



Published in final edited form as:

Drug Alcohol Depend. 2017 April 01; 173: 92–98. doi:10.1016/j.drugalcdep.2016.12.019.

Sex disparities in substance abuse research: Evaluating 23 years of structural neuroimaging studies*

Kimberly E. Lind^{1,2}, Eric J. Gutierrez³, Dorothy J. Yamamoto¹, Michael F. Regner¹, Sherry A. McKee⁴, and Jody Tanabe¹

¹Department of Radiology, University of Colorado School of Medicine, Aurora, CO

²Departments of Health Systems, Management and Policy, University of Colorado School of Public Health, Aurora, CO

³Biostatistics and Informatics, University of Colorado School of Public Health, Aurora, CO

⁴Department of Psychiatry, Yale University School of Medicine, New Haven, CT

Abstract

Background—Sex differences in brain structure and clinical course of substance use disorders underscores the need to include women in structural brain imaging studies. The NIH has supported the need for research to address sex differences. We evaluated female enrollment in substance abuse structural brain imaging research and the methods used to study sex differences in substance effects.

Methods—Structural brain imaging studies published through 2016 (n=230) were evaluated for number of participants by sex and substance use status and methods used to evaluate sex differences. Temporal trends in the numbers of participants by sex and substance use status were analyzed. We evaluated how often sex effects were appropriately analyzed and the proportion of studies that found sex by substance interactions on volumetric measures.

Results—Female enrollment increased over time, but remained significantly lower than male enrollment ($p=0.01$), with the greatest bias for alcohol and opiate studies. 79% of studies included both sexes; however, 74% did not evaluate sex effects or used an analytic approach that precluded detection of sex by substance use interactions. 85% of studies that stratified by sex reported different substance effects on brain volumes. Only 33% of studies examining two-way interactions

*Supplementary material can be found by accessing the online version of this paper at <http://dx.doi.org> and by entering doi:...

Corresponding Author: **Jody Tanabe** Department of Radiology, MS L954 University of Colorado Anschutz Medical Campus Aurora, CO 80045 (720) 848-8572 jody.tanabe@ucdenver.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflicts of Interest

No conflicts declared.

Contributors

KEL and JT designed the study. DJY and MFR conducted the literature review and screened articles for inclusion. EJG extracted and entered data. KEL and EJG completed the statistical analyses. All authors contributed to the writing and critical revision of the manuscript. All authors reviewed and approved the final manuscript.

found significant interactions, highlighting that many studies were underpowered to detect interactions.

Conclusions—Although female participation in substance use studies of brain morphometry has increased, sex disparity persists. Studying adequate numbers of both sexes and employing correct analytic approaches is critical for understanding sex differences in brain morphometric changes in substance abuse.

Keywords

Structural brain imaging; substance abuse; female inclusion; sexual dimorphism; grey matter volume

1. INTRODUCTION

Sex bias in biomedical research has been well documented, with females systematically understudied (see review by Beery and Zucker (Beery and Zucker, 2011)). Institutional efforts have been proposed to address this bias; for example, in 1993 the National Institutes of Health (NIH) initiated a requirement that the inclusion of women be addressed in all NIH-funded research (“NIH Revitalization Act of 1993 Public Law 103-43 - Women and Health Research - NCBI Bookshelf,” 1994). Two decades later, the NIH announced intentions to address sex bias in animal and basic science studies (Clayton and Collins, 2014). In substance abuse research, attitudes have evolved over time from an assumption that sex differences were negligible toward routine scrutiny of research in which sex differences are not modelled or investigated (Wetherington, 2007). Current scientific consensus is that fundamental differences exist in brain structure and function possibly in response to sexually selective evolutionary pressures (Cahill, 2014a). Brain differences between sexes have been shown in terms of neurodevelopmental trajectories (Lenroot et al., 2007), structural morphometry (Clayton and Collins, 2014; Luders et al., 2009; Peelle et al., 2012; Watanabe et al., 2013), connectivity (Cahill, 2014b; Duarte-Carvajalino et al., 2012; Gong et al., 2015; Ingalhalikar et al., 2014), and molecular biology (Al Nadaf et al., 2010; Cahill, 2006; Jazin and Cahill, 2010; Wu et al., 2014). Patients with substance use disorder (SUD) demonstrate sex differences in many natural history features, including age of first use, rate of drug consumption escalation, quantity consumed, affect, and behavior (Becker et al., 2012; Eaton et al., 2012; Hernandez-Avila et al., 2004). Additionally, sex differences in substance use disorders exist across etiological factors, prevalence rates, pharmacokinetics, pharmacodynamics, self-administration behavior, hormonal influences, associated health risks, treatment outcomes, and medication response (Anker and Carroll, 2011; Becker and Koob, 2016; Cooper et al., 2016; Hasin and Grant, 2015; Oberleitner et al., 2015; Smith et al., 2016). These fundamental sex differences in neurobiology and clinical course suggest interactions between neurophysiology, sex, and disease process. Although results of single-sex studies may, at times, generalize to both sexes (Sechzer et al., 1994), it is important to include and appropriately analyze both sexes to fully characterize this disease process. Inclusion and appropriate statistical methods can prevent reporting of an overall effect that may be true for only one sex, or reporting no effect when there may be opposing effects between sexes (Wetherington, 2007).

While adequate inclusion of female participants in SUD research is important, equally important is a statistically valid approach to evaluate and model effects of sex. One editorial reported that most studies that enrolled both sexes did not produce sex-specific analyses (Hayes and Redberg, 2008). Appropriate analytic approach is especially critical in instances where there may be a significant interaction between sex and SUD status. In the presence of an interaction, main effects of sex and SUD alone are likely to be biased estimates of effect size, which may lead to erroneous conclusions or the conclusion that no effect is present (Moyses Szklo and F.Javier Nieto, 2007). This is especially important in structural brain imaging research given known sexual dimorphism of brain structures beginning in childhood and extending over the lifespan (Giedd et al., 2012; Lenroot et al., 2007; Luders et al., 2009; Peelle et al., 2012). Although women's brains are smaller than men's, they have increased regional gray matter, even after accounting for differences in total brain volume (Luders et al., 2009). Sexual dimorphism of brain structure and function has been reported for numerous neurologic and psychiatric disorders, including SUD (Perry et al., 2013; Potenza et al., 2012; Rando et al., 2013a; Regner et al., 2015). As a result, male-biased neuroimaging research in SUD may be limited in generalizability given the aforementioned differences in structure and function between sexes.

This study evaluated trends in female enrollment and analytic approaches to evaluating sex effects within a specific area of SUD research – structural neuroimaging. The goal of this study was to evaluate experimental design and statistical analysis decisions regarding inclusion of female research participants. First, we aimed to evaluate temporal trends in the inclusion of female participants in SUD studies overall and by substance type. This first aim intended to inform general trends of female inclusion in studies to determine if female inclusion changed over time and to determine if disparities in female enrollment had been resolved. We hypothesized that there would be an increase in both number and relative proportion of female participants in SUD research over time. Second, we aimed to evaluate the proportion of studies that applied an appropriate analytic approach to evaluate sex effects in SUD studies. We hypothesized that a significant proportion of studies including both sexes would not apply appropriate analytic methods to evaluate or model sex effects and associated interactions. Third, of studies that either stratified by sex or modelled sex interactions, we aimed to evaluate the proportion of studies that found evidence of sex by SUD interactions. This aim was intended to inform the likelihood that sex by SUD interactions exist and are therefore worth investigating. Based on recent studies demonstrating that sex modulates effects of stimulant dependence on cortical (Rando et al., 2013b; Regner et al., 2015) and subcortical (Chang et al., 2005) volumes, we hypothesized that a majority of studies evaluating sex interactions have reported significant interactions.

2. METHODS

2.1 Literature review

We performed a systematic literature review of all articles studying SUD using structural neuroimaging. We included the following substance abuse categories: alcohol, cannabis, nicotine, opioids, stimulants, and polysubstance. Only studies reporting primary, original data were included; however, non-redundant bibliography entries within systematic reviews

and meta-analyses identified with our search terms were also included. The PubMed search was performed on January 7, 2016 using the search terms in Supplemental Figure S1¹. Abstracts published in English with full-text article access through our institution were included. All abstracts retrieved from this search (n=573) were independently screened and categorized by a post-doctoral neuroimaging fellow (DJY) and radiology resident physician (MFR). Discrepancies in categorization were discussed with the senior author, and a consensus decision was agreed upon by all three individuals. The inclusion process is illustrated in Figure 1.

2.2 Data collection

From each included article we abstracted the number of participants by sex and SUD status. Abstraction was performed by an applied biostatistics graduate student (EJG). We also abstracted the year published, country of corresponding author, funding sources, secondary substances, and analytic approach details. Figure 2 presents the data that were abstracted to determine if sex effects and interactions were tested or included and if they were significant. In some instances, articles did not provide all of the necessary information needed to follow our algorithm. If information was missing or if the analytic approach was unclear, we contacted the corresponding author and requested clarification of their methods.

2.3 Statistical analysis

We produced descriptive statistics for all variables abstracted from the articles including means, medians, and standard deviations for continuous measures and frequencies and percentages for categorical measures. All statistical tests for main effects used a type I error rate of 0.05. We used SAS version 9.4 for all analyses (SAS Institute, Cary, NC).

To investigate Aim 1, we plotted the number of female and male participants included by publication year. Next, we used generalized estimating equations and fit a generalized linear model with a Poisson distribution and log link function to evaluate the outcome of number of study participants by sex and SUD status (i.e., four outcomes for each study: male user, male control, female user, female control). We estimated the association between the number of study participants by sex and SUD to assess trends in participant enrollment irrespective of other study characteristics.

We evaluated covariates that were expected to explain a significant proportion of sample size variance among SUD studies. These additional covariates included: (1) class of substance, (2) funding source, (3) time, and (4) comorbid conditions. Time was defined as the year the study was published and modeled as a continuous numeric variable scaled to years past 1992. Comorbid conditions were defined as any study in which participants with a general medical disorder (e.g., HIV positive status) or psychiatric illness other than SUD (e.g., schizophrenia) were also included. All covariates were modeled as binary indicator variables. We considered all possible three-way and two-way interactions using a backward step-wise modelling approach. First, we fit a saturated model and evaluated the significance of three-way interactions; significant three-way interactions were retained at a significance

¹Supplementary material can be found by accessing the online version of this paper at <http://dx.doi.org> and by entering doi:...

level of $p = 0.20$ (Mickey and Greenland, 1989). After removing non-significant interactions, we then evaluated two-way interactions that were retained at a significance level of $p = 0.20$. Finally, we evaluated main effects that were retained at a significance level of $p = 0.20$. We retained any covariates that appeared to be confounders; confounders were defined as any variable that, after removal from the model, resulted in statistically significant differences in other variable estimates or variance.

To investigate Aim 2, we created a binary indicator of whether or not variance attributable to sex or its interactions with other variables was modelled in an analytically appropriate approach according to our algorithm in Figure 2. The proportion of articles meeting these criteria was calculated with a 95% confidence interval using the Agresti-Coull method (Agresti and Coull, 1998).

To investigate Aim 3, we limited our sample to articles that tested for sex interactions or stratified estimates of substance effect by sex. We calculated the proportion of articles that found evidence of a significant interaction between sex and substance use.

3. RESULTS

A total of 230 articles met the inclusion criteria (Figure 1). An additional 6 studies were excluded from our some of our analyses due to lack of appropriate sample size information. The earliest study was published in 1992, but most studies were published after 2000. Included articles studied the following substances: alcohol (37.8%), cannabis (15.2%), nicotine (11.7%), opioids (4.4%), stimulants (20.0%) and polysubstance (10.9%). NIH funding was reported for a majority (58.7%) of studies, followed by non-US government funding (40.4%), VA funding (9.6%) and 16.1% funded by other sources. A majority of studies were conducted in the US (57.0%), followed by 32.2% conducted in European countries.

3.1 Aim 1

In the unadjusted model evaluating mean sample size by sex and SUD, we found that females were less likely to be included as study participants whether they were non-users ($p=0.02$) or substance users ($p=0.01$). Model-based estimates of sex by user status resulted in estimated study mean sample sizes of 39.3 male substance users, 35.1 male non-users, 15.1 female substance users and 23.4 female non-users. Supplemental Table S1² presents estimates, standard errors, confidence intervals and p-values for the adjusted model. Figure 3 shows the trend in total and mean number of participants by substance use over time for NIH-funded and all studies. Although female participants were effectively excluded prior to 2000 and continue to be underrepresented compared to male participants, there is a slight increase in sampling proportion over time regardless of funding source. While there was a male sampling bias for all drug classes, studies of nicotine were the least biased and studies of alcohol and opioids were the most biased (Figure 4). For example, 24 of 87 alcohol use disorder studies (27.6% of AUD studies) included no females in their sample population;

²Supplementary material can be found by accessing the online version of this paper at <http://dx.doi.org> and by entering doi:...

accordingly, the first quartile includes a female-to-male ratio of zero. Opioid studies had the highest median male to female sampling ratio (Figure 4).

3.2 Aim 2

Of the 182 studies that included both sexes, 134 (73.6%) did not report an attempt to address sex effects (n=79) or did not report an appropriate method for modelling sex effects (n=55). In studies that did not stratify or include an interaction term, but instead used only a main effect of sex as a nuisance covariate, the majority did not report whether or not sex had a significant effect on the outcome (29 of 55); of those that reported sex effects (n=26), 6 found a significant main effect of sex while 20 did not (Figure 2). Of articles in which statistical methodology to address sex effects was unclear or not specified (n=44), we had a 50% response rate to clarification requests sent to study authors.

3.3 Aim 3

In studies that stratified their analyses of substance effect by sex, 81.3% found statistically significant differences in resulting estimates. In studies that used an interaction term, 31.3% found a significant interaction between substance use and sex.

4. DISCUSSION

Significant disparities in female inclusion exist from 1992 – 2015 across all major substances of abuse in structural neuroimaging studies. This finding is contrary to current NIH enrollment in which females are reported to represent roughly half of all clinical research participants (Clayton and Collins, 2014), and more in line with reports that cell and animal studies continue to have sex-bias (Beery and Zucker, 2011). We found that analyses of sex effects were not reported or not evaluated in an analytically appropriate approach in the majority of studies reviewed. In studies that did evaluate sex effects, those that stratified found that estimates of SUD effect differed for males compared to females in 81% of studies. Studies that modelled a two-way interaction for sex and SUD demonstrated a significant interaction in approximately one-third of studies. This difference may reflect statistically insufficient power to detect interactions when sub-groups are small.

Although female enrollment is increasing, in some instances sample sizes are insufficient to detect differences between males and females or to detect interactions between sex and substance use. Our findings demonstrated that there was a male sampling bias for all drug classes, where studies of nicotine were the least biased and studies of alcohol and opiates were the most biased. One factor that may contribute to this effect is the difference in substance use prevalence rates; males generally have higher rates of substance abuse and therefore there is a larger pool of potential research participants to recruit from, although substance use trends have been changing in recent years. Sex differences in prevalence rates finds that rates for males are higher across all classes of licit and illicit substance use; rates of tobacco use are most similar and rates of alcohol use are most dissimilar with male alcohol use disorder (AUD) prevalence 185% that of female AUD prevalence (Substance Abuse and Mental Health Services Administration, 2014). Although prevalence is higher in men, the impact on health care and societal costs may be greater for women possibly due to

increased rates of negative health outcomes, many of which are sex-specific (Greenfield et al., 2007; Weinberger et al., 2013), and co-morbidities with other psychiatric conditions trauma (McKee and Weinberger, 2015; National Institutes of Health, 2015). Another contributing factor may be that veterans are predominantly male, so research done through the Veterans' Administration generally is either male only or predominately male (Cardenas et al., 2005; Durazzo et al., 2015; Fein et al., 2002). We have found no literature indicating females are less interested or motivated to participate in research, however, some research suggests that eligibility criteria regarding pregnancy and birth control may limit female participation and that childcare responsibilities may be barriers to participation in research (Holdcroft, 2007; Liu and Mager, 2016).

While prevalence differs, the sampling bias is out of proportion to prevalence rate differences, and it remains critical that research balance samples across women and men to effectively evaluate for the presence of sex differences. This may require targeted recruitment to over-sample women. Further, sample size may be limited by funding; the financial implications may be a significant challenge for future research (Fields, 2014) especially for neuroimaging studies which are expensive to conduct. An issue that arose during this project was insufficient details on the statistical methods which in turn would impair the reproducibility – an issue which has received substantial attention from the NIH in recent years (Collins and Tabak, 2014) and constitutes a specific criterion of the review process. A significant number of published studies were unclear in their methodology regarding the analytic approach. Critical to ensuring reproducibility, authors should clearly state the analyses completed, describe which covariates were evaluated, and report how they were evaluated. We found that average enrollment of female substance users was 15 and female non-users was 23; this sampling power may have been insufficient to detect sex by SUD interactions compared to a stratified analytic approach. We attempted to evaluate the prevalence of significant sex by SUD interactions as a function of sample size, but we did not see a significant trend likely due to underrepresentation of studies with large sample sizes. To show this trend, we would need a greater number of studies with larger sample sizes, which the current literature lacks.

Most published articles in the area of substance abuse and structural brain imaging were captured by our search method and included articles dating back to the emergence of MRI-based research. Although some articles were excluded due to lack of full text availability (n=4) or not being published in English (n=10), this constituted only 14 articles and is unlikely to have imposed a significant degree of bias in our data sampling. Compared to a previous study examining women in research across a range of disciplines and which sampled articles from specific journals (Beery and Zucker, 2011), our approach provides a complete picture of a specific research area, structural human neuroimaging in SUD. Our search terms may have missed some relevant articles, however, we believe our sample population reflects the target population because: (1) our search terms appear to have captured the majority of articles, and (2) relevant articles not identified by these search terms are unlikely to differ substantially with respect to data abstracted (i.e., the sample is representative of this area of the literature). The most notable limitation in our data abstraction was the absence of certain details about the analytic approach that occurred in 44 studies; in half of these instances, the corresponding author did not respond to our request

for additional information. We are uncertain if non-responding authors were more or less likely to have employed an appropriate analytic approach compared to those who did respond and provide additional information. The most common responses from authors explaining the lack of proper treatment of sex effects were (1) sex effects were outside the scope of their study, or (2) sample population was underpowered to evaluate sex effects.

Our findings have important clinical implications. Because many studies that included women and used an analytically appropriate approach discovered significant sex by SUD interactions on brain structure (81.3% that stratified and 31.3% that used an interaction term), it suggests that sex significantly modulates SUD pathophysiology. Such modulation may be mediated by sex hormones or menstrual cycle, endophenotypes that predate disease, neurobehavioral mechanisms associated with socialized gender roles, co-morbidities, or stress responses known to differ in women compared to men (Becker et al., 2012; Brady and Randall, 1999; Chaplin et al., 2008; Hagemann et al., 2011). SUD is a chronic relapsing-remitting disease that is difficult to treat; if sex-specific patterns of altered brain structure and function can be identified, it could lead to different treatments in women and overall improvement in clinical management. As one example of this approach, Cosgrove and colleagues (Cosgrove et al., 2012) have documented key sex differences in the function of nicotinic receptors in abstinent smokers, providing a neurochemical explanation for why nicotine replacement treatments fail to work as well in women compared to men smokers (Smith et al., 2016).

In terms of generalizability to other areas of biomedical research, we suspect that structural brain imaging in SUD is not unique with respect to these issues of sample size and analytic approach. This suspicion is based on reports from other areas of biomedical research such as cardiology, where it is reported that females are also understudied (Harris and Douglas, 2000; Kim et al., 2008).

5. CONCLUSIONS

In summary, there are specific components of study design and dissemination that can be improved in the field. First, both sexes need to be included in studies. Second, studies need to have adequate sample size to ensure reasonable power to detect sex by substance abuse interactions. Third, analyses need to either stratify by sex or use an interaction term (in regression models) when evaluating substance effects on brain volume or structure. Finally, analytic details need to be provided with clear detail in manuscripts, with specific attention to how sex was evaluated in the analysis. Taken together, these changes would improve treatments for men and women, the end goal of biomedical research.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Role of Funding Source: This work was supported in part by National Institutes of Health grants F32 DA 041011, P50 DA 033945, R01 DA 027748, T32 AA 7464-39.

References

- Agresti A, Coull BA. Approximate Is Better than “Exact” for Interval Estimation of Binomial Proportions. *Am Stat.* 1998; 52:119–126. DOI: 10.2307/2685469
- Al Nadaf S, Waters PD, Koina E, Deakin JE, Jordan KS, Graves JA. Activity map of the tammar X chromosome shows that marsupial X inactivation is incomplete and escape is stochastic. *Genome Biol.* 2010; 11:R122.doi: 10.1186/gb-2010-11-12-r122 [PubMed: 21182760]
- Anker JJ, Carroll ME. Females are more vulnerable to drug abuse than males: evidence from preclinical studies and the role of ovarian hormones. *Curr Top Behav Neurosci.* 2011; 8:73–96. DOI: 10.1007/7854_2010_93 [PubMed: 21769724]
- Becker JB, Koob GF. Sex Differences in Animal Models: Focus on Addiction. *Pharmacol Rev.* 2016; 68:242–263. DOI: 10.1124/pr.115.011163 [PubMed: 26772794]
- Becker JB, Perry AN, Westenbroek C. Sex differences in the neural mechanisms mediating addiction: a new synthesis and hypothesis. *Biol Sex Differ.* 2012; 3:14.doi: 10.1186/2042-6410-3-14 [PubMed: 22676718]
- Beery AK, Zucker I. Sex bias in neuroscience and biomedical research. *Neurosci Biobehav Rev.* 2011; 35:565–572. DOI: 10.1016/j.neubiorev.2010.07.002 [PubMed: 20620164]
- Brady KT, Randall CL. Gender differences in substance use disorders. *Psychiatr Clin North Am.* 1999; 22:241–252. [PubMed: 10385931]
- Cahill L. Equal the same: sex differences in the human brain. *Cerebrum Dana Forum Brain Sci.* 2014a; 2014:5.
- Cahill L. Fundamental sex difference in human brain architecture. *Proc Natl Acad Sci U S A.* 2014b; 111:577–578. DOI: 10.1073/pnas.1320954111 [PubMed: 24381154]
- Cahill L. Why sex matters for neuroscience. *Nat Rev Neurosci.* 2006; 7:477–484. DOI: 10.1038/nrn1909 [PubMed: 16688123]
- Cardenas VA, Studholme C, Meyerhoff DJ, Song E, Weiner MW. Chronic active heavy drinking and family history of problem drinking modulate regional brain tissue volumes. *Psychiatry Res.* 2005; 138:115–130. DOI: 10.1016/j.psychresns.2005.01.002 [PubMed: 15766635]
- Chang L, Cloak C, Patterson K, Grob C, Miller EN, Ernst T. Enlarged striatum in abstinent methamphetamine abusers: a possible compensatory response. *Biol Psychiatry.* 2005; 57:967–974. DOI: 10.1016/j.biopsych.2005.01.039 [PubMed: 15860336]
- Chaplin TM, Hong K, Bergquist K, Sinha R. Gender Differences in Response to Emotional Stress: An Assessment Across Subjective, Behavioral, and Physiological Domains and Relations to Alcohol Craving. *Alcohol Clin Exp Res.* 2008; 32:1242–1250. DOI: 10.1111/j.1530-0277.2008.00679.x [PubMed: 18482163]
- Clayton JA, Collins FS. Policy: NIH to balance sex in cell and animal studies. *Nature.* 2014; 509:282–283. [PubMed: 24834516]
- Collins FS, Tabak LA. Policy: NIH plans to enhance reproducibility. *Nature.* 2014; 505:612–613. [PubMed: 24482835]
- Cooper J, Borland R, McKee SA, Yong HH, Dugue PA. Depression motivates quit attempts but predicts relapse: differential findings for gender from the International Tobacco Control Study. *Addict Abingdon Engl.* 2016; 111:1438–1447. DOI: 10.1111/add.13290
- Cosgrove KP, Esterlis I, McKee SA, Bois F, Seibyl JP, Mazure CM, Krishnan-Sarin S, Staley JK, Picciotto MR, O’Malley SS. Sex differences in availability of $\beta 2^*$ -nicotinic acetylcholine receptors in recently abstinent tobacco smokers. *Arch Gen Psychiatry.* 2012; 69:418–427. DOI: 10.1001/archgenpsychiatry.2011.1465 [PubMed: 22474108]
- Duarte-Carvajalino JM, Jahanshad N, Lenglet C, McMahon KL, de Zubicaray GI, Martin NG, Wright MJ, Thompson PM, Sapiro G. Hierarchical topological network analysis of anatomical human brain connectivity and differences related to sex and kinship. *NeuroImage.* 2012; 59:3784–3804. DOI: 10.1016/j.neuroimage.2011.10.096 [PubMed: 22108644]
- Durazzo TC, Mon A, Gazdzinski S, Yeh PH, Meyerhoff DJ. Serial longitudinal magnetic resonance imaging data indicate non-linear regional gray matter volume recovery in abstinent alcohol-dependent individuals. *Addict Biol.* 2015; 20:956–967. DOI: 10.1111/adb.12180 [PubMed: 25170881]

- Eaton NR, Keyes KM, Krueger RF, Balsis S, Skodol AE, Markon KE, Grant BF, Hasin DS. An invariant dimensional liability model of gender differences in mental disorder prevalence: evidence from a national sample. *J Abnorm Psychol.* 2012; 121:282–288. DOI: 10.1037/a0024780 [PubMed: 21842958]
- Fein G, Di Sclafani V, Cardenas VA, Goldmann H, Tolou-Shams M, Meyerhoff DJ. Cortical gray matter loss in treatment-naive alcohol dependent individuals. *Alcohol Clin Exp Res.* 2002; 26:558–564. [PubMed: 11981133]
- Fields RD. NIH policy: Mandate goes too far. *Nature.* 2014; 510:340–340. DOI: 10.1038/510340a
- Giedd JN, Raznahan A, Mills KL, Lenroot RK. Review: magnetic resonance imaging of male/female differences in human adolescent brain anatomy. *Biol Sex Differ.* 2012; 3:19. doi: 10.1186/2042-6410-3-19 [PubMed: 22908911]
- Gong Q, Dazzan P, Scarpazza C, Kasai K, Hu X, Marques TR, Iwashiro N, Huang X, Murray RM, Koike S, David AS, Yamasue H, Lui S, Mechelli A. A Neuroanatomical Signature for Schizophrenia Across Different Ethnic Groups. *Schizophr Bull.* 2015; 41:1266–1275. DOI: 10.1093/schbul/sbv109 [PubMed: 26264820]
- Greenfield SF, Brooks AJ, Gordon SM, Green CA, Kropp F, McHugh RK, Lincoln M, Hien D, Miele GM. Substance abuse treatment entry, retention, and outcome in women: a review of the literature. *Drug Alcohol Depend.* 2007; 86:1–21. DOI: 10.1016/j.drugalcdep.2006.05.012 [PubMed: 16759822]
- Hagemann G, Ugur T, Schleussner E, Mentzel HJ, Fitzek C, Witte OW, Gaser C. Changes in Brain Size during the Menstrual Cycle. *PLoS ONE.* 2011; 6doi: 10.1371/journal.pone.0014655
- Harris DJ, Douglas PS. Enrollment of Women in Cardiovascular Clinical Trials Funded by the National Heart, Lung, and Blood Institute. *N Engl J Med.* 2000; 343:475–480. DOI: 10.1056/NEJM200008173430706 [PubMed: 10944565]
- Hasin DS, Grant BF. The National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) Waves 1 and 2: review and summary of findings. *Soc Psychiatry Psychiatr Epidemiol.* 2015; 50:1609–1640. DOI: 10.1007/s00127-015-1088-0 [PubMed: 26210739]
- Hayes SN, Redberg RF. Dispelling the myths: calling for sex-specific reporting of trial results. *Mayo Clin Proc.* 2008; 83:523–525. DOI: 10.4065/83.5.523 [PubMed: 18452679]
- Hernandez-Avila CA, Rounsaville BJ, Kranzler HR. Opioid-, cannabis- and alcohol-dependent women show more rapid progression to substance abuse treatment. *Drug Alcohol Depend.* 2004; 74:265–272. DOI: 10.1016/j.drugalcdep.2004.02.001 [PubMed: 15194204]
- Holdcroft A. Gender bias in research: how does it affect evidence based medicine? *J R Soc Med.* 2007; 100:2–3.
- Ingalhalikar M, Smith A, Parker D, Satterthwaite TD, Elliott MA, Ruparel K, Hakonarson H, Gur RE, Gur RC, Verma R. Sex differences in the structural connectome of the human brain. *Proc Natl Acad Sci U S A.* 2014; 111:823–828. DOI: 10.1073/pnas.1316909110 [PubMed: 24297904]
- Jazin E, Cahill L. Sex differences in molecular neuroscience: from fruit flies to humans. *Nat Rev Neurosci.* 2010; 11:9–17. DOI: 10.1038/nrn2754 [PubMed: 20019686]
- Kim ESH, Carrigan TP, Menon V. Enrollment of Women in National Heart, Lung, and Blood Institute-Funded Cardiovascular Randomized Controlled Trials Fails to Meet Current Federal Mandates for Inclusion. *J Am Coll Cardiol.* 2008; 52:672–673. DOI: 10.1016/j.jacc.2008.05.025 [PubMed: 18702973]
- Lenroot RK, Gogtay N, Greenstein DK, Wells EM, Wallace GL, Clasen LS, Blumenthal JD, Lerch J, Zijdenbos AP, Evans AC, Thompson PM, Giedd JN. Sexual dimorphism of brain developmental trajectories during childhood and adolescence. *NeuroImage.* 2007; 36:1065–1073. DOI: 10.1016/j.neuroimage.2007.03.053 [PubMed: 17513132]
- Liu KA, Mager NAD. Women's involvement in clinical trials: historical perspective and future implications. *Pharm Pract.* 2016; 14doi: 10.18549/PharmPract.2016.01.708
- Luders E, Gaser C, Narr KL, Toga AW. Why sex matters: brain size independent differences in gray matter distributions between men and women. *J Neurosci Off J Soc Neurosci.* 2009; 29:14265–14270. DOI: 10.1523/JNEUROSCI.2261-09.2009
- McKee SA, Weinberger AH. Innovations in Translational Sex and Gender-Sensitive Tobacco Research. *Nicotine Tob Res.* 2015; 17:379–381. DOI: 10.1093/ntr/ntu335 [PubMed: 25762746]

- Mickey RM, Greenland S. The impact of confounder selection criteria on effect estimation. *Am J Epidemiol.* 1989; 129:125–137. [PubMed: 2910056]
- Szklo, Moyses, Nieto, F Javier. *Epidemiology Beyond the Basics*. 2nd. Jones and Bartlett Publishers; Sudbury Massachusetts: 2007.
- National Institutes of Health. Women | National Institute on Alcohol Abuse and Alcoholism (NIAAA) [WWW Document]. 2015. URL <https://www.niaaa.nih.gov/alcohol-health/special-populations-co-occurring-disorders/women> (accessed 12.2.16)
- NIH Revitalization Act of 1993 Public Law 103-43 - Women and Health Research - NCBI Bookshelf [WWW Document]. 1994. URL <http://www.ncbi.nlm.nih.gov/books/NBK236531/> (accessed 8.12.16)
- Oberleitner LMS, Smith PH, Weinberger AH, Mazure CM, McKee SA. Impact of Exposure to Childhood Maltreatment on Transitions to Alcohol Dependence in Women and Men. *Child Maltreat.* 2015; 20:301–308. DOI: 10.1177/1077559515591270 [PubMed: 26130105]
- Peelle JE, Cusack R, Henson RNA. Adjusting for global effects in voxel-based morphometry: gray matter decline in normal aging. *NeuroImage.* 2012; 60:1503–1516. DOI: 10.1016/j.neuroimage.2011.12.086 [PubMed: 22261375]
- Perry RI, Krmpotich T, Thompson LL, Mikulich-Gilbertson SK, Banich MT, Tanabe J. Sex modulates approach systems and impulsivity in substance dependence. *Drug Alcohol Depend.* 2013; 133:222–227. DOI: 10.1016/j.drugalcdep.2013.04.032 [PubMed: 23725607]
- Potenza MN, Hong KA, Lacadie CM, Fulbright RK, Tuit KL, Sinha R. Neural correlates of stress-induced and cue-induced drug craving: influences of sex and cocaine dependence. *Am J Psychiatry.* 2012; 169:406–414. DOI: 10.1176/appi.ajp.2011.11020289 [PubMed: 22294257]
- Rando K, Tuit K, Hannestad J, Guarnaccia J, Sinha R. Sex differences in decreased limbic and cortical grey matter volume in cocaine dependence: a voxel-based morphometric study. *Addict Biol.* 2013a; 18:147–160. DOI: 10.1111/adb.12008 [PubMed: 23167305]
- Rando K, Tuit K, Hannestad J, Guarnaccia J, Sinha R. Sex differences in decreased limbic and cortical grey matter volume in cocaine dependence: a voxel-based morphometric study. *Addict Biol.* 2013b; 18:147–160. DOI: 10.1111/adb.12008 [PubMed: 23167305]
- Regner MF, Dalwani M, Yamamoto D, Perry RI, Sakai JT, Honce JM, Tanabe J. Sex Differences in Gray Matter Changes and Brain-Behavior Relationships in Patients with Stimulant Dependence. *Radiology.* 2015; :142541.doi: 10.1148/radiol.2015142541
- Sechzer JA, Rabinowitz VC, Denmark FL, McGinn MF, Weeks BM, Wilkens CL. Sex and gender bias in animal research and in clinical studies of cancer, cardiovascular disease, and depression. *Ann N Y Acad Sci.* 1994; 736:21–48. [PubMed: 7710209]
- Smith PH, Weinberger AH, Zhang J, Emme E, Mazure CM, McKee SA. Sex Differences in Smoking Cessation Pharmacotherapy Comparative Efficacy: A Network Meta-analysis. *Nicotine Tob Res Off J Soc Res Nicotine Tob.* 2016; doi: 10.1093/ntr/ntw144
- Substance Abuse and Mental Health Services Administration. Summary of National Findings, NSDUH Series H-48, HHS Publication No (SMA) 14-4863. Substance Abuse and Mental Health Services Administration; Rockville, MD: 2014. Results from the 2013 National Survey on Drug Use and Health.
- Watanabe M, Liao JH, Jara H, Sakai O. Multispectral quantitative MR imaging of the human brain: lifetime age-related effects. *Radiogr Rev Publ Radiol Soc N Am Inc.* 2013; 33:1305–1319. DOI: 10.1148/rg.335125212
- Weinberger AH, Mazure CM, Morlett A, McKee SA. Two decades of smoking cessation treatment research on smokers with depression: 1990–2010. *Nicotine Tob Res Off J Soc Res Nicotine Tob.* 2013; 15:1014–1031. DOI: 10.1093/ntr/nts213
- Wetherington CL. Sex-gender differences in drug abuse: a shift in the burden of proof? *Exp Clin Psychopharmacol.* 2007; 15:411–417. DOI: 10.1037/1064-1297.15.5.411 [PubMed: 17924775]
- Wu H, Luo J, Yu H, Rattner A, Mo A, Wang Y, Smallwood PM, Erlanger B, Wheelan SJ, Nathans J. Cellular resolution maps of X chromosome inactivation: implications for neural development, function, and disease. *Neuron.* 2014; 81:103–119. DOI: 10.1016/j.neuron.2013.10.051 [PubMed: 24411735]

Highlights

- Female participation in substance use studies of structural brain imaging has increased over time.
- Females are included as research participants at significantly lower rates than males.
- Many studies do not analyze sex effects or use an analytic approach that cannot detect sex effects.
- 85% of studies that stratified by sex reported different substance effects on brain volumes.

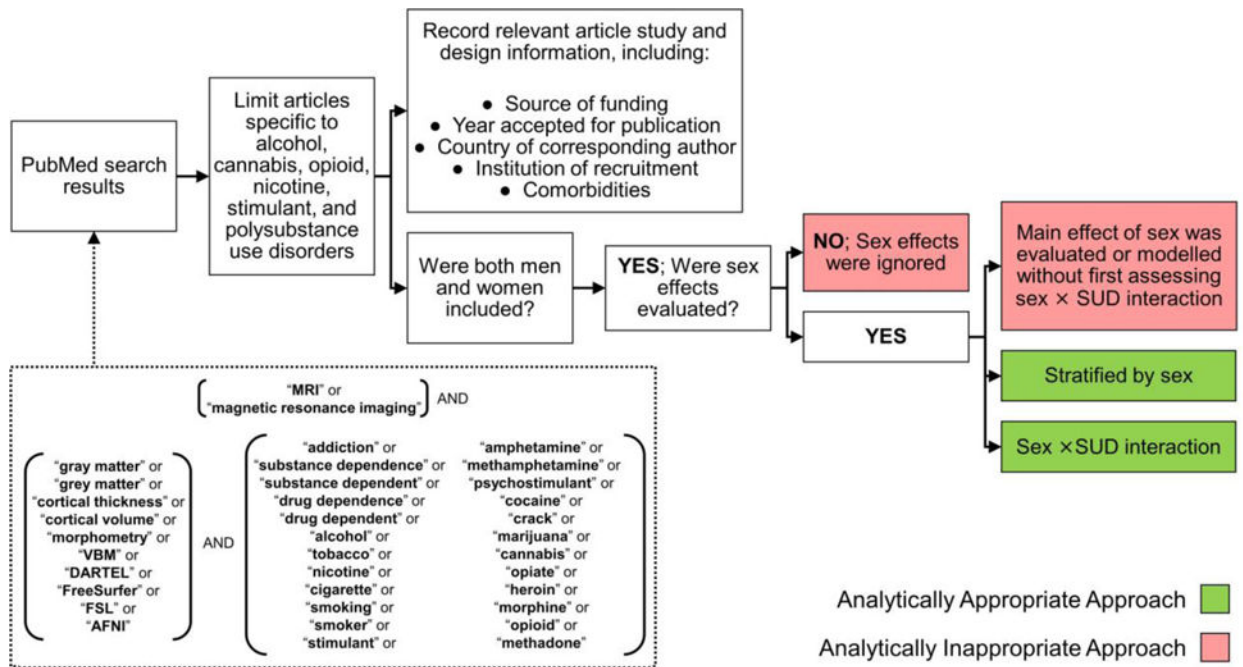


Figure 1. Article inclusion and exclusion (PRISMA Diagram).

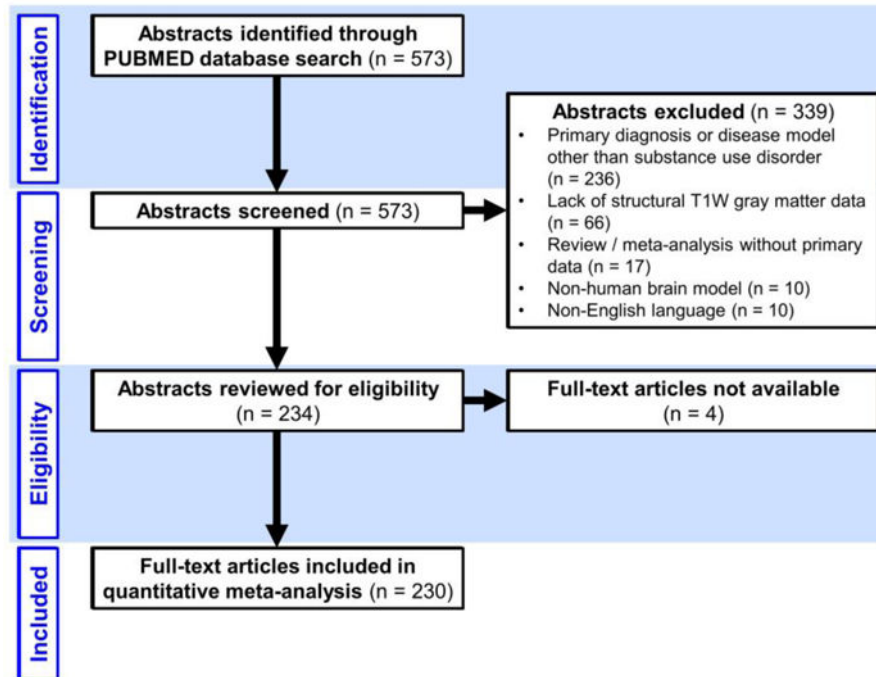


Figure 2.
Evaluation Results of Statistical Methodology.

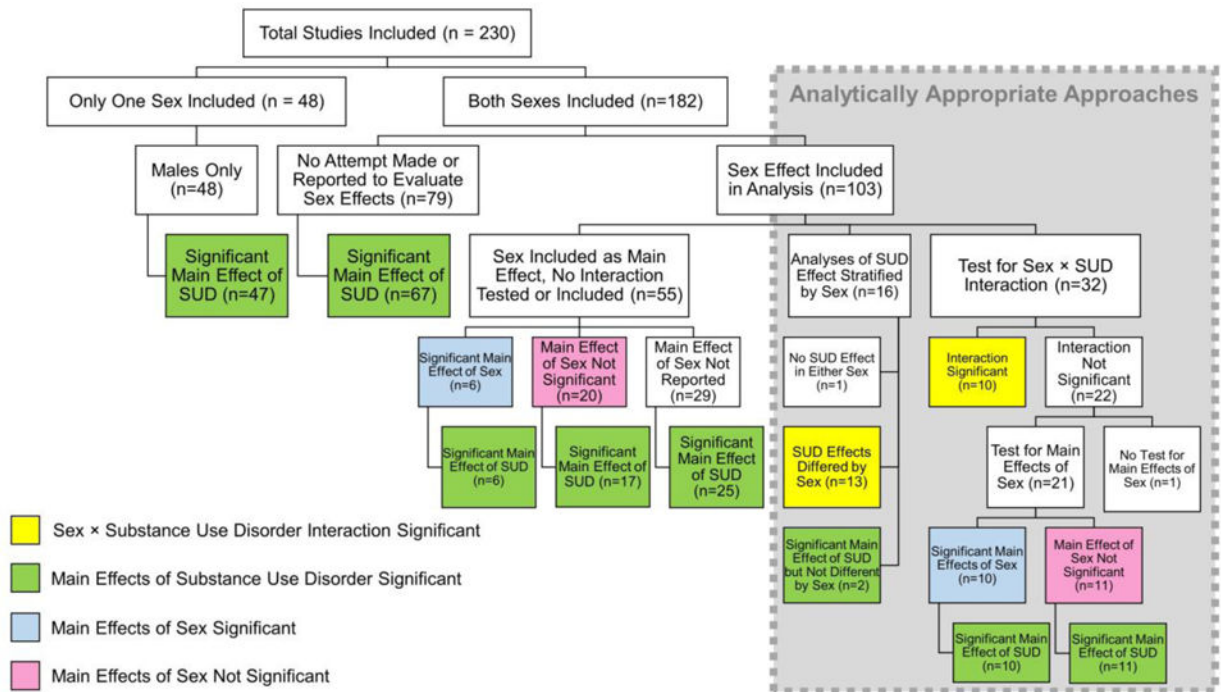


Figure 3. Enrollment by sex and substance use status over time

Top, number of NIH studies and all studies meeting inclusion criteria each year. Second box from top, total number of participants in the literature each year by group and sex. Third box from top, mean number of participants per study per year by group and sex. Bottom box, mean number of participants in NIH-funded studies per study per year by group and sex. Male and female enrollment were significantly different in SUD ($p=0.01$) and controls ($p=0.02$), with systematically higher enrollment of male participants in both groups over time.

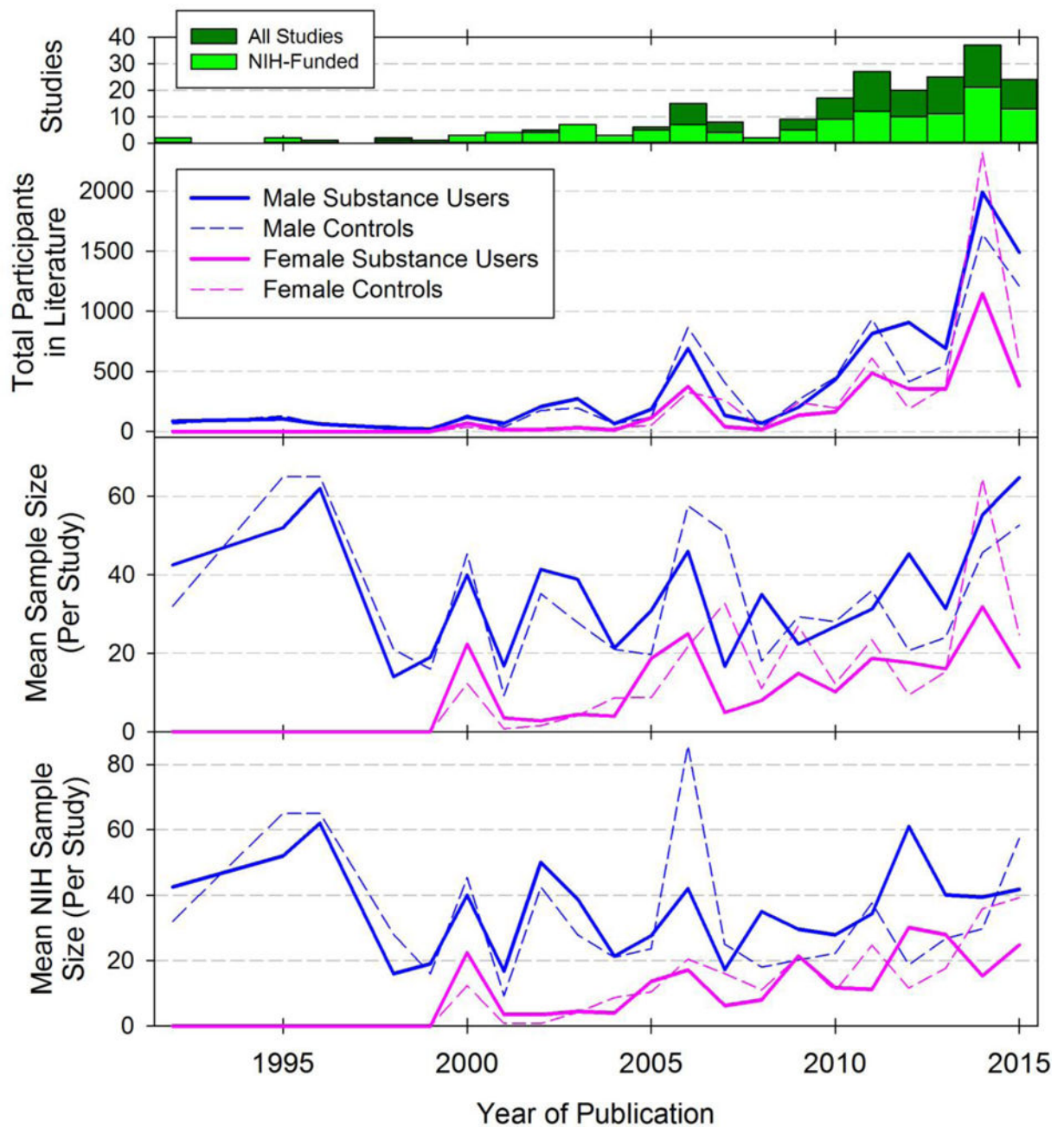


Figure 4. Female to male sampling ratio by substance type

For each study in the sample population, the ratio of total females to males was calculated to form a distribution with cardinality equal to the number of included studies. The top and bottom edges of each box represent the 3rd and 1st quartiles, and lines within the boxes indicate medians. Whiskers above and below each box represent the 90th and 10th percentiles. Statistical outliers are plotted as points.