated number of newly incident users during a specified time interval, and its denominator is based on the estimated size of the ‘at risk’ population. That is, this denominator excludes individuals who had started extra-medical use in the past, and includes only those individuals ‘at risk’ of starting for the first time. Consistent with prior National Comorbidity Survey articles (e.g. [Anthony et al., 1994](#)), and many other more recent articles based on NSDUH data (e.g., [Meier et al., 2012](#); [DeAndrea et al., 2013](#); [Seedall and Anthony, 2013](#)), the term “extra-medical” refers to using the drug outside of boundaries intended by the prescribing clinician (e.g., “to get high” or for other such feelings).

With this methods focus, this article has been written to help potential users of R-DAS and other online data analysis systems by (i) clarifying statistical issues involving approximation of confidence intervals (CI), (ii) providing a way to estimate CI when tabular output is suppressed with an ‘error message’ based on an R-DAS confidentiality restriction, and (iii) showing how to make pairwise comparisons of estimates not otherwise allowed in R-DAS. The specific substantive aims pertain to state-specific incidence estimates just mentioned, including estimates for middle-aged and elderly US residents that otherwise are suppressed with a R-DAS confidentiality error message. We anticipate that many DAD readers will be interested in how this particular obstacle can be overcome.

Due to the methods focus of this article, we do not provide a detailed overview of prior important contributions in this substantive research domain. Fortunately, in addition to the background literature we have summarized in our own research group’s past contributions of epidemiological evidence on prescription pain relievers, as well as male-female and age-related differences in incidence for extra-medical drug use (e.g., [Meier et al., 2012](#); [DeAndrea et al., 2013](#); [Seedall and Anthony, 2013](#)), there are new reviews from other research teams (e.g., [Dowling et al., 2006](#); [Martins et al., 2009](#); [Manchikanti et al., 2012](#); [Volkow et al., 2014](#)). In addition to US estimates, there is a growing body of contributions, including Germany, the United Kingdom, Oceania, and Asia ([Larance et al., 2011](#); [Roxburgh et al., 2011](#); [Schubert et al., 2013](#); [Radbruch et al., 2013](#); [Hawkes, 2013](#); [Nicholas et al., 2013](#); [Degenhardt et al., 2014](#)).

2 Materials and Methods

2.1 The Epidemiological Field Surveys

The NSDUH have been described in many past articles of this journal, such as those cited above. There also is extensive online documentation about NSDUH study methods and statistical inference approaches (e.g., [Aldworth et al. 2011](#)). In brief, the NSDUH is conducted annually with field operations that involve multistage area sampling and recruitment. As for materials for this study, each year from 2002 through 2011, the yield has been large scale nationally representative samples of community-dwelling (non-institutionalized) U.S. citizens, age 12 and older, such that this survey no longer qualifies as a “household survey” per se, even though it recently has been characterized as such (<https://apha.confex.com/apha/141am/webprogram/Paper280610.html>, last accessed 9 May 2014). NSDUH participation levels at 70% or better generally have been achieved, and confidential audio computer-assisted self-interviewing has been used for data gathering. Examples of standardized survey items are available online (<http://oas.samhsa.gov/nsduh/methods.cfm>, last accessed 9 May 2014). Various government reports provide basic estimates for the US population as a whole and subgroups judged to be especially important (e.g., overall male-female differences of the type reported here). These online reports provide details on potential limitations, as well as a comparison with alternative field survey estimates such as those from the Monitoring the Future school surveys (e.g., <http://www.samhsa.gov/data/nsduh/2k11results/nsduhresults2011.htm>)

In our illustration, we describe male-female variation in incidents estimates for EMPPR across 51 state-level jurisdictions (including the District of Columbia). Here, estimation is based entirely on the Restricted-use Data Analysis System of the pooled 2002-2009 NSDUH survey data; the 10-year analysis weights for 2002-2011 data were not yet available at the time this project was being conducted. R-DAS involves use of correct analysis weights for the pooled incidence estimates, with corresponding 8-year post-stratification adjustment factors built into the weights as described in Section 1.1. Here, R-DAS applies Taylor series linearization for estimation of variances under conditions of complex survey sampling. Once the correct pooled estimates are produced by R-DAS, various methods to summarize and compare incidence estimates may be employed, as illustrates below.

2.2 Confidence Intervals for Incidence Rates

In general epidemiology, an incidence rate for a disease in a specified population is defined as the ratio of the estimated number of newly incident cases arising during an interval of time to some measure of total or at-risk population size (or person-time experience).¹ R-DAS makes it possible to estimate the number of newly incident extra-medical drug users in a defined time interval, divided by the estimated size of the at-risk population under study, with exclusions and inclusions already specified in Section 1.2.² Specofoc R-DAS constructed variable specifications for this article’s estimates are presented in online supplementary materials that accompany this article. In many cases, the R-DAS-derived incidence estimates are small proportions (i.e., close to 0). For these small estimates, the standard and correctly implemented procedures for approximating 95% confidence intervals may yield a negative sign on the lower confidence bound. This negative value clearly falls outside a proportion’s allowable range of (0,1). In such cases, a logit transformation of estimated incidence rates provides mathematically acceptable results. The logit transformation

¹In our NSDUH incidence estimates, we form a numerator based on the number of newly incident users during a specified time interval, and we form a denominator based on the number of individuals who had not tried the drug prior to the start of that interval (as best these quantities can be measured in NSDUH field surveys). That is, the denominator is based on an ‘at risk’ specification as opposed to a ‘total population’ specification. DAD readers interested in alternative specifications for incidence estimation may find useful information at this US Centers for Disease Control and Prevention web site: <http://www.cdc.gov/ophs/csels/dsepd/SS1978/Lesson3/Section2.html>

²Example of generic R-DAS table specification can be found in Supplementary Section A.3.

of the incidence rate, p , is defined as

$$L = \ln \frac{p}{1-p},$$

where $\ln$ denotes a natural logarithm. If $\hat{p}$ is an estimate of the incidence rate and $\hat{L} = \ln \frac{\hat{p}}{1-\hat{p}}$ is the corresponding logit transformation, then an approximate $1 - \alpha$ confidence interval for the incidence rate, p , is given by

$$\left(1 + e^{-\hat{L} + z_{\alpha/2} \frac{\sqrt{Var(\hat{p})}}{\hat{p}(1-\hat{p})}}\right)^{-1} \leq p \leq \left(1 + e^{-\hat{L} - z_{\alpha/2} \frac{\sqrt{Var(\hat{p})}}{\hat{p}(1-\hat{p})}}\right)^{-1}, \quad (1)$$

where $Var(\hat{p})$ is the variance estimate of $\hat{p}$ and $z_{\alpha/2}$ is $(1 - \alpha/2)$ percentile of the standard normal distribution (i.e., $N(0,1)$). Via correspondence with the NSDUH R-DAS support group, we have learned that R-DAS uses the logit transformation to yield asymmetric confidence bounds that always fall within the $(0,1)$ interval.

2.3 Statistical Significance of Differences

Customarily, the observed difference between estimates is evaluated in terms of a test statistic, with a p-value to help quantify departures from a null hypothesis of no difference. A test of the hypothesis $H_0 : p_1 = p_2$ versus a two-sided alternative $H_A : p_1 \neq p_2$ (i.e., a test for a statistically significant difference between two incidence estimates) currently is not implemented in R-DAS. Nonetheless, the contrast can be made with ease “by hand”. Under the null hypothesis, a test can be based on the statistic

$$T = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{Var(\hat{p}_1) + Var(\hat{p}_2) - 2Cov(\hat{p}_1, \hat{p}_2)}}. \quad (2)$$

where $Cov(\cdot, \cdot)$ is a covariance term. Although R-DAS is not set up to provide $Cov(\hat{p}_1, \hat{p}_2)$, we will work from an assumption that the covariance term is not appreciably different from zero, and generally can be ignored in approximation of the test statistic (Aldworth et al. (2011)). We return to this issue in our discussion section.

2.3.1 Control of Error When Multiple Tests Are Performed

Whenever multiple inter-related tests of statistical significance are performed (e.g., testing for male-female differences for each of 51 state-level jurisdictions), one needs to be cautious in presentation of the results of the analysis due to the inflated probability of at least one false positive result (i.e., erroneous conclusion of a statistically significant non-null difference even when the null is true). Elevated probability of at least one false positive result has been called a “multiple-testing problem” – the more tests performed, the more false positive results to be expected, as shown in numerical examples provided elsewhere (e.g., [Rice, 1988](#)). Some type of correction for multiple testing may be needed. For example, in application of a commonly used Bonferroni correction, a hypothesis is rejected if the corresponding p-value is below the statistical threshold, α , divided by the number of tests. Interested readers can judge when and whether these corrections for multiple testing are needed, whether the Bonferroni correction is optimal, or whether alternatives to the Bonferroni approach might be useful ([Cribbie, 2007](#)).

2.4 Estimation of Suppressed Output

We have discovered that it often is the case that R-DAS estimates for extra-medical drug use can be produced for the entire population of interest (e.g., males and females as an overall population of interest), and for sub-populations that include relatively large unweighted numerators and denominators (e.g., 12-44 year olds). Nevertheless, drug use estimates for other sub-populations often are suppressed due to the small number confidentiality/privacy issue mentioned in Section 1.1 (e.g., when there is a small sample size for EMPPR users age 65 years and older). In this instance, when sub-populations can be defined to be mutually exclusive, it becomes possible for investigators to estimate the suppressed output “by hand” even when R-DAS will not produce the complementary tables.

To illustrate using an example of two mutually exclusive sub-populations, we let $\hat{p}$ and $\hat{N}$ denote the estimated percentage and the total weighted population size. Let $\hat{p}_1$ and $\hat{N}_1$ denote the estimated percentage and the sub-population weighted size for which the R-DAS delivers a table of estimates, and let $\hat{p}_2$ denote the estimated percentage from the sub-population with weighted $\hat{N}_2$,

for which the R-DAS output is suppressed. We can write $\hat{p}$ as a weighted sum of $\hat{p}_1$ and $\hat{p}_2$,

$$\hat{p} = \frac{\hat{N}_1}{\hat{N}}\hat{p}_1 + \frac{\hat{N} - \hat{N}_1}{\hat{N}}\hat{p}_2.$$

Then,

$$\begin{aligned} E(\hat{p}) &= \frac{\hat{N}_1}{\hat{N}}E(\hat{p}_1) + \frac{\hat{N} - \hat{N}_1}{\hat{N}}E(\hat{p}_2), \\ Var(\hat{p}) &= \left(\frac{\hat{N}_1}{\hat{N}}\right)^2 Var(\hat{p}_1) + \left(\frac{\hat{N} - \hat{N}_1}{\hat{N}}\right)^2 Var(\hat{p}_2) + 2\left(\frac{\hat{N}_1}{\hat{N}}\right)\left(\frac{\hat{N} - \hat{N}_1}{\hat{N}}\right)Cov(\hat{p}_1, \hat{p}_2), \end{aligned} \quad (3)$$

where E is the expected value, Var is the variance, and Cov is covariance. Therefore,

$$\begin{aligned} E(\hat{p}_2) &= \frac{\hat{N}}{\hat{N} - \hat{N}_1}E(\hat{p}) - \frac{\hat{N}_1}{\hat{N} - \hat{N}_1}E(\hat{p}_1), \\ Var(\hat{p}_2) &= \left(\frac{\hat{N}}{\hat{N} - \hat{N}_1}\right)^2 Var(\hat{p}) - \left(\frac{\hat{N}_1}{\hat{N} - \hat{N}_1}\right)^2 Var(\hat{p}_1) - 2\left(\frac{\hat{N}_1}{\hat{N}}\right)\left(\frac{\hat{N} - \hat{N}_1}{\hat{N}}\right)Cov(\hat{p}_1, \hat{p}_2). \end{aligned} \quad (4)$$

In practice, an R-DAS analyst generally must presume a zero covariance term as discussed in Section 2.3. In Eq.(4), this zero covariance assumption may introduce some bias into the variance estimation process of the suppressed output. In Section 3.3 we show that this bias tends to be minimal in this particular NSDUH estimation context.³

3 Results

The results section has three parts: (3.1) Data Preparation, (3.2) Incidence Estimation for Males and Females, By State, and (3.3) State-by-State Variation in Age-Specific Incidence Estimates.

3.1 Data Preparation

The illustration of these statistical methods in this article involves estimation of incidence rates for extra-medical use of prescription-type pain relievers (mainly opioids) in the US, as an extension of prior reports (see Section 1.2). Via R-DAS, novel estimates for 2002 to 2009 are used to exam-

³See Supplementary Section A.4 for the additional discussion of zero covariance term assumption.

ine male-female differences by geographical location (state) and by age. At present, R-DAS does not provide a convenient way to transfer its output to a document or a spreadsheet. A current online recommendation (provided in [Help with the Restricted-use Data Analysis System \(R-DAS\)](#), 2013) is to copy and paste the output, but this output can not be read into a statistical package for an immediate analysis. An alternative might be to enter the output manually, but this alternative is time consuming and creates opening for inadvertent transcription errors. We produce a methodological solution to this problem, with elimination of transcription errors, via a python script ([Python Software Foundation](#), 2013) for extraction of all relevant information from the R-DAS tabular output. Specifically, for each table, the script extracts all control variable names, as well as percentages, standard errors, and weighted sample sizes (except totals). We have shared a sample script in online document <http://www.epi.msu.edu/vsevoloz/scripts/>, also available by request to O.A.V. The script can be modified quite readily to accommodate extraction of different information (e.g., R-DAS-approximated confidence bounds). The authors welcome requests for help with such modifications.

3.2 Incidence Rate Estimation for Males & Females, By State

Combined 8-year R-DAS incidence estimates for EMPPR use, estimated for males and for females separately, in each of 51 state-level jurisdictions (hereinafter, “state”) are summarized in Figures 1 and 2.⁴ As noted in Section 2.1, within these figures, the 48 “lower” states and the District of Columbia are shaded to reflect the size of each incidence estimate. Since estimated standard errors for these estimates are quite small relative to the incidence estimates themselves, shading based just on the estimates provides an adequate representation of state-level variation in the risk of becoming a newly incident EMPPR user.⁵

As shown in Figures 1 and 2, the state of Utah has top-ranked empirical estimates for becoming a newly incident EMPPR user, and this is seen for males as well as for females. For males, the estimate is 1.5% (95% CI = 1.1, 2.2); females it is 1.8% (95% CI = 1.3, 2.5). It is not our intention to claim statistically significant results for Utah. Moreover, in the current application, a rigorous

⁴R-DAS tables based on which Figures 1 and 2 were generated can be found http://www.epi.msu.edu/vsevoloz/scripts/RDAS/rdas_incidents_gender.html

⁵Complete list of numerical estimates, including Alaska and Hawaii, is presented in Supplementary Table 1 and shows the very narrow 95% CI bounds.

statistical analysis would be hard to perform due to multiplicity concerns (i.e., adjusting for 51 pair-wise comparisons), as well as correlation among incidence rate estimates in nearby states. However, we still think that it is epidemiologically noteworthy that Utah has come up as one of the top-ranked states for both males and females. Additionally, estimates for Wisconsin stood out with high estimated rates for both females (1.5%; CI = 1.0, 2.1) and males (1.4%; CI = 1.0, 1.8). For many other states, a substantial male-female difference in estimated incidence rates can be seen.

To see whether females or males are at greater risk of becoming newly incident users at the state level, we conducted a statistical significance test (level of significance $\alpha = 0.05\%$) for the difference in R-DAS incidence estimates based on Eq. (2). The results of our findings (with no adjustment for multiple testing and zero covariance assumption) are summarized in Fig. 3; states with statistically significant male-female differences ($p < 0.05$) are highlighted in color. Additionally, in Fig. 4, we provide 95% confidence intervals for the differences, as calculated by inverting the test statistic in Eq. (2) (Casella and Berger, 2002).

For all highlighted states, female incidence estimates were larger than corresponding male estimates. We note that if these state-specific estimated differences are regarded as exchangeable replications within a Bonferroni family of estimates, one might be led to the conclusion that females are not more likely than males to become newly incident extra-medical users of these pain relievers; none of the state-level p-values passed the Bonferroni corrected threshold of $5\% \times 51$.

Nonetheless, despite the fact that none of the state-level p-values passed the Bonferroni corrected thresholds, we think it should be argued that the Bonferroni correction is too stringent in this instance, and that our results support general pattern of female excess in the risk of becoming an extra-medical user of these drugs, with a difference of these sex-specific risk estimates falling above the null value of 0.0 for 31 states and with only 18 states having point estimates below the null value (i.e., 2:1 odds of female excess risk overall).⁶

Our judgment that the Bonferroni correction is too stringent in the current analysis is based on an understanding of the Bonferroni correction as one that has been designed to control the probability of at least one false finding among independent tests within a family of exchangeable tests. Accordingly, whenever tests are interdependent, the Bonferroni correction becomes too stringent.

⁶See Supplementary Table 1 for the complete list of female and male incidence estimates; our discussion provides evidence of an overall female excess based on Fisher combined testing.

Moreover, for the extreme opposite case of perfectly correlated tests, the correction is not required. We do not judge that these state-specific comparisons of males and females are truly independent. Rather, we judge that it is reasonable to assume that there is some spatial correlation in incidence estimates among nearby states (Barrios et al., 2012), given considerations such as interstate drug trafficking patterns and shared methods covariation from several potential sources. Examples of ‘shared methods covariation’ involve the possibility that NSDUH uses the same field staff in adjacent states; or that there is across-state sharing of respondent tendencies to disclose or not disclose illegal or sensitive aspects of behavior.⁷

3.3 State-by-State Variation in Age-Specific Incidence Estimates

It was not difficult to identify an illustration of potential substantive importance so as to foster the reader’s understanding of an R-DAS ‘error’ message that is generated when tabular output has been suppressed due to a confidentiality restriction. For this illustration, we turned to an issue of whether prevention programs for EMPPR users might be needed among older persons, or whether these prevention programs might be focused solely upon young persons, a topic addressed by Colliver and colleagues (Colliver et al., 2006). Fairly regular R-DAS system response error messages were generated when we tried to estimate state-specific incidence of EMPPR use among 45-106 year olds, which might lead an epidemiologist to think that there is good reason to believe that older US community residents are not becoming newly incident EMPPR users.⁸ That is, the epidemiologist might reason that the size of the ‘at-risk’ study population of 46-106 year olds is quite massive, and therefore the confidentiality restriction must be due to the trivial size of each estimate’s numerator (translated as “too few” newly incident users in that age range, so few that no R-DAS estimate can be produced).

In order to illustrate the utility of Eq. (4) as described in Section 2.4, we constructed R-DAS analyses to derive the estimates for older adults and their standard errors. That is, the approach here involved taking differences between the total population estimates and the estimates for the 12-44 year olds, as described in Section 2.4.

⁷Section A.2 of online Supplemental Material provides mathematical clarification of spatial correlation issues in relation to possibly overly conservative Bonferroni correction.

⁸R-DAS screenshots with our specifications for the numerator and the denominator for the age-stratified incidence estimation can be found in Supplementary Section A.3.

Figures 5 and 6 illustrate incidence estimates for the 12-44-year-olds and for the older adults.⁹ Among 12-44-year-olds, incidence estimates are noticeably larger than corresponding estimates for older adults. In addition, there are states with non-zero estimated values for older adults. That is, the estimates show non-zero risk of becoming a newly incident EMPPR user beyond age 44 years.

A close look at Fig. 6 shows that Tennessee and Minnesota have a markedly deeper shade of color as compared to other states (i.e., larger incidence estimates), but we do not claim that these two states are significantly different from other states (i.e., at $p < 0.05$). Unlike percentages in Figures 1-5, for which the corresponding standard errors are trivial, the estimated percentages in Fig. 6 are small, but have corresponding standard errors that are relatively large (traceable to the relatively small numerators for newly incident users among the older adults).

Overall, a close inspection of Figures 5 and 6 discloses that some newly incident cases of EMPPR use in the US are arising after age 44 years. Nevertheless, we note that the main purpose of this analysis was to illustrate that, in certain situations, the R-DAS output suppressed due to confidentiality concerns can be re-estimated “by hand” using Eq. (4) with 0.0 for its covariance term, $Cov(\hat{p}_1, \hat{p}_2)$.

To assess an aspect of potential bias of Eq. (4) estimators, we compared our results to R-DAS output available for large states that did not require Eq (4). Some bias might be expected because a reduced number of significant digits displayed in R-DAS tables might lead to some rounding error. Table 1 shows the estimated and the observed R-DAS output for these larger states. It may be somewhat reassuring that the observed bias in the incidence estimates, defined as a difference between the observed and the estimated values, is negligible (on the order of a thousandth of 1%). The apparent bias among standard errors is larger but still less than tenth of a percent. This additional bias may be due to a non-zero covariance term involving the two proportions being compared. Nevertheless, based on the results in Table 1, the additional bias introduced by a non-zero covariance term might often turn out to be quite small. We conclude that the estimation procedure presented in Eq. (4) should work reasonably well in many contexts of estimation and inference from R-DAS output.

⁹Additional numerical estimates for each state and age subgroup are provided in Supplementary Table 2.

Table 1:
Bias assessment in re-estimation of the suppressed R-DAS output.

State	Incidence Estimate (%)			Standard Errors (%)		
	Estimated	Observed	Bias	Estimated	Observed	Bias
California	0.122	0.122	0.000	0.058	0.054	0.004
Colorado	0.343	0.343	0.000	0.249	0.220	0.029
Florida	0.318	0.318	0.000	0.108	0.097	0.011
Illinois	0.088	0.087	0.001	0.049	0.044	0.005
Michigan	0.168	0.169	0.001	0.070	0.065	0.005
Minnesota	0.443	0.441	0.002	0.250	0.209	0.042
New Mexico	0.222	0.220	0.002	0.153	0.130	0.023
New York	0.133	0.132	0.001	0.061	0.056	0.005
Oklahoma	0.344	0.344	0.000	0.231	0.238	-0.007
Pennsylvania	0.097	0.097	0.000	0.070	0.063	0.007
Tennessee	0.777	0.777	0.000	0.304	0.338	-0.033
Texas	0.199	0.199	0.000	0.095	0.088	0.007

4 Discussion

The main focus of this article is methodological. In this article, we have shared some research ideas and approaches stimulated by our initial experiences in use of the SAMHSA R-DAS for estimation and inference. That is, once we had devised this set of research approaches, we had to decide whether to keep the approaches to ourselves, or to share them with others who might find them to be useful. We chose to share what we had learned and worked out, in hope that we can create a community of scientists whose use of R-DAS will encourage the SAMHSA tradition of innovation in data sharing. In time, we hope that SAMHSA will commission refinements of the R-DAS as it has done with its SDA system, including options for use of the generalized linear model with link functions when R-DAS is tapped for more advanced epidemiological analyses.

Before any discussion of new substantive findings from this work, several study limitations deserve attention. For example, in contrast to medical examiner toxicological reports on overdose deaths, the NSDUH estimates are based on self-report data. When the goal is to study newly incident EMPPR use in national-scale surveys, there seems to be no logistically feasible alternative to the self-report. Moreover, we acknowledge that there might be opioids-related sample attrition. For example, as for any individual who engages in extra-medical use of an opioid pain reliever for the first time and who dies by overdose on that first occasion of use, this experience cannot be captured in any cross-sectional national scale survey of this type.

In Section 2.3’s coverage of assessing statistical significance of observed differences in R-DAS estimates, we drew attention to the covariance term of the standard equation. We noted an assumption – namely, that for a great many R-DAS difference tests, if the subpopulations are defined as mutually exclusive and truly are independent, this covariance term should be equal to zero or negligibly different from zero. Although we do not now have access to NSDUH data that would allow us to probe for violations of this assumption, the SAMHSA research team and the R-DAS Help Desk staff do have access to all of the data required to investigate this important topic in a systematic fashion. We hope that this article will stimulate them to provide us with the data required to probe for violations of the zero covariance assumption, or to complete and publish a report of their own analyses on this topic.

Notwithstanding limitations such as these, the article draws attention to (1) the observed state-level variations in the incidence of extra-medical use of prescription pain relievers, as well as some state-level male-female differences that are consistent with recently published estimates for young people in the US as a whole (Seedall and Anthony, 2013) and (2) the observation that in the United States there is evidence that newly incident EMPPR use is occurring after age 45. Given this article’s methods focus, our discussion of these observations necessarily is constrained. Readers are referred to other sources on the topic of male-female differences and age-related differences (e.g., Simoni-Wastila et al., 2004; Kelly et al., 2008; Seedall and Anthony, 2013 Colliver et al., 2006)

It might be argued that in this study’s estimates the preponderance of evidence actually is balanced against the idea of a male-female difference. Among 51 jurisdictions for which we have produced estimates, there was a statistically significant ($p < 0.05$) female excess only in Nebraska, South Carolina, New Hampshire, Oklahoma, and New Mexico. Nonetheless, in order to substantiate our summary statements about the observed overall female excess, we employed the classical Fisher test (Fisher, 1925), by combining the one-sided p-values (i.e., $H_a : p_F > p_M$) over 51 states. The resulting p-value for the overall national over-representation of females was very small (p-value = 3.49×10^{-7}). Thus, we conclude that during the years under study the women in the US were over-represented among newly incident EMPPR users. That is, rather than dismiss these state-level differences via a Bonferroni correction, future investigators can ask whether there might be other evidence consistent with the observed female excess risk of starting EMPPR use.

In light of this study’s observation that Utah’s incidence estimates are in the top rank position

among states, we draw attention to some separate epidemiological observations of note. Namely, [McDonald et al., 2012](#) also provided state-level estimates of mean milligrams of opioids (in morphine equivalents) dispensed per resident per state in 2008, and drew attention to the state of Utah, where there was an estimated average of 672 milligrams per resident (roughly 120-130 milligrams above the national average). Furthermore, in 2007, Utah was found to have had more deaths due to opioids overdose than it had motor vehicle crash deaths ([Utah Department of Health Violence & Injury Prevention Program, 2011](#)).

As for the observations about newly incident EMPPR use in middle and later adulthood, we note that Colliver and colleagues ([Colliver et al., 2006](#)) studied prevalence of EMPPR use among US residents age 50 years and older, and made projections for the future, discovering that the number of prevalent users might well double by the year 2020. These prior findings were in the background of our decision to contrast the adult population to older adults (age 45-106 years) and our discovery that there are some older adults who are becoming EMPPR users in later life, even though original R-DAS tabular output would suggest that the numbers are too small for estimation at the state level. Accordingly, these estimates offer an epidemiological rationale to think through prevention of EMPPR use in each state's older adult population, although the incidence estimates for these older adults do tend to be smaller than incidence estimates for 12-44 year olds. Indeed, some survey statisticians might insist that at least a few of our state-level estimates for 45-106 year olds have minimal scientific utility, in that the relative size of the standard errors to the point estimates sometimes is less than optimal, as shown in the online appendix.

We note that all of the study analyses were performed using the NSDUH R-DAS. In our judgment, R-DAS creates an unprecedented opportunity to study the distribution (and to some extent the dynamics) of extra-medical drug use in the United States. As such, an important project goal was to clarify current R-DAS research approaches. In addition, we have outlined an approach that can be used to estimate the R-DAS output for a population subgroup, for which the tabulated results ordinarily are suppressed due to confidentiality restrictions. This approach includes no inherent violations of confidentiality or privacy protection principles, although the approach is one that requires each end-user investigator to be vigilant about the possibility of inadvertent invasions of privacy and inadvertent disclosure of drug use and other sensitive behaviors of individual NSDUH participants. That is, the approach we outline is no substitute for each investigator and

research team evaluating whether individual-level participant confidentiality or privacy violations might occur when our approach is applied; this judgment must be made on a case-by-case basis. The R-DAS system response of “error” when there is a confidentiality restriction or suppression of output must serve as a warning to investigators that they are reaching boundary lines of a territory where protection of privacy is crucial; special care and inspection of output are required. Consultation with the R-DAS Help Desk may be needed.

At the end of the day, United States federal law dictates specific roles for government staff who must manage confidentiality of data from surveys like NSDUH, as well as penalties for mis-steps, but we believe it is an ethical responsibility of each individual end-user investigator to join in the federal effort to protect study participant privacy, ensuring that the study data being reported cannot be used to identify or impose risks on individual subjects or population subgroups. End-user investigators have two basic options, the first of which involves becoming more informed about disclosure risk assessment at a basic level, and applying suggested ‘best practices’. For example, the “Rainbow Series” of reports published by the US National Center on Health Statistics now is available as in a searchable online document file (<http://www.cdc.gov/nchs/products/series.htm>, last accessed 11 May 2014), as are online tutorials and other tools created by the Federal Committee on Statistical Methodology of the US Office of Management and Budget (<http://www.fcsfm.gov/committees/cdac/cdac.html>, last accessed 9 May 2014). When these files are searched for terms such as ‘disclosure analysis’ or ‘confidentiality,’ the search results include many useful suggestions that any end-user investigator can apply before disseminating or publishing estimates from shared data sources. One suggestion is for an end-user to seek out examples of crosstabulations with table cells that contain information on very few respondents (often with a rule of thumb that a value should be suppressed when the cell count is below a threshold set variously at $4 < n < 21$), much as the PUF creator might employ ‘data coarsening’ to preclude small cell counts. Another suggestion that requires no special statistical expertise is a ‘k-anonymity’ approach characterized as ‘the most rigorous standard’ for evaluation of whether a microdata set or analysis is ‘safe’. That is, a data set should be regarded as “safe” if there are at least k (usually $k = 3$) records that are identical with respect to the set of key variables [that might be used to identify a study respondent indirectly]” (Steel, 2004).

Here, R-DAS end-users might face a small obstacle because R-DAS does not yield unweighted

cell counts. Nonetheless, in experience to date, the R-DAS Help Desk staff members have been willing to confirm whether R-DAS cell counts fall below investigator-specified thresholds and in some instances have provided the actual unweighted cell values. In addition, our research team has prepared formulae that can be used to derive an approximate unweighted cell count for each cell of a table. These formulae can be used to screen for cell counts that may require a R-DAS Help Desk consultation (formulae available upon request to the first author).

A second option exists for end-user investigators who are methodologists interested in scientific disclosure risk assessment *per se*. The recent creation of a portal to NSDUH microdata files now makes it possible for methodologists to participate in disclosure risk assessment research, as illustrated in important methods research that challenged a prevailing sensitivity test based on the (n,k) Dominance rule ([Domingo-Ferrer and Torra, 2004](#)). A more recent example is the work of [Loong et al., 2013](#), who compared and contrasted inferences based on original microdata from the Cancer Care Outcomes Research and Surveillance Project versus inferences based on partially synthesized data intended to protect confidentiality by placing sensitive variable values with substitute values based on multiple imputation procedures. With NSDUH microdata in hand, methodologists should be able to conduct similar NSDUH studies, opening up possibilities that some of the current SDA restrictions might be relaxed (i.e., via use of using partially synthesized values for sensitive variables rather than suppression of all variable values).

Finally, we hope that this article might be of use to DAD journal readers from other countries, who can begin to make data sharing part and parcel of their public health research activities, if they have not done so already. Of course, efforts to protect study participant privacy, always will be needed to ensure that the study data being reported cannot be used to identify individual subjects. Traditions of data sharing, with disclosure risk assessment and protection of confidentiality, have become a regular part of international symposia, roundtable discussions, and other activities of leading researchers in Australia, the US, and other countries. Interested DAD readers may wish to consult already published reports on this topic (e.g., [Malin et al., 2013](#)), including downloadable books made available free of charge (e.g., [United States, 1985](#); [United States, 2002](#); [United States, 2004](#); [United States, 2005](#); [United States, 2007](#); [United States, 2014](#)).

Acknowledgment

This work is supported by a National Institute on Drug Abuse T32 research training program grant award (T32DA021129), JCA's NIDA Senior Scientist Award (K05DA015799), and by Michigan State University. The content is the sole responsibility of the authors and does not necessarily represent the official views of Michigan State University, the National Institute on Drug Abuse, or the National Institutes of Health, or SAMHSA. In addition, we would like to thank the United States Substance Abuse and Mental Health Services Administration Center for Behavioral Health Statistics and Quality for sponsoring the National Surveys on Drug Use and Health and for making the NSDUH datasets available for public use to allow research of this nature.

References

- Aldworth, J., Chromy, J. R., Davis, T. R., Foster, M. S., and Kathryn Spagnola, L. E. P., 2011. NSDUH statistical inference report.
URL <http://www.samhsa.gov/data/2k12/NSDUH2011MRB/NSDUHmrStatInference2011.pdf>
- Anthony, J. C., Warner, L. A., Kessler, R. C., 1994. Comparative epidemiology of dependence on tobacco, alcohol, controlled substances, and inhalants: Basic findings from the national comorbidity survey. *Exp. Clin. Psychopharmacol.* 2 (3), 244–268.
- Barrios, T., Diamond, R., Imbens, G. W., Kolesar, M., 2012. Clustering, spatial correlations, and randomization inference. *Journal of the American Statistical Association* 107 (498), 578–591.
- Casella, G., Berger, R. L., 2002. *Statistical Inference*, 2nd Edition. DUXBURY.
- Colliver, J. D., Compton, W. M., Gfroerer, J. C., Condon, T., 2006. Projecting drug use among aging baby boomers in 2020. *Annals of Epidemiology* 16 (4), 257–265.
- Cribbie, R. A., 2007. Multiplicity control in structural equation modeling. *Structural Equation Modeling* 14 (1), 98–112.
- DeAndrea, D. C., Troost, J. P., Anthony, J. C., 2013. Toward primary prevention of extra-medical OxyContin use among young people. *Prev. Med.* 57 (3), 244–246.

- Degenhardt, L., Whiteford, H., Hall, W. D., 2014. The global burden of disease projects: What have we learned about illicit drug use and dependence and their contribution to the global burden of disease? *Drug and Alcohol Review* 33 (1), 4–12.
- Deming, W. E., Stephan, F. F., et al., 1940. On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *The Annals of Mathematical Statistics* 11 (4), 427–444.
- Domingo-Ferrer, J., Torra, V., 2004. Disclosure risk assessment in statistical data protection. *Journal of Computational and Applied Mathematics* 164, 285–293.
- Dowling, K., Storr, C. L., Chilcoat, H. D., 2006. Potential influences on initiation and persistence of extramedical prescription pain reliever use in the us population. *Clin. J. Pain.* 22 (9), 776–783.
- Fisher, R., 1925. *Statistical Methods For Research Workers*. Cosmo study guides. Cosmo Publications.
- URL <http://books.google.com/books?id=4bTttAJR5kEC>
- Hawkes, N., 2013. Deaths from tramadol and legal highs reach new highs in england and wales. *BMJ* 347, f5336.
- Heeringa, S. G., West, B. T., Berglund, P. A., 2010. *Applied Survey Data Analysis*. Chapman & Hall/CRC Statistics in the Social and Behavioral Sciences.
- Help with the Restricted-use Data Analysis System (R-DAS), 2013. Substance abuse and mental health data archive (SAMHDA).
- URL <http://www.icpsr.umich.edu/icpsrweb/SAMHDA/support/faqs?label=rdas>
- Kelly, J. P., Cook, S. F., Kaufmana, D. W., Andersona, T., Rosenberga, L., Mitchell, A. A., 2008. Prevalence and characteristics of opioid use in the US adult population. *Pain* 138 (3), 507–513.
- Larance, B., Ambekar, A., Azim, T., Murthy, P., Panda, S., Degenhardt, L., Mathers, B., 2011. The availability, diversion and injection of pharmaceutical opioids in south asia. *Drug and Alcohol Review* 30 (3), 246–254.

- Loong, B., Zaslavsky, A. M., He, Y., Harrington, D. P., 2013. Disclosure control using partially synthetic data for large-scale health surveys, with applications to cancers. *Statistics in Medicine* 32 (24), 4139–4161.
- Malin, B. A., El Emam, K., O’Keefe, C. M., 2013. Biomedical data privacy: problems, perspectives, and recent advances. *Journal of the American Medical Informatics Association* 20 (1), 2–6.
- Manchikanti, L., Helm, S., Fellows, B., Janata, J. W., Pampati, V., Grider, J. S., Boswell, M. V., 2012. Opioid epidemic in the united states. *Pain Physician* 15, 2150–1149.
- Martins, S. S., Keyes, K. M., Storr, C. L., Zhu, H., Chilcoat, H. D., 2009. Pathways between non-medical opioid use/dependence and psychiatric disorders: results from the national epidemiologic survey on alcohol and related conditions. *Drug. Alcohol. Depend.* 103 (1), 16–24.
- McDonald, D. C., Carlson, K., Izrael, D., 2012. Geographic variation in opioid prescribing in the us. *The Journal of Pain* 13 (10), 988–996.
- Meier, E. A., Troost, J. P., Anthony, J. C., 2012. Extramedical use of prescription pain relievers by youth aged 12 to 21 years in the united states: national estimates by age and by year. *Arch. Pediatr. Adolesc. Med.* 166 (9), 803–807.
- Mohammed, A. M., Anthony, J. C., 2014. Female Cocaine Users Are At Excess Risk Of Becoming Cocaine Dependent Soon After Onset Of Cocaine Use: Estimates For The United States, 2002-2011. CPDD 76th Annual Meeting (Abstract).
URL <http://www.cpdd.vcu.edu/Pages/Meetings/CPDD14AbstractBook.pdf>
- Nicholas, R., Roche, A., Dobbin, M., Lee, N., 2013. Beyond the paper trail: using technology to reduce escalating harms from opioid prescribing in australia. *Aust. N. Z. J. Public. Health.* 37 (2), 139–147.
- Online Help for Analysis Programs-SDA 3.5, 2013. Substance abuse and mental health data archive (SAMHDA).
URL <http://www.icpsr.umich.edu/icpsrweb/content/SAMHDA/help/helpan.htm>
- Parker, M. A., Anthony, J. C., 2014. Should anyone be riding to glory on the now-descending limb of the crack-cocaine epidemic curve in the united states? *Drug. Alcohol. Depend.* 138, 225–228.

Python Software Foundation, 2013. Python Language Reference, version 2.7.

URL <http://www.python.org>

Radbruch, L., Glaeske, G., Grond, S., Münchberg, F., Scherbaum, N., Storz, E., Tholen, K., Zagermann-Muncke, P., Zieglänsberger, W., Hoffmann-Menzel, H., et al., 2013. Topical review on the abuse and misuse potential of tramadol and tilidine in germany. *Substance Abuse* 34 (3), 313–320.

Rice, W., 1988. Analyzing tables of statistical tests. *Evolution* 43 (1), 223–225.

Roxburgh, A., Bruno, R., Larance, B., Burns, L., 2011. Prescription of opioid analgesics and related harms in australia. *Med. J. Aust.* 195 (5), 280.

Schubert, I., Ihle, P., Sabatowski, R., 2013. Increase in opiate prescription in germany between 2000 and 2010: a study based on insurance data. *Deutsches Ärzteblatt International* 110 (4), 45–51.

Seedall, R. B., Anthony, J. C., 2013. Risk estimates for starting tobacco, alcohol, and other drug use in the united states: Malefemale differences and the possibility that ”limiting time with friends” is protective. *Drug. Alcohol. Depend.* 133 (2), 751–753.

Simoni-Wastila, L., Ritter, G., Strickler, G., 2004. Gender and other factors associated with the nonmedical use of abusable prescription drugs. *Subst. Use. Misuse.* 39 (1), 1–23.

Steel, P. M., 2004. Disclosure Risk Assessment for Microdata.

URL https://www.census.gov/srd/sdc/Steel_Disclosure_Risk_Assessment_for_Microdata.pdf

United States, 1985. Sharing Research Data. United States. National Research Council, Washington, DC.

United States, 2002. Access to Research Data in the 21st Century: An Ongoing Dialogue Among Interested Parties Report of a Workshop. United States. National Research Council, Washington, DC.

United States, 2004. Open Access and the Public Domain in Digital Data and Information for Science: Proceedings of an International Symposium. United States. National Research Council, Washington, DC.

United States, 2005. Expanding Access to Research Data: Reconciling Risks and Opportunities. United States. National Research Council, Washington, DC.

United States, 2007. Putting People on the Map: Protecting Confidentiality with Linked Social-Spatial Data. United States. National Research Council, Washington, DC.

United States, 2014. Proposed Revisions to the Common Rule for the Protection of Human Subjects in the Behavioral and Social Sciences. United States. National Research Council, Washington, DC.

United States Census Bureau, 2011. Age and sex composition: 2010.

URL <http://www.census.gov/prod/cen2010/briefs/c2010br-03.pdf>

Utah Department of Health Violence & Injury Prevention Program, 2011. Prescription Opioid Deaths in Utah, 2011.

Volkow, N. D., Frieden, T. R., Hyde, P. S., Cha, S. S., 2014. Medication-assisted therapiestackling the opioid-overdose epidemic. N. Engl. J. Med.

Table 1.

Bias assessment in re-estimation of the suppressed R-DAS output.

State	Incidence Estimate (%)			Standard Errors (%)		
	Estimated	Observed	Bias	Estimated	Observed	Bias
California	0.122	0.122	0.000	0.058	0.054	0.004
Colorado	0.343	0.343	0.000	0.249	0.220	0.029
Florida	0.318	0.318	0.000	0.108	0.097	0.011
Illinois	0.088	0.087	0.001	0.049	0.044	0.005
Michigan	0.168	0.169	0.001	0.070	0.065	0.005
Minnesota	0.443	0.441	0.002	0.250	0.209	0.042
New Mexico	0.222	0.220	0.002	0.153	0.130	0.023
New York	0.133	0.132	0.001	0.061	0.056	0.005
Oklahoma	0.344	0.344	0.000	0.231	0.238	-0.007
Pennsylvania	0.097	0.097	0.000	0.070	0.063	0.007
Tennessee	0.777	0.777	0.000	0.304	0.338	-0.033
Texas	0.199	0.199	0.000	0.095	0.088	0.007

Fig. 1: 8-Year R-DAS (2002 to 2009) incidence rate estimates for extra-medical use of pain relievers among Males.

Fig. 2: 8-Year R-DAS (2002 to 2009) incidence rate estimates for extra-medical use of pain relievers among Females.

Fig. 3: Results of significance testing for differences between males and females EMPPR incidence rates. States with statistically significant differences are highlighted in color.

Fig. 4: Estimated differences in incidence estimates with the corresponding confidence intervals.

Fig. 5: 8-Year R-DAS (2002 to 2009) incidence estimates for extra-medical use of prescription pain relievers among 12-44-year-olds.

Fig. 6: 8-Year R-DAS (2002 to 2009) incidence estimates for extra-medical use of prescription pain relievers among older adults.

A.1. Clarification of incidence estimation

In our NSDUH incidence estimates, we form a numerator based on the number of newly incident users during a specified time interval, and we form a denominator based on the number of individuals who had not tried the drug prior to the start of that interval (as best these quantities can be measured in NSDUH field surveys). That is, the denominator is based on an 'at risk' specification as opposed to a 'total population' specification. DAD readers interested in alternative specifications for incidence estimation may find useful information at this US Centers for Disease Control and Prevention web site: <http://www.cdc.gov/ophss/csels/dsepd/SS1978/Lesson3/Section2.html>

A.2. Clarification of spatial correlation issues in relation to possibly overly conservative Bonferroni correction.

Given these patterns, we would not say that the male-female comparisons necessarily involve spatial correlation, but we consider spatial correlation to be plausible. Plausibility is substantiated by interstate commerce by drug traffickers (e.g., those in Kansas City, MO, supplying users in Kansas City, KS, and vice versa), as well as plausible shared methods covariation (e.g., NSDUH interviewers employed for Kansas City working both states), among other sources of spatial correlation among individual values across state lines.

As for the math, to illustrate, let p_1 and q_1 be the male/female incidence estimates for state 1; p_2 and q_2 in state 2. Now, envision an extreme case with $\text{Corr}(p_1, p_2) = 1$ and $\text{Corr}(q_1, q_2) = 0$ (i.e., males estimates are perfectly correlated, while female estimates are independent).

Then,

$$\begin{aligned}\text{Cov}(p_1 - q_1, p_2 - q_2) &= \text{Cov}(p_1 - q_1, p_1 - q_2) = \text{Cov}(p_1, p_1) - \text{Cov}(p_1, q_2) - \text{Cov}(q_1, p_1) + \\ &\text{Cov}(q_1, q_2) = \text{Cov}(p_1, p_1) = \text{Var}(p_1) > 0.\end{aligned}$$

With non-zero covariance, there is going to be a non-zero correlation.

What does this have to do with the Bonferroni correction? In Section 3.2, we perform 51 tests. If the state level estimated differences are judged to be independent, it can be argued that the Bonferroni correction is required to take multiple testing into account with proper control of the probability of a falsely positive finding. However, due to plausible correlation of our observed state level estimates, we argue that the Bonferroni correction might well be too conservative, with tangible loss of power [Cribbie \(2007\)](#).

A.3. Example of generic R-DAS table specifications for these male-female and age-related analyses.

The first screenshot of R-DAS shows our specifications for the numerator and the denominator for the male-female incidence estimation table. The resulting R-DAS tabulated values are available at http://www.epi.msu.edu/vsevoloz/scripts/RDAS/rdas_incidents_gender.html. Notice that there is no age restriction on these analyses, and note that the R-DAS constructed variables have been used for specification of the newly incident users in the numerator and for the 'at risk' ("eligibles") in the denominator. This specification was repeated for each of the 51 jurisdictions shown in Table 1 (i.e., with each state and D.C. as a stratum).

The screenshot shows the R-DAS interface for the National Survey on Drug Use and Health: 8-Year R-DAS (2002 to 2009). The interface includes a header with the SAMHDA logo and a navigation bar. Below the header, there is a section titled "R-DAS Home > Select Variables" with a sub-header "Depending upon your setup, you may be able to drag variables to the boxes in the form, or simply type the variable name into the box." The form contains several input fields: "Row" with "RSEX", "Column" with "RECANL_B", "Control" with "STATE", and "Selection Filter(s)" with "ELGANL_B(1)". There are also checkboxes for "Percentage" (Column, Row, Total), "Summary statistics" (Show), and "Missing data" (Show). A "Run the Table" button is at the bottom.

Screenshot of R-DAS specification for male-female incidence estimation.

The second screenshot shows our specifications for the numerator and the denominator for the age-stratified incidence estimation table. The specifications for ages 45 to 106 produce an "error" message due to confidentiality restrictions. Notice that there is no sex or gender restriction on these analyses, and note that the R-DAS constructed variables have been used for specification of the newly incident users in the numerator and for the 'at risk' ("eligibles") in the denominator. This specification was repeated for each of the 51 jurisdictions shown in Table 2 (i.e., with each state and D.C. as a stratum).

The image shows two side-by-side screenshots of the R-DAS interface. The left screenshot shows the "Row" field set to "AGE(r12-44)" and the "Column" field set to "RECANL_B". The right screenshot shows the "Row" field set to "AGE(r12-109)" and the "Column" field set to "RECANL_B". Both screenshots show the same "Control" field set to "STATE" and "Selection Filter(s)" set to "ELGANL_B(1)". The "Percentage" checkboxes are also present in both.

Screenshots of R-DAS specification for different ages.

A.4. Clarification of zero covariance term assumption in Equations 2 and 4.

In practice, an R-DAS analyst generally must presume a zero covariance term as discussed in Sections 2.3 and 2.4. However, Eq.(2) and Eq.(4) are addressing different statistical problems and the same assumption of $Cov(\hat{p}_1, \hat{p}_2) = 0$ leads to different consequences. In Eq.(2), this zero covariance term assumption may introduce some bias into the calculation of the test statistic. In Eq.(4), the same assumption may introduce some bias into the variance estimation process of the suppressed output. In Section 3.3 we show that this latter bias tends to be minimal in this particular NSDUH estimation context.

A.5. Additional tables

Table 1.

Incidence estimates for extra-medical use of prescription pain relievers by state.

Male			Female		
State	Incidence Estimate (%)	SE	State	Incidence Estimate (%)	SE
Washington	1.57	0.26	Utah	1.81	0.29
Utah	1.54	0.27	Oklahoma	1.69	0.28
Arizona	1.39	0.30	Wisconsin	1.51	0.28
Wisconsin	1.38	0.21	New Mexico	1.50	0.19
Oregon	1.36	0.20	Louisiana	1.47	0.25
Indiana	1.35	0.22	Kentucky	1.41	0.23
Colorado	1.35	0.18	New Hampshire	1.37	0.17
Alabama	1.32	0.23	Tennessee	1.36	0.26
Nevada	1.30	0.21	Delaware	1.34	0.18
Michigan	1.25	0.10	Arizona	1.34	0.20
Tennessee	1.25	0.23	North Carolina	1.34	0.19
Montana	1.24	0.24	Rhode Island	1.33	0.18
Idaho	1.22	0.20	Oregon	1.33	0.15
Minnesota	1.19	0.23	Indiana	1.30	0.20
Rhode Island	1.17	0.19	Michigan	1.28	0.09
Missouri	1.17	0.18	Colorado	1.27	0.28
Virginia	1.16	0.27	Nevada	1.26	0.23
Ohio	1.11	0.09	Washington	1.25	0.15
Vermont	1.11	0.18	Kansas	1.25	0.22
Texas	1.07	0.09	Idaho	1.24	0.17
Louisiana	1.04	0.16	Missouri	1.20	0.20
Oklahoma	1.03	0.15	Wyoming	1.19	0.19
Kentucky	1.02	0.17	South Carolina	1.18	0.16
Arkansas	1.02	0.18	Ohio	1.15	0.08
South Dakota	0.98	0.16	Arkansas	1.11	0.16
California	0.98	0.08	Texas	1.10	0.10
Wyoming	0.97	0.12	Georgia	1.09	0.15
Maine	0.97	0.18	Connecticut	1.08	0.20
Delaware	0.97	0.18	Iowa	1.07	0.19
Alaska	0.95	0.15	Minnesota	1.07	0.19
Illinois	0.94	0.08	Alaska	1.05	0.17
Pennsylvania	0.92	0.08	Virginia	1.05	0.18
Massachusetts	0.92	0.13	Massachusetts	1.04	0.18
Florida	0.92	0.09	California	1.02	0.08
North Carolina	0.91	0.13	Nebraska	1.01	0.17
Kansas	0.90	0.19	Vermont	1.00	0.16
West Virginia	0.90	0.16	Florida	0.97	0.10
Iowa	0.87	0.12	Montana	0.95	0.12
New Jersey	0.83	0.15	New York	0.94	0.08
New Mexico	0.82	0.14	Maine	0.93	0.14
New Hampshire	0.79	0.12	Alabama	0.91	0.16
North Dakota	0.77	0.12	North Dakota	0.89	0.12
New York	0.77	0.07	South Dakota	0.88	0.14
Georgia	0.73	0.16	Pennsylvania	0.85	0.09
Mississippi	0.73	0.11	Mississippi	0.85	0.15
Connecticut	0.69	0.14	West Virginia	0.84	0.11
Hawaii	0.68	0.13	Illinois	0.83	0.07
South Carolina	0.67	0.12	Hawaii	0.82	0.24
Maryland	0.67	0.12	New Jersey	0.73	0.14
Nebraska	0.58	0.11	Maryland	0.68	0.13
DC	0.39	0.08	DC	0.64	0.14

Table 2.

Incidence estimates for extra-medical use of prescription pain relievers by age.

Age:12–44				Age:45–106		
State	Incidence Estimate (%)	SE		State	Incidence Estimate (%)	SE
Oregon	2.60	0.23		Tennessee	0.78	0.30
Washington	2.52	0.24		Wisconsin	0.66	0.25
Utah	2.51	0.27		Minnesota	0.44	0.25
Rhode Island	2.46	0.24		Oklahoma	0.34	0.23
Arizona	2.45	0.31		Colorado	0.34	0.25
Indiana	2.36	0.24		Florida	0.32	0.11
Oklahoma	2.33	0.20		Utah	0.28	0.25
Michigan	2.29	0.11		Louisiana	0.28	0.19
Nevada	2.27	0.27		Alabama	0.26	0.18
Kentucky	2.27	0.25		Hawaii	0.22	0.22
Delaware	2.27	0.25		New Mexico	0.22	0.15
Vermont	2.22	0.25		Virginia	0.21	0.22
Idaho	2.22	0.21		Texas	0.20	0.10
New Hampshire	2.21	0.20		Indiana	0.20	0.17
Montana	2.18	0.26		Kentucky	0.19	0.13
Ohio	2.18	0.12		Michigan	0.17	0.07
Wyoming	2.18	0.22		Iowa	0.16	0.14
Wisconsin	2.18	0.24		Kansas	0.15	0.14
Missouri	2.17	0.22		Washington	0.15	0.07
Louisiana	2.16	0.23		Maine	0.14	0.10
North Carolina	2.10	0.21		Arizona	0.14	0.14
Colorado	2.10	0.25		Missouri	0.14	0.13
New Mexico	2.09	0.21		New York	0.13	0.06
Arkansas	1.99	0.22		California	0.12	0.06
Alabama	1.92	0.18		Alaska	0.12	0.17
Virginia	1.91	0.23		Idaho	0.12	0.20
Kansas	1.90	0.25		Montana	0.12	0.12
Maine	1.85	0.20		Connecticut	0.11	0.13
South Carolina	1.82	0.20		DC	0.10	0.10
Tennessee	1.82	0.18		Pennsylvania	0.10	0.07
Massachusetts	1.78	0.18		Arkansas	0.10	0.00
West Virginia	1.77	0.20		Massachusetts	0.09	0.10
South Dakota	1.76	0.17		Illinois	0.09	0.05
Texas	1.73	0.10		North Dakota	0.08	0.06
Iowa	1.72	0.16		Nebraska	0.08	0.12
Minnesota	1.71	0.18		Rhode Island	0.08	0.12
Alaska	1.70	0.17		South Dakota	0.08	0.09
Pennsylvania	1.69	0.11		Nevada	0.07	0.13
California	1.66	0.09		Mississippi	0.07	0.07
Connecticut	1.65	0.23		West Virginia	0.06	0.08
Florida	1.61	0.10		North Carolina	0.06	0.07
Georgia	1.61	0.18		Ohio	0.06	0.05
North Dakota	1.55	0.15		New Hampshire	0.00	0.07
Illinois	1.54	0.09		South Carolina	0.00	0.08
New York	1.50	0.09		Delaware	0.00	0.05
New Jersey	1.48	0.19		Georgia	0.00	0.03
Nebraska	1.46	0.17		Maryland	0.00	0.01
Mississippi	1.43	0.16		New Jersey	0.00	0.05
Hawaii	1.28	0.17		Oregon	0.00	0.07
Maryland	1.25	0.16		Vermont	0.00	0.08
DC	0.85	0.12		Wyoming	0.00	0.06

References

Cribbie, R. A., 2007. Multiplicity control in structural equation modeling. *Structural Equation Modeling* 14 (1), 98–112.