PHP 2550 Project 1

Himashree Chandru

2023-10-05

Objective

The main objective of this exploratory data analysis is to determine the presence of any associations between exposure to smoking during pregnancy (SDP) and environmental tobacco smoke (ETS), and child outcomes related to substance use, self-regulation, internalizing problems (IP), externalizing problems (EP), attention problems (AP), Attention Deficit Hyperactivity Disorder (ADHD), and Autism Spectrum Disorder (ASD).

Data

The data was collected from a subset of 100 mothers who participated in a previous randomized controlled trial aimed at reducing SDP and ETS using a tailored video intervention, and their children, as part of Dr. Lauren Micalizzi's research on prenatal tobacco exposure and child outcomes at the Brown University School of Public Health.

There are 49 observations of 78 variables in this data set. Each observation corresponds to a particular parent-child pair and the variables contain information on demographics, on exposure, and on outcomes in children collected from both parents and the children. The data also contains information on outcomes in the parents, but these are not of interest in our current study.

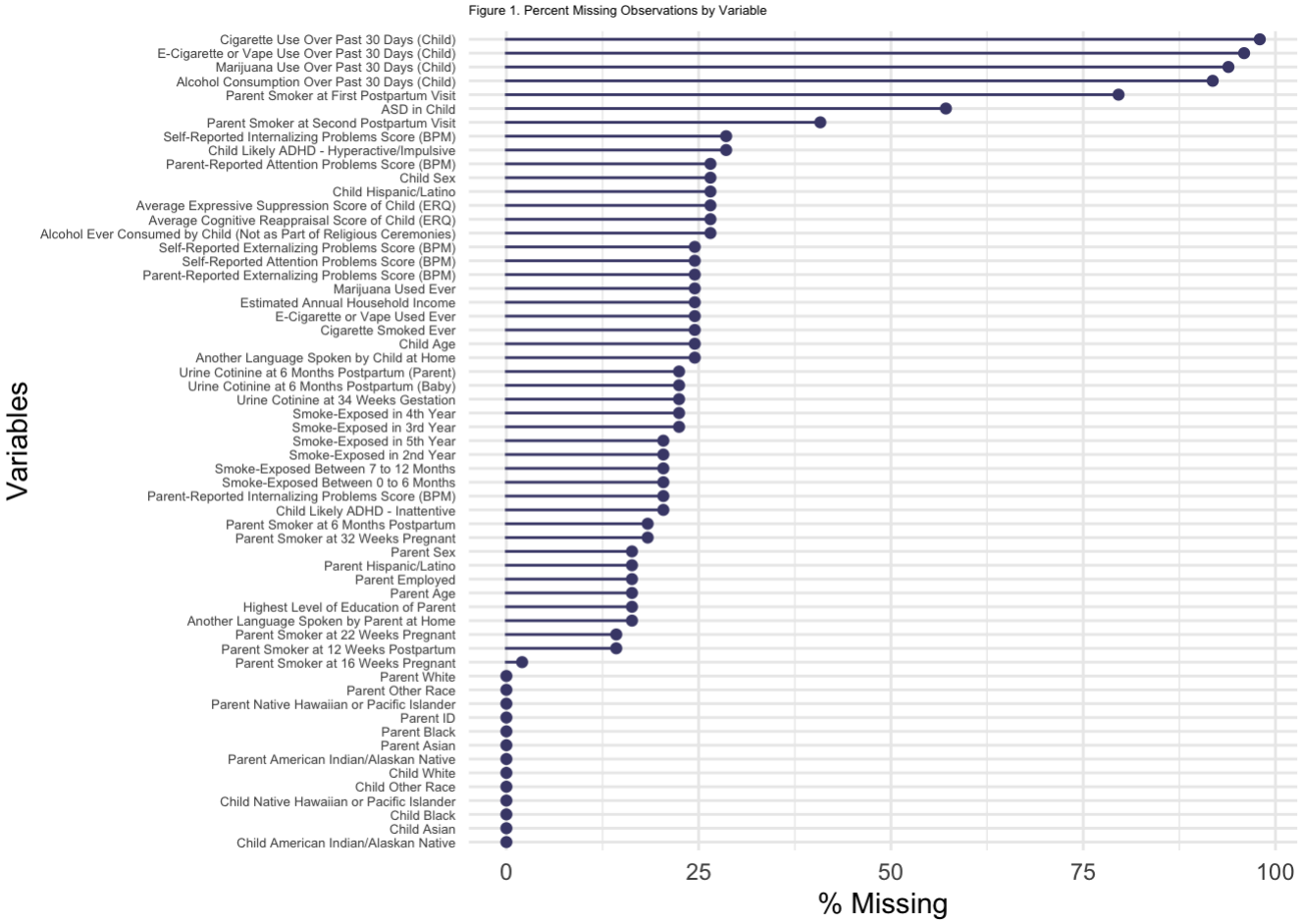
Variables

The variables included in this analysis include those on:

- Parent demographics (race, age, sex, number of languages spoken, employment, highest level of education and annual household income),
- Child demographics (race, age, sex, number of languages spoken),
- Exposure to SDP (self-reported information from the parent about whether they smoked when 16 weeks, 22 weeks and 32 weeks pregnant and urine cotinine (UC) levels from parent at 34 weeks pregnant)
- Exposure to ETS (self-reported information on whether the parent or partner smoked during the first 5 years of the child's life, on whether the parent smoked at the first and second postpartum visits, and at 12 weeks and 6 months postpartum, and urine cotinine levels from both mom and baby at 6 months postpartum)
- Substance use in child (if cigarettes, e-cigarettes/vape, marijuana, or alcohol were ever used and if so, how frequently in the last 30 days),
- Internalizing, attention and externalizing problems in child (self-reported and parent-reported scores on the Brief Problem Monitor),
- Self-regulation in child (self-reported average cognitive reappraisal (CR) and expressive suppression (ES) scores on the Emotion Regulation Questionnaire),
- ADHD in child (SWAN scores on the inattentive and hyperactive items),
- ASD in child (absence, presence or suspicion)

Exploratory Data Analysis

A Summary of the Missing Data



From Figure 1, it can be seen that almost all of the variables have some amount of missing data. In particular, the variables indicating the frequency of substance use in child over the past 30 days, and the presence of ASD in child, are missing greater than 50% of observations. The variables indicating if the parent was a smoker at the second postpartum visit, self-reported IP score, SWAN score on hyperactive items, parent-reported AP score, sex of the child, whether the child was Hispanic/Latino, average CR and ES scores of the child, and if alcohol was ever consumed by the child have greater than 25% but less than 50% of observations missing. However, since the sample size is quite small (n = 49), imputation is not considered.

Demographics

Tables 1 and 2 summarize the demographics information of the parents and children, respectively. From Table 1, it can be seen that the majority of parents are White and/or Hispanic/Latino, and the median parent age is 37. The majority of parents are employed full-time and appear to have received some college-level education. The median estimated annual household income is $42,000. There appears to be a parent that reported their sex as male, which might be the result of an entry error.

From Table 2, it can be seen that the majority of children are also White and/or Hispanic/Latino, and the median child age is 14. 64% of the children are boys and 36% are girls.

Table 1. Parent Demographics

Characteristic	**N = 49**[1]
Parent Race	
Hispanic/Latino	13 (32%)
American Indian/Alaskan Native	4 (8.2%)
Asian	0 (0%)

[1] n (%); Median (IQR)

Table 1. Parent Demographics

Characteristic	N = 49[1]
Native Hawaiian or Pacific Islander	8 (16%)
Black	0 (0%)
White	26 (53%)
Other Race	6 (12%)
Parent Age	37 (35, 39)
Parent Sex	
Male	1 (2.4%)
Female	40 (98%)
Another Language Spoken by Parent at Home	15 (37%)
Parent Employed	
No	12 (29%)
Part-Time	7 (17%)
Full-Time	22 (54%)
Highest Level of Education of Parent	
Some High School	3 (7.3%)
High School	3 (7.3%)
GED	5 (12%)
Some College	15 (37%)
2 Year Degree	3 (7.3%)
4 Year Degree	10 (24%)
Postgraduate Degree	2 (4.9%)
Estimated Annual Household Income	42,000 (20,000, 65,000)

[1] n (%); Median (IQR)

Table 2. Child Demographics

Characteristic	N = 49[1]
Child Race	
Hispanic/Latino	15 (42%)
American Indian/Alaskan Native	5 (10%)
Asian	0 (0%)
Native Hawaiian or Pacific Islander	0 (0%)
Black	15 (31%)
White	19 (39%)
Other Race	5 (10%)
Child Age	14 (13, 15)
Child Sex	
Male	23 (64%)
Female	13 (36%)
Another Language Spoken by Child at Home	11 (30%)

[1] n (%); Median (IQR)

Substance Use in Child

Table 3 summarizes information on whether the child has ever used cigarettes, e-cigarettes/vape, marijuana, or alcohol, and the frequency of usage over the past 30 days. Alcohol seems to have ever been used the most (14%), and marijuana seems to be the most frequently used (median frequency of 12 days).

Table 3. Substance Use in Child

Characteristic	N = 49[1]
Substance Used (Ever)	
Cigarette	1 (2.7%)
E-Cigarette or Vape	3 (8.1%)
Marijuana	3 (8.1%)
Alcohol (Not as Part of Religious Ceremonies)	5 (14%)
Number of Days Used in the Past 30 Days	

[1] n (%); Median (IQR)

Table 3. Substance Use in Child	
Characteristic	**N = 49**[1]
Cigarette	0 (0, 0)
E-Cigarette or Vape	1 (1, 2)
Marijuana	12 (8, 15)
Alcohol	1 (0, 3)

[1] n (%); Median (IQR)

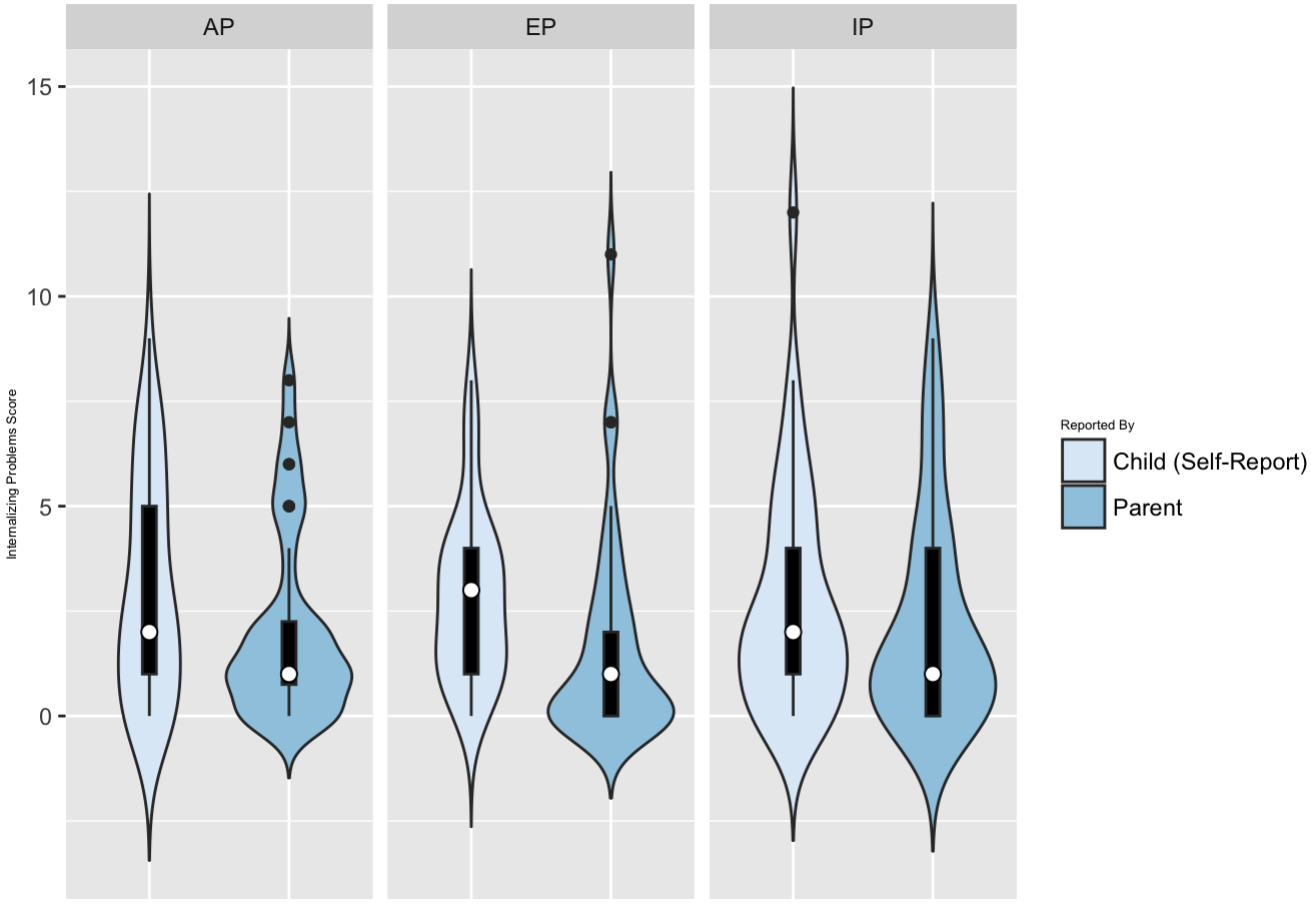
Self-Regulation, Internalizing Problems, Externalizing Problems, Attention Problems, ASD and ADHD in Child

Table 4 summarizes information about self-regulation, IP, EP, AP, ASD and ADHD in the children. The distributions of the self-reported and parent-reported scores for IP, EP and AP seem to be different, as can be visually seen in Figure 2. This difference indicates that we need to explore associations of between the exposures and both self-reported and parent-reported scores.

Table 4. Self-Regulation, Internalizing Problems, Externalizing Problems, Attention Problems, ASD and ADHD in Child	
Characteristic	**N = 49**[1]
Self-Regulation	
Average Cognitive Reappraisal Score	3.00 (2.83, 3.83)
Average Expressive Suppression Score	2.50 (2.25, 3.31)
Internalizing Problems	
Self-Reported Score	2 (1, 4)
Parent-Reported Score	1 (0, 4)
Externalizing Problems	
Self-Reported Score	3 (1, 4)
Parent-Reported Score	1 (0, 2)
Attention Problems	
Self-Reported Score	2 (1, 5)
Parent-Reported Score	1 (1, 2)
ASD in Child	
No	19 (90%)
Diagnosed	1 (4.8%)
Suspected	1 (4.8%)
ADHD in Child	
Likely ADHD - Hyperactive/Impulsive	20 (57%)
Likely ADHD - Inattentive	33 (85%)

[1] Median (IQR); n (%)

Figure 2. Difference in Distributions of Child-Reported and Parent-Reported Scores



Smoking During Pregnancy

Table 5 summarizes smoking by the parent at 16, 22 and 32 weeks pregnant, and urine cotinine levels of the parent at 34 weeks pregnant.

Table 5. Smoking During Pregnancy

Characteristic	**N = 49**[1]
Smoker At 16 Weeks	12 (25%)
Smoker At 22 Weeks	13 (31%)
Smoker At 32 Weeks	10 (25%)
UC At 34 Weeks Gestation	1 (0, 37)

[1] n (%); Median (IQR)

Postpartum Smoking

Table 6 summarizes postpartum smoking by the parent at the 1st and 2nd postpartum visits, and at 12 weeks and 6 months postpartum. About 20-40% of parents seem to have been smoking upto 6 months after giving birth.

Table 6. Postpartum Smoking

Characteristic	**N = 49**[1]
At First Postpartum Visit	3 (30%)
At Second Postpartum Visit	7 (24%)
At 12 Weeks Postpartum	12 (29%)
At 6 Months Postpartum	16 (40%)

[1] n (%)

Environmental Tobacco Smoke

Table 7 summarizes smoke exposure of the baby due to smoking by either the parent or partner, and the urine cotinine levels of the baby and parent at 6 months postpartum. About 20-25% of children seem to have been exposed to ETS during at least one of the first 5 years of their life. It is worth noting here that the median UC levels are higher for both baby (1.5) and parent (15) at 6 months postpartum than the median UC level at 34 weeks gestation (1).

Table 7. Postpartum Environmental Tobacco Smoke Exposure

Characteristic	N = 49[1]
Smoke Exposure	
0 to 6 Months	10 (26%)
7 to 12 Months	9 (23%)
2nd Year	11 (28%)
3rd Year	11 (29%)
4th Year	10 (26%)
5th Year	10 (26%)
UC levels	
At 6 Months Postpartum (Parent)	15 (1, 119)
At 6 Months Postpartum (Baby)	1.5 (0.6, 4.0)

[1] n (%); Median (IQR)

Association Between Timing of SDP and Outcomes

Substance Use by SDP Timing

Table 8 summarizes substance use by child by the timing of SDP. Overall, children whose parent smoked during their pregnancy seem to be more likely to use substances compared to those whose parent did not. Among the parents who smoked, the timing of SDP does not seem to be associated with whether or not the children used substances, except in the case of e-cigarettes/vape where the children of those who smoked at 22 weeks pregnant were more likely to use them than the children of those who smoked at 16 or 32 weeks.

Table 9 summarizes the frequency of substance use by child by timing of SDP. In the case of marijuana, the children of those in the non-smoking group had a greater median frequency of usage. Otherwise, the children of those in the smoking group had a greater median frequency of usage than the children of those in the smoking group. There does not seem to be an association between timing of SDP and frequency of substance usage except in the case of alcohol, where the children of those who smoked at 22 weeks had the lowest median frequency of usage compared to the other two timings.

Table 8. Substance Use (Ever) by Timing of SDP

	At 16 Weeks		At 22 Weeks		At 32 Weeks	
Characteristic	**No**, N = 36[1]	**Yes**, N = 12[1]	**No**, N = 29[1]	**Yes**, N = 13[1]	**No**, N = 30[1]	**Yes**, N = 10[1]
Cigarette	0 (0%)	1 (11%)	0 (0%)	1 (10%)	0 (0%)	1 (13%)
E-Cigarette or Vape	2 (7.1%)	1 (11%)	1 (4.5%)	2 (20%)	1 (4.2%)	1 (13%)
Marijuana	1 (3.6%)	2 (22%)	1 (4.5%)	2 (20%)	1 (4.2%)	2 (25%)
Alcohol	3 (11%)	2 (25%)	2 (9.1%)	3 (33%)	3 (13%)	2 (29%)

[1] n (%)

Table 9. Frequency of Substance Use by Timing of SDP

	At 16 Weeks		At 22 Weeks		At 32 Weeks	
Characteristic	**No**, N = 36[1]	**Yes**, N = 12[1]	**No**, N = 29[1]	**Yes**, N = 13[1]	**No**, N = 30[1]	**Yes**, N = 10[1]
Cigarette	NA (NA, NA)	0 (0, 0)	NA (NA, NA)	0 (0, 0)	NA (NA, NA)	0 (0, 0)

[1] Median (IQR)

Table 9. Frequency of Substance Use by Timing of SDP

Characteristic	At 16 Weeks		At 22 Weeks		At 32 Weeks	
	No, N = 36[1]	**Yes**, N = 12[1]	**No**, N = 29[1]	**Yes**, N = 13[1]	**No**, N = 30[1]	**Yes**, N = 10[1]
E-Cigarette or Vape	0 (0, 0)	2 (2, 2)	0 (0, 0)	2 (2, 2)	NA (NA, NA)	2 (2, 2)
Marijuana	18 (18, 18)	8 (5, 10)	18 (18, 18)	8 (5, 10)	18 (18, 18)	8 (5, 10)
Alcohol	0 (0, 0)	6 (3, 8)	0 (0, 0)	1 (1, 6)	0 (0, 0)	6 (3, 8)

[1] Median (IQR)

Self-Regulation, Internalizing Problems, Externalizing Problems, Attention Problems, ASD and ADHD by SDP Timing

Table 10 summarizes self-regulation, IP, EP, AP, ASD and ADHD by SDP timing. With regard to self-regulation, those whose parents smoked during pregnancy seemed to fare worse (lower median cognitive reappraisal score and higher median expressive suppression score) than their peers whose parents did not smoke during pregnancy. With regard to timing, children whose parent smoked at 22 weeks fared the worst compared to children of non-smoking parents at 22 weeks.

With regard to internalizing problems scores, children of smoking parents had both higher median self-reported and parent-reported scores compared to the children of non-smoking parents. A similar pattern can be observed for both externalizing problems scores and attention problems scores.

The timing of SDP did not seem to be associated with internalizing problems scores. However, the timing of SDP did seem to be associated with both self-reported and parent-reported externalizing problems scores, and with parent-reported attention problems scores.

Considering ASD, those with parents in the non-smoking group were more likely to be ASD-diagnosed or suspected. Regarding ADHD, those with parents who smoked during pregnancy are more likely to have the hyperactive/impulsive type regardless of timing of SDP, and are more likely to have the inattentive type, but only in the case of smoking at 16 weeks. There seems to be a mild association between the timing of SDP and likely ADHD.

Table 10. Self-Regulation, Internalizing Problems, Externalizing Problems, Attention Problems, ASD and ADHD in Child by Timing of SDP

Characteristic	At 16 Weeks		At 22 Weeks		At 32 Weeks	
	No, N = 36[1]	**Yes**, N = 12[1]	**No**, N = 29[1]	**Yes**, N = 13[1]	**No**, N = 30[1]	**Yes**, N = 10[1]
Self-Regulation						
Average CR Score	3.08 (2.75, 3.83)	3.00 (3.00, 3.92)	3.25 (2.83, 4.00)	3.00 (3.00, 3.83)	3.08 (2.83, 3.88)	3.00 (3.00, 3.42)
Average ES Score	2.50 (2.00, 3.13)	2.75 (2.50, 3.50)	2.50 (2.00, 3.00)	3.00 (2.50, 3.69)	2.50 (2.13, 3.25)	2.63 (2.50, 3.56)
Internalizing Problems						
Self-Reported Score	2 (1, 5)	2 (0, 3)	2 (1, 4)	3 (1, 4)	2 (0, 5)	3 (2, 3)
Parent-Reported Score	1 (0, 3)	2 (1, 4)	1 (0, 4)	2 (1, 4)	1 (0, 4)	3 (1, 5)
Externalizing Problems						
Self-Reported Score	2 (1, 4)	3 (1, 4)	3 (1, 4)	4 (2, 6)	2 (1, 4)	4 (3, 5)
Parent-Reported Score	1 (0, 2)	2 (0, 4)	1 (0, 2)	2 (0, 4)	1 (0, 2)	3 (0, 4)
Attention Problems						
Self-Reported Score	2 (0, 4)	5 (2, 7)	2 (0, 5)	5 (2, 7)	2 (0, 5)	5 (2, 7)
Parent-Reported Score	1 (0, 2)	2 (1, 6)	1 (0, 2)	2 (1, 6)	1 (0, 2)	4 (1, 6)
ASD in Child						
No	15 (88%)	3 (100%)	11 (92%)	4 (100%)	12 (92%)	3 (100%)
Diagnosed	1 (5.9%)	0 (0%)	1 (8.3%)	0 (0%)	1 (7.7%)	0 (0%)
Suspected	1 (5.9%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
ADHD in Child						
Likely ADHD-H	13 (50%)	7 (78%)	11 (55%)	8 (80%)	11 (52%)	6 (75%)
Likely ADHD-I	25 (83%)	8 (89%)	21 (91%)	9 (90%)	22 (88%)	7 (88%)

[1] Median (IQR); n (%)

Associations Between Timing of Postpartum Smoking and Outcomes

Substance Use by Postpartum Smoking Timing

Table 11 summarizes substance use in child by whether the parent smoked postpartum. Overall, it can be seen that the children whose parent smoked postpartum were more likely to try using substances.

Table 12 summarizes the frequency of substance use in child by whether the parent smoked postpartum. It can be seen that in a lot of the cases, there is not enough information to draw a conlusion (indicated by NA's). But where there is enough information (i.e. alcohol use), it can be seen that children whose parent smoked postpartum had a greater median frequency of substance usage.

Table 11. Substance Use (Ever) by Timing of Postpartum Smoking

	1st PP Visit		2nd PP Visit		At 12 Weeks PP		At 6 Months PP	
Characteristic	**No**, N = 7[1]	**Yes**, N = 3[1]	**No**, N = 22[1]	**Yes**, N = 7[1]	**No**, N = 30[1]	**Yes**, N = 12[1]	**No**, N = 24[1]	**Yes**, N = 16[1]
Cigarette	0 (0%)	0 (0%)	0 (0%)	1 (20%)	0 (0%)	1 (11%)	0 (0%)	1 (8.3%)
E-Cigarette or Vape	1 (17%)	0 (0%)	0 (0%)	1 (20%)	1 (3.8%)	2 (22%)	1 (5.0%)	2 (17%)
Marijuana	0 (0%)	1 (50%)	1 (5.3%)	1 (20%)	1 (3.8%)	2 (22%)	0 (0%)	3 (25%)
Alcohol	2 (33%)	1 (100%)	1 (5.3%)	1 (20%)	2 (7.7%)	3 (38%)	1 (5.0%)	4 (36%)

[1] n (%)

Table 12. Frequency of Substance Use by Timing of Postpartum Smoking

	1st PP Visit		2nd PP Visit		At 12 Weeks PP		At 6 Months PP	
Characteristic	**No**, N = 7[1]	**Yes**, N = 3[1]	**No**, N = 22[1]	**Yes**, N = 7[1]	**No**, N = 30[1]	**Yes**, N = 12[1]	**No**, N = 24[1]	**Yes**, N = 16[1]
Cigarette	NA (NA, NA)	NA (NA, NA)	NA (NA, NA)	0 (0, 0)	NA (NA, NA)	0 (0, 0)	NA (NA, NA)	0 (0, 0)
E-Cigarette or Vape	NA (NA, NA)	NA (NA, NA)	NA (NA, NA)	2 (2, 2)	0 (0, 0)	2 (2, 2)	0 (0, 0)	2 (2, 2)
Marijuana	NA (NA, NA)	12 (12, 12)	18 (18, 18)	3 (3, 3)	18 (18, 18)	8 (5, 10)	NA (NA, NA)	12 (8, 15)
Alcohol	0 (0, 0)	1 (1, 1)	0 (0, 0)	10 (10, 10)	0 (0, 0)	1 (1, 6)	0 (0, 0)	1 (1, 6)

[1] Median (IQR)

Self-Regulation, Internalizing Problems, Externalizing Problems, Attention Problems, ASD and ADHD

Table 13 summarizes the self-regulation, IP, EP, AP, ASD and ADHD in child by timing of parent's postpartum smoking. In terms of self-regulation, those with non-smoking parents tended to fare better or same as those with smoking parents except in the 1st Postpartum Visit group. In terms of internalizing problems score, the children of the non-smoking group fared better or same except in the case of self-reported scores for the 1st Postpartum Visit group. For the externalizing problems score, again, the children of the non-smoking group fared better or same except in the case of parent-reported scores in the 2nd Postpartum Visit group. With regard to the attention problems score, the children of the non-smoking group fared better or the same as the smoking group with no exceptions.

Children of those in the non-smoking group were more likely to be ASD-diagnosed or suspected. Children of those in the non-smoking group were less likely to have ADHD of the inattentive type, while children of those only in the non-smoking group at 12 weeks and at 6 months postpartum were less likely to have ADHD of the hyperactive/impulsive type.

Table 13. Self-Regulation, Internalizing Problems, Externalizing Problems, Attention Problems, ASD and ADHD in Child by Timing of Postpartum Smoking

	1st PP Visit		2nd PP Visit		At 12 Weeks PP		At 6 Months PP	
Characteristic	**No**, N = 7[1]	**Yes**, N = 3[1]	**No**, N = 22[1]	**Yes**, N = 7[1]	**No**, N = 30[1]	**Yes**, N = 12[1]	**No**, N = 24[1]	**Yes**, N = 16[1]
Self-Regulation								
Average CR Score	3.00 (2.88, 3.63)	3.92 (3.46, 4.38)	3.33 (2.83, 4.00)	3.00 (2.96, 3.00)	3.25 (2.83, 3.96)	3.00 (3.00, 3.21)	3.00 (2.71, 3.71)	3.00 (3.00, 4.58)
Average ES Score	3.25 (2.69, 3.81)	2.50 (2.38, 2.63)	2.50 (2.00, 3.00)	2.50 (2.50, 3.50)	2.50 (2.00, 3.00)	2.75 (2.50, 3.75)	2.50 (2.13, 3.00)	2.63 (2.44, 3.75)
Internalizing Problems								
Self-Reported Score	4 (1, 7)	3 (2, 3)	2 (0, 4)	3 (2, 4)	2 (1, 4)	3 (2, 4)	2 (0, 5)	2 (2, 4)
Parent-Reported Score	2 (0, 2)	3 (3, 4)	1 (0, 4)	1 (1, 4)	1 (0, 3)	3 (1, 4)	1 (0, 4)	2 (1, 4)
Externalizing Problems								
Self-Reported Score	2 (2, 4)	6 (5, 6)	2 (1, 4)	3 (3, 4)	2 (1, 4)	4 (3, 6)	2 (1, 4)	3 (2, 4)
Parent-Reported Score	0 (0, 1)	3 (2, 3)	1 (0, 3)	0 (0, 4)	0 (0, 2)	3 (0, 5)	0 (0, 1)	2 (0, 4)
Attention Problems								
Self-Reported Score	3 (0, 5)	8 (8, 9)	2 (0, 5)	2 (2, 5)	2 (0, 4)	5 (2, 7)	2 (0, 5)	3 (2, 6)
Parent-Reported Score	1 (0, 1)	6 (5, 6)	1 (0, 2)	1 (1, 2)	1 (0, 2)	2 (1, 6)	1 (0, 2)	2 (1, 4)
ASD in Child								
No	2 (100%)	1 (100%)	11 (92%)	2 (100%)	12 (86%)	4 (100%)	7 (78%)	8 (100%)
Diagnosed	0 (0%)	0 (0%)	1 (8.3%)	0 (0%)	1 (7.1%)	0 (0%)	1 (11%)	0 (0%)
Suspected	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (7.1%)	0 (0%)	1 (11%)	0 (0%)
ADHD in Child								
Likely ADHD-H	4 (80%)	2 (100%)	8 (50%)	3 (60%)	12 (55%)	7 (70%)	9 (47%)	8 (80%)
Likely ADHD-I	5 (100%)	2 (100%)	17 (85%)	4 (80%)	21 (81%)	9 (90%)	18 (86%)	10 (91%)

[1] Median (IQR); n (%)

Association Between SDP Dosage and Outcomes

SDP Dosage and Substance Use

Table 14 summarizes the substance use in child by SDP dosage (measured by parent's UC levels at 34 weeks gestation). UC levels are usually less than 10 ng/mL in non-smokers, between 11-30 ng/mL in light smokers or those exposed to second-hand smoke, and more than 500 ng/mL in those who are heavy smokers. There do not seem to be any heavy smokers during pregnancy in our data. In the case of cigarette and e-cigarette/vape use, the dosage of SDP exposure seems to be associated with how likely the children are to use them. In the case of marijuana and alcohol, no clear pattern can be deciphered.

Table 15 summarizes the frequency of substance use in child by SDP dosage. There is too much missing data in the case of cigarette, e-cigarette/vape and marijuana use to decipher any associations. In the case of alcohol use, those with moderate prenatal smoke exposure (UC level of 31-500 ng/mL) have the highest median frequency of use.

Table 14. Substance Use (Ever) by SDP Dosage

Characteristic	**Less than 11**, N = 25[1]	**11 to 30**, N = 2[1]	**31 to 500**, N = 10[1]	**Above 500**, N = 0[1]
Cigarette	0 (0%)	0 (0%)	1 (13%)	0 (NA%)
E-Cigarette or Vape	0 (0%)	1 (50%)	1 (13%)	0 (NA%)
Marijuana	1 (5.0%)	0 (0%)	2 (25%)	0 (NA%)
Alcohol	2 (10%)	1 (50%)	2 (29%)	0 (NA%)

[1] n (%)

Table 15. Frequency of Substance Use by SDP Dosage

Characteristic	Less than 11, N = 25[1]	11 to 30, N = 2[1]	31 to 500, N = 10[1]	Above 500, N = 0[1]
Cigarette	NA (NA, NA)	NA (NA, NA)	0 (0, 0)	NA (NA, NA)
E-Cigarette or Vape	NA (NA, NA)	NA (NA, NA)	2 (2, 2)	NA (NA, NA)
Marijuana	18 (18, 18)	NA (NA, NA)	8 (5, 10)	NA (NA, NA)
Alcohol	0 (0, 0)	0 (0, 0)	6 (3, 8)	NA (NA, NA)

[1] Median (IQR)

Association between SDP Dosage and Self-Regulation, Internalizing Problems, Externalizing Problems, Attention Problems, ASD and ADHD

Table 16 summarizes self-regulation, IP, EP, AP, ASD and ADHD by SDP dosage. With regard to self-regulation, no pattern can be seen for the CR score, but the median ER score seems to be highest for the the light exposure group (11-30 ng/mL UC). The median self-reported and parent-reported scores for both internalizing and externalizing problems seem to be highest for the light exposure group. The median attention problems score increases with increase in SDP dosage. ASD diagnosis rate is highest in the low exposure group (less than 11 ng/mL UC). Both types of ADHD are most likely in the light exposure group, and they are more likely in the moderate exposure group than in the low exposure group.

Table 16. Self-Regulation, Internalizing Problems, Externalizing Problems, Attention Problems, ASD and ADHD in Child by SDP Dosage

Characteristic	Less than 11, N = 25[1]	11 to 30, N = 2[1]	31 to 500, N = 10[1]	Above 500, N = 0[1]
Self-Regulation				
Average Cognitive Reappraisal Score	3.08 (2.67, 3.88)	3.00 (3.00, 3.00)	3.00 (3.00, 3.42)	NA (NA, NA)
Average Expressive Suppression Score	2.50 (2.00, 3.06)	4.50 (4.50, 4.50)	2.63 (2.50, 3.56)	NA (NA, NA)
Internalizing Problems				
Self-Reported Score	2 (0, 5)	12 (12, 12)	3 (2, 3)	NA (NA, NA)
Parent-Reported Score	1 (0, 2)	5 (3, 6)	3 (1, 5)	NA (NA, NA)
Externalizing Problems				
Self-Reported Score	2 (1, 4)	6 (5, 6)	4 (3, 5)	NA (NA, NA)
Parent-Reported Score	0 (0, 2)	6 (3, 8)	3 (0, 4)	NA (NA, NA)
Attention Problems				
Self-Reported Score	2 (0, 5)	5 (4, 6)	5 (2, 7)	NA (NA, NA)
Parent-Reported Score	1 (0, 2)	2 (2, 2)	4 (1, 6)	NA (NA, NA)
ASD in Child				
No	9 (90%)	2 (100%)	3 (100%)	0 (NA%)
Diagnosed	1 (10%)	0 (0%)	0 (0%)	0 (NA%)
Suspected	0 (0%)	0 (0%)	0 (0%)	0 (NA%)
ADHD in Child				
Likely ADHD - Hyperactive/Impulsive	8 (47%)	2 (100%)	6 (75%)	0 (NA%)
Likely ADHD - Inattentive	18 (86%)	2 (100%)	7 (88%)	0 (NA%)

[1] Median (IQR); n (%)

Interrelatedness of Prenatal and Postnatal Exposure

Self-Report Variables

Table 17 summarizes postnatal exposure to tobacco smoke by timing of SDP exposure. For all three timings of SDP exposure and all postnatal exposure variables, the child is more likely to be exposed to tobacco smoke after birth if they were also exposed to SDP. This association appears to be strongest between the parent smoking during pregnancy (all three timings) and the parent smoking at 12 weeks and 6 months postpartum.

Table 17. Postnatal Smoke Exposure by Exposure to SDP

Characteristic	At 16 Weeks		At 22 Weeks		At 32 Weeks	
	No, N = 36[1]	**Yes**, N = 12[1]	**No**, N = 29[1]	**Yes**, N = 13[1]	**No**, N = 30[1]	**Yes**, N = 10[1]
Parent Smoker at First Postpartum Visit	0 (0%)	3 (75%)	0 (0%)	3 (60%)	0 (0%)	3 (100%)
Parent Smoker at Second Postpartum Visit	0 (0%)	7 (88%)	0 (0%)	7 (88%)	0 (0%)	6 (86%)
Parent Smoker at 12 Weeks Postpartum	2 (6.5%)	10 (91%)	1 (3.8%)	11 (92%)	2 (7.1%)	9 (100%)
Parent Smoker at 6 Months Postpartum	6 (21%)	10 (91%)	5 (19%)	11 (92%)	5 (20%)	9 (100%)
Smoke-Exposed Between 0 to 6 Months	6 (20%)	4 (44%)	5 (22%)	5 (50%)	6 (24%)	4 (50%)
Smoke-Exposed Between 7 to 12 Months	5 (17%)	4 (44%)	3 (13%)	5 (50%)	5 (20%)	4 (50%)
Smoke-Exposed in 2nd Year	6 (20%)	5 (56%)	3 (13%)	6 (60%)	5 (20%)	5 (63%)
Smoke-Exposed in 3rd Year	6 (21%)	5 (56%)	3 (14%)	6 (60%)	5 (21%)	5 (63%)
Smoke-Exposed in 4th Year	5 (17%)	5 (56%)	2 (9.1%)	6 (60%)	4 (17%)	5 (63%)
Smoke-Exposed in 5th Year	5 (17%)	5 (56%)	2 (8.7%)	6 (60%)	4 (16%)	5 (63%)

[1] n (%)

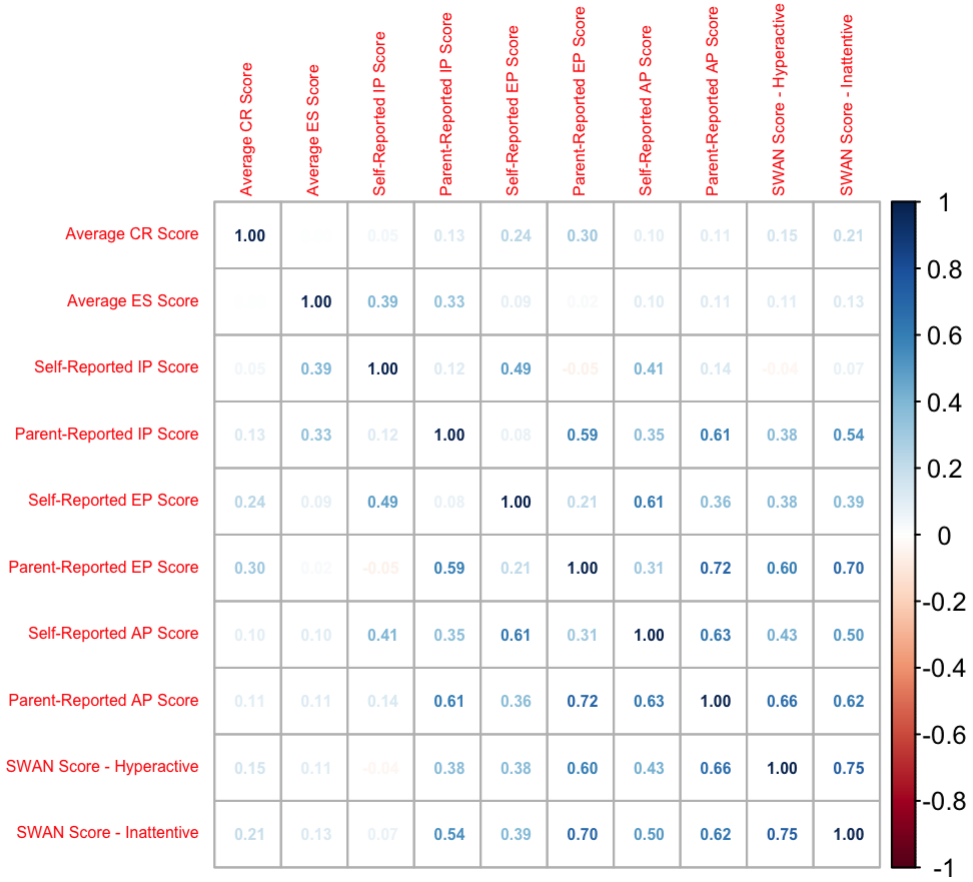
Urine Cotinine

The correlation between the UC levels at 34 weeks gestation, and 6 months postpartum from baby was found to be 0.54. This indicates a moderate association between prenatal and postnatal exposure to tobacco smoke.

Interrelations Among Self-Regulation Data

Figure 3 shows the correlations between the different self-regulation variables. Parent-reported EP and AP scores, parent-reported EP score and SWAN score on inattentive items, and the SWAN scores on hyperactive and inattentive items seem to be the most strongly positively correlated (correlation coefficient >= 0.7).

Figure 3. Correlations Between Self-Regulation Variables



Limitations

One of the main limitations of this analysis is the small sample size (n = 49). Any conclusions drawn from such a small sample size might not hold for the whole population due to sampling bias. It would be better to repeat this analysis after more data is collected to determine more concrete and reliable associations. Another main limitation is the high rate of missingness in the data. Coupled with the small sample size, this further limits the credibility of any associations drawn from this data.

References

Nicotine Cotinine (Urine). https://www.urmc.rochester.edu/encyclopedia/content.aspx?contentid=nicotine_cotinine&contenttypeid=167# (https://www.urmc.rochester.edu/encyclopedia/content.aspx?contentid=nicotine_cotinine&contenttypeid=167#)

Code Appendix

```
knitr::opts_chunk$set(echo = TRUE)
knitr::opts_knit$set(root.dir = "~/Downloads")

#Packages
install.packages("remotes", repos = "http://cran.us.r-project.org")
remotes::install_github("ddsjoberg/bstfun")

library(tidyverse)
library(naniar)
library(readr)
library(gtsummary)
library(kableExtra)
library(ggplot2)
library(stringr)
library(labelled)
library(bstfun)
library(forcats)
library(ggpubr)
library(corrplot)
#loading in the data
df <- read.csv("project1.csv")
#Data cleaning

lapply(df, class)

#One of the values of the mom_numcig variable is a range. We convert this to a mean of t
he low and high of the range. Another value of this column is "None". We change this to
0.

df[df$mom_numcig == "20-25", "mom_numcig"] <- mean(c(20,25))
df[df$mom_numcig == "None", "mom_numcig"] <- 0

#We will change the income variable and mom_numcig from character to numeric.
char_cols <- c("income", "mom_numcig")
df[char_cols] <- lapply(df[char_cols], parse_number)

#0's on SWAN inattentive and hyperactive are actually NA's
df$swan_hyperactive[df$swan_hyperactive == 0] <- NA
df$swan_inattentive[df$swan_inattentive == 0] <- NA

#Transforming the SWAN variables so that they indicate if the child is likely hyperactiv
e/inattentive
df$swan_hyperactive[df$swan_hyperactive < 6] <- 0
df$swan_hyperactive[df$swan_hyperactive >= 6] <- 1
df$swan_inattentive[df$swan_inattentive < 6] <- 0
df$swan_inattentive[df$swan_inattentive >= 6] <- 1

#Transforming SDP and ETS variables so that they can be summarized easier
sdp_ets <- c("mom_smoke_16wk", "mom_smoke_22wk", "mom_smoke_32wk", "mom_smoke_pp1", "mom
_smoke_pp2", "mom_smoke_pp12wk", "mom_smoke_pp6mo")
```

```
ind_func <- function(col){
  col_func <- df[,col]
  col_func <- case_when(col_func == "2=No" ~ 0,
                        col_func == "1=Yes" ~ 1)
  return(col_func)
}

df[sdp_ets] <- lapply(sdp_ets, ind_func)

#Turning any remaining blank values into NAs
df[df == ""] <- NA

#Change variables to factors and recode their levels
factors <- c("psex", "employ", "pedu", "tsex", "childasd")
df[,factors] <- lapply(factors, function(x){df[,x] <- as.factor(df[,x])})

df$psex <- fct_recode(df$psex,
                      "Male" = "0",
                      "Female" = "1")
df$employ <- fct_recode(df$employ,
                        "No" = "0",
                        "Part-Time" = "1",
                        "Full-Time" = "2")
df$pedu <- fct_recode(df$pedu,
                      "Some High School" = "0",
                      "High School" = "1",
                      "GED" = "2",
                      "Some College" = "3",
                      "2 Year Degree" = "4",
                      "4 Year Degree" = "5",
                      "Postgraduate Degree" = "6")
df$tsex <- fct_recode(df$tsex,
                      "Male" = "0",
                      "Female" = "1")
df$childasd <- fct_recode(df$childasd,
                          "No" = "0",
                          "Diagnosed" = "1",
                          "Suspected" = "2")

#Convert "Prefer Not to Say" to NA
df$tethnic[df$tethnic == 2] <- NA

#Create variable labels
var_label(df) <- list(
  parent_id = "Parent ID",
  page = "Parent Age",
  psex = "Parent Sex",
  plang = "Another Language Spoken by Parent at Home",
  pethnic = "Parent Hispanic/Latino",
  paian = "Parent American Indian/Alaskan Native",
  pasian = "Parent Asian",
```

```
  pnhpi = "Parent Native Hawaiian or Pacific Islander",
  pblack = "Parent Black",
  pwhite = "Parent White",
  prace_other = "Parent Other Race",
  employ = "Parent Employed",
  pedu = "Highest Level of Education of Parent",
  income = "Estimated Annual Household Income",
  childasd = "ASD in Child",
  nidaalc = "Parent Alcohol Use in the Past 6 Months",
  nidatob = "Parent Tobacco Product Use in the Past 6 Months",
  nidapres = "Parent Prescription Drug Use for Non-Medical Reasons in the Past 6 Month
s",
  nidaill = "Parent Illegal Drug Use in the Past 6 Months",
  momcig = "Number of Days Parent Smoked in the Past 30 Days",
  mom_numcig = "Cigarettes Smoked per Day by Parent",
  mom_smoke_16wk = "Parent Smoker at 16 Weeks Pregnant",
  mom_smoke_22wk = "Parent Smoker at 22 Weeks Pregnant",
  mom_smoke_32wk = "Parent Smoker at 32 Weeks Pregnant",
  mom_smoke_pp1 = "Parent Smoker at First Postpartum Visit",
  mom_smoke_pp2 = "Parent Smoker at Second Postpartum Visit",
  mom_smoke_pp12wk = "Parent Smoker at 12 Weeks Postpartum",
  mom_smoke_pp6mo = "Parent Smoker at 6 Months Postpartum",
  cotimean_34wk = "Urine Cotinine at 34 Weeks Gestation",
  cotimean_pp6mo_baby = "Urine Cotinine at 6 Months Postpartum (Baby)",
  cotimean_pp6mo = "Urine Cotinine at 6 Months Postpartum (Parent)",
  swan_inattentive = "Child Likely ADHD - Inattentive",
  swan_hyperactive = "Child Likely ADHD - Hyperactive/Impulsive",
  bpm_att_p = "Parent-Reported Attention Problems Score (BPM)",
  bpm_ext_p = "Parent-Reported Externalizing Problems Score (BPM)",
  bpm_int_p = "Parent-Reported Internalizing Problems Score (BPM)",
  smoke_exposure_6mo = "Smoke-Exposed Between 0 to 6 Months",
  smoke_exposure_12mo = "Smoke-Exposed Between 7 to 12 Months",
  smoke_exposure_2yr = "Smoke-Exposed in 2nd Year",
  smoke_exposure_3yr = "Smoke-Exposed in 3rd Year",
  smoke_exposure_4yr = "Smoke-Exposed in 4th Year",
  smoke_exposure_5yr = "Smoke-Exposed in 5th Year",
  ppmq_parental_knowledge = "Parent-Reported Average Response on Parental Knowledge (PK
Q)",
  ppmq_child_disclosure = "Parent-Reported Average Response on Child Disclosure (PKQ)",
  ppmq_parental_solicitation = "Parent-Reported Average Response on Parental Solicitatio
n (PKQ)",
  ppmq_parental_control = "Parent-Reported Average Response on Parental Control (PKQ)",
  bpm_att_a = "Self-Reported Attention Score for Parent (BPM)",
  bpm_ext_a = "Self-Reported Externalizing Problems Score for Parent (BPM)",
  bpm_int_a = "Self-Reported Internalizing Problems Score for Parent (BPM)",
  erq_cog_a = "Self-Reported Average Cognitive Reappraisal Score for Parent (ERQ)",
  erq_exp_a = "Self-Reported Average Expressive Suppression Score for Parent (ERQ)",
  tage = "Child Age",
  tsex = "Child Sex",
  language = "Another Language Spoken by Child at Home",
  tethnic = "Child Hispanic/Latino",
  taian = "Child American Indian/Alaskan Native",
```

```
  tasian = "Child Asian",
  tnhpi = "Child Native Hawaiian or Pacific Islander",
  tblack = "Child Black",
  twhite = "Child White",
  trace_other = "Child Other Race",
  cig_ever = "Cigarette Smoked Ever",
  num_cigs_30 = "Cigarette Use Over Past 30 Days (Child)",
  e_cig_ever = "E-Cigarette or Vape Used Ever",
  num_e_cigs_30 = "E-Cigarette or Vape Use Over Past 30 Days (Child)",
  mj_ever = "Marijuana Used Ever",
  num_mj_30 = "Marijuana Use Over Past 30 Days (Child)",
  alc_ever = "Alcohol Ever Consumed by Child (Not as Part of Religious Ceremonies)",
  num_alc_30 = "Alcohol Consumption Over Past 30 Days (Child)",
  bpm_att = "Self-Reported Attention Problems Score (BPM)",
  bpm_ext = "Self-Reported Externalizing Problems Score (BPM)",
  bpm_int = "Self-Reported Internalizing Problems Score (BPM)",
  erq_cog = "Average Cognitive Reappraisal Score of Child (ERQ)",
  erq_exp = "Average Expressive Suppression Score of Child (ERQ)",
  pmq_parental_knowledge = "Child-Reported Average Response on Parental Knowledge (PK
Q)",
  pmq_child_disclosure = "Child-Reported Average Response on Child Disclosure (PKQ)",
  pmq_parental_solicitation = "Child-Reported Average Response on Parental Solicitation
(PKQ)",
  pmq_parental_control = "Child-Reported Average Response on Parental Control (PKQ)"
)

#Remove variables not used in analysis
df <- df %>%
  select(-c(nidaalc, nidatob, nidapres, nidaill, momcig, mom_numcig, erq_cog_a, erq_exp_
a, pmq_parental_knowledge, pmq_child_disclosure, pmq_parental_solicitation, pmq_parental
_control, ppmq_parental_knowledge, ppmq_child_disclosure, ppmq_parental_solicitation, pp
mq_parental_control, bpm_att_a, bpm_int_a, bpm_ext_a))

#Missing values
label_func <- function(var){
  return(attr(df[,var], "label"))
}

### Exploratory Data Analysis

#### A Summary of the Missing Data
df_miss <- df
colnames(df_miss) <- lapply(colnames(df_miss), label_func)

gg_miss_var(df_miss, show_pct = TRUE) +
  theme(axis.text.y = element_text(size = 5)) +
  ggtitle("Figure 1. Percent Missing Observations by Variable") +
  theme(plot.title = element_text(size = 5))
#### Demographics
var_label(df) <- list(
  pethnic = "Hispanic/Latino",
  paian = "American Indian/Alaskan Native",
```

```
  pasian = "Asian",
  pnhpi = "Native Hawaiian or Pacific Islander",
  pblack = "Black",
  pwhite = "White",
  prace_other = "Other Race",
  tethnic = "Hispanic/Latino",
  taian = "American Indian/Alaskan Native",
  tasian = "Asian",
  tnhpi = "Native Hawaiian or Pacific Islander",
  tblack = "Black",
  twhite = "White",
  trace_other = "Other Race"
)

#Tables summarizing parent demographics
pdems <- c("page", "psex", "plang", "employ", "pedu", "income")
prace_vars <- c("pethnic", "paian", "pasian", "pnhpi", "pblack", "pwhite", "prace_othe
r")

theme_gtsummary_compact(set_theme=TRUE, font_size = 10)

df %>%
  select(all_of(prace_vars), all_of(pdems)) %>%
  tbl_summary(missing = "no") %>%
  add_variable_grouping("Parent Race" = prace_vars) %>%
  as_gt() %>%
  gt::tab_header(title = "Table 1. Parent Demographics")

#Table summarizing child demographics
cdems <- c("tage", "tsex", "language")
crace_vars <- c("tethnic", "taian", "tasian", "tnhpi", "tblack", "twhite", "trace_othe
r")

df %>%
  select(all_of(crace_vars), all_of(cdems)) %>%
  tbl_summary(missing = "no", type = list(tage ~ "continuous")) %>%
  add_variable_grouping("Child Race" = crace_vars) %>%
  as_gt() %>%
  gt::tab_header(title = "Table 2. Child Demographics")
#### Substance Use in Child

su_ever_child <- c("cig_ever", "e_cig_ever", "mj_ever", "alc_ever")
su_dosage_child <- c("num_cigs_30", "num_e_cigs_30", "num_mj_30", "num_alc_30")

var_label(df) <- list(
  cig_ever = "Cigarette",
  e_cig_ever = "E-Cigarette or Vape",
  mj_ever = "Marijuana",
  alc_ever = "Alcohol (Not as Part of Religious Ceremonies)",
  num_cigs_30 = "Cigarette",
  num_e_cigs_30 = "E-Cigarette or Vape",
  num_mj_30 = "Marijuana",
```

```
  num_alc_30 = "Alcohol"
)

df %>%
  select(all_of(su_ever_child), all_of(su_dosage_child)) %>%
  tbl_summary(missing = "no",
              type = list(c("num_cigs_30", "num_e_cigs_30", "num_mj_30", "num_alc_30") ~
"continuous")) %>%
  add_variable_grouping("Substance Used (Ever)" = su_ever_child, "Number of Days Used in
the Past 30 Days" = su_dosage_child) %>%
  as_gt() %>%
  gt::tab_header(title = "Table 3. Substance Use in Child")
#### Self-Regulation, Internalizing Problems, Externalizing Problems, Attention Problem
s, ASD and ADHD in Child

sr_child <- c("erq_cog", "erq_exp")
ip_child <- c("bpm_int", "bpm_int_p")
ep_child <- c("bpm_ext", "bpm_ext_p")
ap_child <- c("bpm_att", "bpm_att_p")
adhd <- c("swan_hyperactive", "swan_inattentive")

var_label(df) <- list(
  erq_cog =  "Average Cognitive Reappraisal Score",
  erq_exp = "Average Expressive Suppression Score",
  bpm_int = "Self-Reported Score",
  bpm_int_p = "Parent-Reported Score",
  bpm_ext = "Self-Reported Score",
  bpm_ext_p = "Parent-Reported Score",
  bpm_att = "Self-Reported Score",
  bpm_att_p = "Parent-Reported Score",
  swan_hyperactive = "Likely ADHD - Hyperactive/Impulsive",
  swan_inattentive = "Likely ADHD - Inattentive"
)

df %>%
  select(all_of(sr_child), all_of(ip_child), all_of(ep_child), all_of(ap_child), childas
d, all_of(adhd)) %>%
  tbl_summary(missing = "no",
              type = list(c(bpm_ext, bpm_ext_p, bpm_att, bpm_att_p) ~ "continuous")) %>%
  add_variable_grouping("Self-Regulation" = sr_child,
                        "Internalizing Problems" = ip_child,
                        "Externalizing Problems" = ep_child,
                        "Attention Problems" = ap_child,
                        "ADHD in Child" = adhd) %>%
  as_gt() %>%
  gt::tab_header(title = "Table 4. Self-Regulation, Internalizing Problems, Externalizin
g Problems, Attention Problems, ASD and ADHD in Child")
#Differing distributions of parent-reported and child-reported IP, EP and AP scores
scores_df <- df %>%
  select(all_of(ip_child), all_of(ep_child), all_of(ap_child)) %>%
  pivot_longer(cols = c(all_of(ip_child), all_of(ep_child), all_of(ap_child)), names_to
= "score_type", values_to = "bpm_score") %>%
```

```
  mutate(reporter = case_when(score_type == "bpm_int_p" ~ "Parent",
                              score_type == "bpm_ext_p" ~ "Parent",
                              score_type == "bpm_att_p" ~ "Parent",
                              score_type == "bpm_int" ~ "Child (Self-Report)",
                              score_type == "bpm_ext" ~ "Child (Self-Report)",
                              score_type == "bpm_att" ~ "Child (Self-Report)"),
         score_type = case_when(score_type == "bpm_int_p" ~ "IP",
                              score_type == "bpm_ext_p" ~ "EP",
                              score_type == "bpm_att_p" ~ "AP",
                              score_type == "bpm_int" ~ "IP",
                              score_type == "bpm_ext" ~ "EP",
                              score_type == "bpm_att" ~ "AP"))

ggplot(data = scores_df, aes(x = reporter, y = bpm_score, fill = reporter)) +
  geom_violin(trim = FALSE) +
  geom_boxplot(width = .1, fill = "black") +
  stat_summary(fun.y = median, geom = "point", fill = "white", shape = 21, size = 2.5) +
  ylab("Internalizing Problems Score") +
  labs(fill = "Reported By") +
  scale_fill_brewer() +
  facet_wrap(~score_type) +
  ggtitle("Figure 2. Difference in Distributions of Child-Reported and Parent-Reported S
cores") +
  theme(axis.text.x = element_blank(), axis.ticks.x = element_blank(), axis.title.x = el
ement_blank(), title = element_text(size = 5))
#### Smoking During Pregnancy

sdp <- c("mom_smoke_16wk", "mom_smoke_22wk", "mom_smoke_32wk")

var_label(df) <- list(
  mom_smoke_16wk = "Smoker At 16 Weeks",
  mom_smoke_22wk = "Smoker At 22 Weeks",
  mom_smoke_32wk = "Smoker At 32 Weeks",
  cotimean_34wk = "UC At 34 Weeks Gestation"
)

df %>%
  select(all_of(sdp), cotimean_34wk) %>%
  tbl_summary(missing = "no") %>%
  as_gt() %>%
  gt::tab_header(title = "Table 5. Smoking During Pregnancy")
#### Postpartum Smoking

pp_smoke <- c("mom_smoke_pp1", "mom_smoke_pp2", "mom_smoke_pp12wk", "mom_smoke_pp6mo")

var_label(df) <- list(
  mom_smoke_pp1 = "At First Postpartum Visit",
  mom_smoke_pp2 = "At Second Postpartum Visit",
  mom_smoke_pp12wk = "At 12 Weeks Postpartum",
  mom_smoke_pp6mo = "At 6 Months Postpartum"
)
```

```
df %>%
  select(all_of(pp_smoke)) %>%
  tbl_summary(missing = "no") %>%
  as_gt() %>%
  gt::tab_header(title = "Table 6. Postpartum Smoking")
#### Environmental Tobacco Smoke

var_label(df) <- list(
  smoke_exposure_6mo = "0 to 6 Months",
  smoke_exposure_12mo = "7 to 12 Months",
  smoke_exposure_2yr = "2nd Year",
  smoke_exposure_3yr = "3rd Year",
  smoke_exposure_4yr = "4th Year",
  smoke_exposure_5yr = "5th Year",
  cotimean_34wk = "At 34 Weeks Gestation",
  cotimean_pp6mo = "At 6 Months Postpartum (Parent)",
  cotimean_pp6mo_baby = "At 6 Months Postpartum (Baby)"
)

#Smoke exposure
ets_self <- c("smoke_exposure_6mo", "smoke_exposure_12mo", "smoke_exposure_2yr", "smoke_
exposure_3yr", "smoke_exposure_4yr", "smoke_exposure_5yr")

#Urine cotinine
uc <- c("cotimean_pp6mo", "cotimean_pp6mo_baby")

df %>%
  select(all_of(ets_self), all_of(uc)) %>%
  tbl_summary(missing = "no") %>%
  add_variable_grouping("Smoke Exposure" = ets_self,
                        "UC levels" = uc) %>%
  as_gt() %>%
  gt::tab_header(title = "Table 7. Postpartum Environmental Tobacco Smoke Exposure")
#### Association Between Timing of SDP and Outcomes

##### Substance Use by SDP Timing

#Outcomes - Substance use
df[,sdp] <- lapply(sdp, function(x){df[,x] <- as.factor(df[,x])})

df$mom_smoke_16wk <- fct_recode(df$mom_smoke_16wk,
                                "No" = "0",
                                "Yes" =  "1")
df$mom_smoke_22wk <- fct_recode(df$mom_smoke_22wk,
                                "No" = "0",
                                "Yes" =  "1")
df$mom_smoke_32wk <- fct_recode(df$mom_smoke_32wk,
                                "No" = "0",
                                "Yes" =  "1")

var_label(df) <- list(
  alc_ever = "Alcohol"
```

```
)

#Ever Used
df %>%
  select(all_of(sdp), all_of(su_ever_child)) %>%
  pivot_longer(cols = all_of(sdp), names_to = "sdp_timing", values_to = "exposed") %>%
  mutate(sdp_timing = case_when(sdp_timing == "mom_smoke_16wk" ~ "At 16 Weeks",
            sdp_timing == "mom_smoke_22wk" ~ "At 22 Weeks",
            sdp_timing == "mom_smoke_32wk" ~ "At 32 Weeks")) %>%
  tbl_strata(strata = sdp_timing,
             .tbl_fun =
               ~ .x %>%
               tbl_summary(by = exposed, missing = "no")) %>%
  as_gt() %>%
  gt::tab_header(title = "Table 8. Substance Use (Ever) by Timing of SDP")

#Frequency of Usage
df %>%
  select(all_of(sdp), all_of(su_dosage_child)) %>%
  pivot_longer(cols = all_of(sdp), names_to = "sdp_timing", values_to = "exposed") %>%
  mutate(sdp_timing = case_when(sdp_timing == "mom_smoke_16wk" ~ "At 16 Weeks",
            sdp_timing == "mom_smoke_22wk" ~ "At 22 Weeks",
            sdp_timing == "mom_smoke_32wk" ~ "At 32 Weeks")) %>%
  tbl_strata(strata = sdp_timing,
             .tbl_fun =
               ~ .x %>%
               tbl_summary(by = exposed, missing = "no",
                           type = list(where(is.numeric) ~ "continuous"))) %>%
  as_gt() %>%
  gt::tab_header(title = "Table 9. Frequency of Substance Use by Timing of SDP")
##### Self-Regulation, Internalizing Problems, Externalizing Problems, Attention Problem
s, ASD and ADHD by SDP Timing

var_label(df) <- list(
  erq_cog =  "Average CR Score",
  erq_exp = "Average ES Score",
  bpm_int = "Self-Reported Score",
  bpm_int_p = "Parent-Reported Score",
  bpm_ext = "Self-Reported Score",
  bpm_ext_p = "Parent-Reported Score",
  bpm_att = "Self-Reported Score",
  bpm_att_p = "Parent-Reported Score",
  swan_hyperactive = "Likely ADHD-H",
  swan_inattentive = "Likely ADHD-I"
)

df %>%
  select(all_of(sdp), all_of(sr_child), all_of(ip_child), all_of(ep_child), all_of(ap_ch
ild), childasd, all_of(adhd)) %>%
  pivot_longer(cols = all_of(sdp), names_to = "sdp_timing", values_to = "exposed") %>%
  mutate(sdp_timing = case_when(sdp_timing == "mom_smoke_16wk" ~ "At 16 Weeks",
            sdp_timing == "mom_smoke_22wk" ~ "At 22 Weeks",
```

```
            sdp_timing == "mom_smoke_32wk" ~ "At 32 Weeks")) %>%
  tbl_strata(strata = sdp_timing,
             .tbl_fun =
               ~ .x %>%
               tbl_summary(by = exposed, missing = "no",
                           type = list(c(bpm_ext, bpm_ext_p, bpm_att, bpm_att_p, bpm_in
t, bpm_int_p) ~ "continuous"))) %>%
  add_variable_grouping("Self-Regulation" = sr_child,
                        "Internalizing Problems" = ip_child,
                        "Externalizing Problems" = ep_child,
                        "Attention Problems" = ap_child,
                        "ADHD in Child" = adhd) %>%
  as_gt() %>%
  gt::tab_header("Table 10. Self-Regulation, Internalizing Problems, Externalizing Probl
ems, Attention Problems, ASD and ADHD in Child by Timing of SDP")
#### Associations Between Timing of Postpartum Smoking and Outcomes

##### Substance Use by Postpartum Smoking Timing
var_label(df) <- list(
  mom_smoke_pp1 = "At First PP Visit",
  mom_smoke_pp2 = "At Second PP Visit",
  mom_smoke_pp12wk = "At 12 Weeks PP",
  mom_smoke_pp6mo = "At 6 Months PP"
)

df[,pp_smoke] <- lapply(pp_smoke, function(x){df[,x] <- as.factor(df[,x])})

df$mom_smoke_pp1 <- fct_recode(df$mom_smoke_pp1,
                               "No" = "0",
                               "Yes" =  "1")
df$mom_smoke_pp2 <- fct_recode(df$mom_smoke_pp2,
                               "No" = "0",
                               "Yes" =  "1")
df$mom_smoke_pp12wk <- fct_recode(df$mom_smoke_pp12wk,
                               "No" = "0",
                               "Yes" =  "1")
df$mom_smoke_pp6mo <- fct_recode(df$mom_smoke_pp6mo,
                               "No" = "0",
                               "Yes" =  "1")

#Ever Used
df %>%
  select(all_of(pp_smoke), all_of(su_ever_child)) %>%
  pivot_longer(cols = all_of(pp_smoke), names_to = "pp_smoke_timing", values_to = "expos
ed") %>%
  mutate(pp_smoke_timing = case_when(pp_smoke_timing == "mom_smoke_pp1" ~ "1st PP Visi
t",
            pp_smoke_timing == "mom_smoke_pp2" ~ "2nd PP Visit",
            pp_smoke_timing == "mom_smoke_pp12wk" ~ "At 12 Weeks PP",
            pp_smoke_timing == "mom_smoke_pp6mo" ~ "At 6 Months PP")) %>%
  tbl_strata(strata = pp_smoke_timing,
             .tbl_fun =
```

```
              ~ .x %>%
              tbl_summary(by = exposed, missing = "no")) %>%
  as_gt() %>%
  gt::tab_header("Table 11. Substance Use (Ever) by Timing of Postpartum Smoking")

#Frequency of Usage
df %>%
  select(all_of(pp_smoke), all_of(su_dosage_child)) %>%
  pivot_longer(cols = all_of(pp_smoke), names_to = "pp_smoke_timing", values_to = "expos
ed") %>%
  mutate(pp_smoke_timing = case_when(pp_smoke_timing == "mom_smoke_pp1" ~ "1st PP Visi
t",
            pp_smoke_timing == "mom_smoke_pp2" ~ "2nd PP Visit",
            pp_smoke_timing == "mom_smoke_pp12wk" ~ "At 12 Weeks PP",
            pp_smoke_timing == "mom_smoke_pp6mo" ~ "At 6 Months PP")) %>%
  tbl_strata(strata = pp_smoke_timing,
             .tbl_fun =
              ~ .x %>%
              tbl_summary(by = exposed, missing = "no",
                          type = list(where(is.numeric) ~ "continuous"))) %>%
  as_gt() %>%
  gt::tab_header("Table 12. Frequency of Substance Use by Timing of Postpartum Smoking")
##### Self-Regulation, Internalizing Problems, Externalizing Problems, Attention Problem
s, ASD and ADHD

var_label(df) <- list(
  erq_cog =  "Average CR Score",
  erq_exp = "Average ES Score",
  bpm_int = "Self-Reported Score",
  bpm_int_p = "Parent-Reported Score",
  bpm_ext = "Self-Reported Score",
  bpm_ext_p = "Parent-Reported Score",
  bpm_att = "Self-Reported Score",
  bpm_att_p = "Parent-Reported Score",
  swan_hyperactive = "Likely ADHD-H",
  swan_inattentive = "Likely ADHD-I"
)

df %>%
  select(all_of(pp_smoke), all_of(sr_child), all_of(ip_child), all_of(ep_child), all_of
(ap_child), childasd, all_of(adhd)) %>%
  pivot_longer(cols = all_of(pp_smoke), names_to = "pp_smoke_timing", values_to = "expos
ed") %>%
  mutate(pp_smoke_timing = case_when(pp_smoke_timing == "mom_smoke_pp1" ~ "1st PP Visi
t",
            pp_smoke_timing == "mom_smoke_pp2" ~ "2nd PP Visit",
            pp_smoke_timing == "mom_smoke_pp12wk" ~ "At 12 Weeks PP",
            pp_smoke_timing == "mom_smoke_pp6mo" ~ "At 6 Months PP")) %>%
  tbl_strata(strata = pp_smoke_timing,
             .tbl_fun =
              ~ .x %>%
              tbl_summary(by = exposed, missing = "no",
```

```
                           type = list(c(erq_cog, erq_exp, bpm_ext, bpm_ext_p, bpm_att,
bpm_att_p, bpm_int, bpm_int_p) ~ "continuous"))) %>%
  add_variable_grouping("Self-Regulation" = sr_child,
                        "Internalizing Problems" = ip_child,
                        "Externalizing Problems" = ep_child,
                        "Attention Problems" = ap_child,
                        "ADHD in Child" = adhd) %>%
  as_gt() %>%
  gt::tab_header("Table 13. Self-Regulation, Internalizing Problems, Externalizing Probl
ems, Attention Problems, ASD and ADHD in Child by Timing of Postpartum Smoking")
#### Association Between Timing of ETS and Outcomes

##### Substance Use in Child by ETS Timing

var_label(df) <- list(
  smoke_exposure_6mo = "0 to 6 Months",
  smoke_exposure_12mo = "7 to 12 Months",
  smoke_exposure_2yr = "2nd Year",
  smoke_exposure_3yr = "3rd Year",
  smoke_exposure_4yr = "4th Year",
  smoke_exposure_5yr = "5th Year",
  cotimean_34wk = "At 34 Weeks Gestation",
  cotimean_pp6mo = "At 6 Months Postpartum (Parent)",
  cotimean_pp6mo_baby = "At 6 Months Postpartum (Baby)"
)

df[,ets_self] <- lapply(ets_self, function(x){df[,x] <- as.factor(df[,x])})

df$smoke_exposure_6mo <- fct_recode(df$smoke_exposure_6mo,
                                    "No" = "0",
                                    "Yes" =  "1")
df$smoke_exposure_12mo <- fct_recode(df$smoke_exposure_12mo,
                                "No" = "0",
                                "Yes" =  "1")
df$smoke_exposure_2yr <- fct_recode(df$smoke_exposure_2yr,
                                "No" = "0",
                                "Yes" =  "1")
df$smoke_exposure_3yr <- fct_recode(df$smoke_exposure_3yr,
                                "No" = "0",
                                "Yes" =  "1")
df$smoke_exposure_4yr <- fct_recode(df$smoke_exposure_4yr,
                                "No" = "0",
                                "Yes" =  "1")
df$smoke_exposure_5yr <- fct_recode(df$smoke_exposure_5yr,
                                "No" = "0",
                                "Yes" =  "1")

#Ever Used
df %>%
  select(all_of(ets_self), all_of(su_ever_child)) %>%
  pivot_longer(cols = all_of(ets_self), names_to = "ets_timing", values_to = "exposed")
%>%
```

```
  mutate(ets_timing = case_when(ets_timing == "smoke_exposure_6mo" ~ "Months 0 to 6",
            ets_timing == "smoke_exposure_12mo" ~ "Months 7 to 12",
            ets_timing == "smoke_exposure_2yr" ~ "Year 2",
            ets_timing == "smoke_exposure_3yr" ~ "Year 3",
            ets_timing == "smoke_exposure_4yr" ~ "Year 4",
            ets_timing == "smoke_exposure_5yr" ~ "Year 5")) %>%
  tbl_strata(strata = ets_timing,
             .tbl_fun =
               ~ .x %>%
               tbl_summary(by = exposed, missing = "no")) %>%
  modify_caption("**Substance Use (Ever) by Timing of ETS**")

#Frequency of Usage
df %>%
  select(all_of(ets_self), all_of(su_dosage_child)) %>%
    pivot_longer(cols = all_of(ets_self), names_to = "ets_timing", values_to = "expose
d") %>%
  mutate(ets_timing = case_when(ets_timing == "smoke_exposure_6mo" ~ "Months 0 to 6",
            ets_timing == "smoke_exposure_12mo" ~ "Months 7 to 12",
            ets_timing == "smoke_exposure_2yr" ~ "Year 2",
            ets_timing == "smoke_exposure_3yr" ~ "Year 3",
            ets_timing == "smoke_exposure_4yr" ~ "Year 4",
            ets_timing == "smoke_exposure_5yr" ~ "Year 5")) %>%
  tbl_strata(strata = ets_timing,
             .tbl_fun =
               ~ .x %>%
               tbl_summary(by = exposed, missing = "no",
                           type = list(where(is.numeric) ~ "continuous"))) %>%
  modify_caption("**Frequency of Substance Use by Timing of ETS**")
##### Self-Regulation, Internalizing Problems, Externalizing Problems, Attention Problem
s, ASD and ADHD in Child by ETS Timing

var_label(df) <- list(
  erq_cog =  "Average Cognitive Reappraisal Score",
  erq_exp = "Average Expressive Suppression Score",
  bpm_int = "Self-Reported Score",
  bpm_int_p = "Parent-Reported Score",
  bpm_ext = "Self-Reported Score",
  bpm_ext_p = "Parent-Reported Score",
  bpm_att = "Self-Reported Score",
  bpm_att_p = "Parent-Reported Score",
  swan_hyperactive = "Likely ADHD - Hyperactive/Impulsive",
  swan_inattentive = "Likely ADHD - Inattentive"
)

df %>%
  select(all_of(ets_self), all_of(sr_child), all_of(ip_child), all_of(ep_child), all_of
(ap_child), childasd, all_of(adhd)) %>%
  pivot_longer(cols = all_of(ets_self), names_to = "ets_timing", values_to = "exposed")
%>%
  mutate(ets_timing = case_when(ets_timing == "smoke_exposure_6mo" ~ "Months 0 to 6",
            ets_timing == "smoke_exposure_12mo" ~ "Months 7 to 12",
```

```
            ets_timing == "smoke_exposure_2yr" ~ "Year 2",
            ets_timing == "smoke_exposure_3yr" ~ "Year 3",
            ets_timing == "smoke_exposure_4yr" ~ "Year 4",
            ets_timing == "smoke_exposure_5yr" ~ "Year 5")) %>%
  tbl_strata(strata = ets_timing,
             .tbl_fun =
               ~ .x %>%
               tbl_summary(by = exposed, missing = "no",
                           type = list(c(erq_cog, erq_exp, bpm_ext, bpm_ext_p, bpm_att,
bpm_att_p, bpm_int, bpm_int_p) ~ "continuous"))) %>%
  add_variable_grouping("Self-Regulation" = sr_child,
                        "Internalizing Problems" = ip_child,
                        "Externalizing Problems" = ep_child,
                        "Attention Problems" = ap_child,
                        "ADHD in Child" = adhd) %>%
  modify_caption("**Self-Regulation, Internalizing Problems, Externalizing Problems, Att
ention Problems, ASD and ADHD in Child by Timing of ETS**")
#### Association Between SDP Dosage and Outcomes

##### SDP Dosage and Substance Use
uc_df_prenatal <- df %>%
  select(cotimean_34wk, all_of(su_ever_child), all_of(su_dosage_child), all_of(sr_chil
d), all_of(ip_child), all_of(ep_child), all_of(ap_child), childasd, all_of(adhd))

#coverting to categorical
uc_df_prenatal$cotimean_34wk <- cut(uc_df_prenatal$cotimean_34wk,
      breaks = c(0, 11, 30, 500, Inf),
      labels = c("Less than 11", "11 to 30", "31 to 500", "Above 500"))

#Ever Used
uc_df_prenatal %>%
  select(cotimean_34wk, all_of(su_ever_child)) %>%
  tbl_summary(by = cotimean_34wk, missing = "no") %>%
  as_gt() %>%
  gt::tab_header("Table 14. Substance Use (Ever) by SDP Dosage")

#Frequency of Usage
uc_df_prenatal %>%
  select(cotimean_34wk, all_of(su_dosage_child)) %>%
  tbl_summary(by = cotimean_34wk, missing = "no", type = list(where(is.numeric) ~ "conti
nuous")) %>%
  as_gt() %>%
  gt::tab_header("Table 15. Frequency of Substance Use by SDP Dosage")
##### Association between SDP Dosage and Self-Regulation, Internalizing Problems, Extern
alizing Problems, Attention Problems, ASD and ADHD
var_label(uc_df_prenatal) <- list(
  erq_cog =  "Average Cognitive Reappraisal Score",
  erq_exp = "Average Expressive Suppression Score",
  bpm_int = "Self-Reported Score",
  bpm_int_p = "Parent-Reported Score",
  bpm_ext = "Self-Reported Score",
  bpm_ext_p = "Parent-Reported Score",
```

```
  bpm_att = "Self-Reported Score",
  bpm_att_p = "Parent-Reported Score",
  swan_hyperactive = "Likely ADHD - Hyperactive/Impulsive",
  swan_inattentive = "Likely ADHD - Inattentive"
)

uc_df_prenatal %>%
  select(cotimean_34wk, all_of(sr_child), all_of(ip_child), all_of(ep_child), all_of(ap_
child), childasd, all_of(adhd)) %>%
  tbl_summary(by = cotimean_34wk, missing = "no",
                            type = list(c(erq_cog, erq_exp, bpm_ext, bpm_ext_p, bpm_att,
bpm_att_p, bpm_int, bpm_int_p) ~ "continuous")) %>%
  add_variable_grouping("Self-Regulation" = sr_child,
                        "Internalizing Problems" = ip_child,
                        "Externalizing Problems" = ep_child,
                        "Attention Problems" = ap_child,
                        "ADHD in Child" = adhd) %>%
  as_gt() %>%
  gt::tab_header("Table 16. Self-Regulation, Internalizing Problems, Externalizing Probl
ems, Attention Problems, ASD and ADHD in Child by SDP Dosage")
#### Interrelatedness of Prenatal and Postnatal Exposure

##### Self-Report Variables
#Exposure self-report
var_label(df) <- list(
  mom_smoke_pp1 = "Parent Smoker at First Postpartum Visit",
  mom_smoke_pp2 = "Parent Smoker at Second Postpartum Visit",
  mom_smoke_pp12wk = "Parent Smoker at 12 Weeks Postpartum",
  mom_smoke_pp6mo = "Parent Smoker at 6 Months Postpartum",
  smoke_exposure_6mo = "Smoke-Exposed Between 0 to 6 Months",
  smoke_exposure_12mo = "Smoke-Exposed Between 7 to 12 Months",
  smoke_exposure_2yr = "Smoke-Exposed in 2nd Year",
  smoke_exposure_3yr = "Smoke-Exposed in 3rd Year",
  smoke_exposure_4yr = "Smoke-Exposed in 4th Year",
  smoke_exposure_5yr = "Smoke-Exposed in 5th Year"
)

df %>%
  select(all_of(sdp), all_of(pp_smoke), all_of(ets_self)) %>%
  pivot_longer(cols = all_of(sdp), names_to = "sdp_timing", values_to = "exposed") %>%
  mutate(sdp_timing = case_when(sdp_timing == "mom_smoke_16wk" ~ "At 16 Weeks",
            sdp_timing == "mom_smoke_22wk" ~ "At 22 Weeks",
            sdp_timing == "mom_smoke_32wk" ~ "At 32 Weeks")) %>%
  tbl_strata(strata = sdp_timing,
             .tbl_fun =
               ~ .x %>%
               tbl_summary(by = exposed, missing = "no")) %>%
  as_gt() %>%
  gt::tab_header("Table 17. Postnatal Smoke Exposure by Exposure to SDP")
##### Urine Cotinine
uc_df_prepost <- df %>%
  select(cotimean_34wk, cotimean_pp6mo_baby)
```

```
ggplot(data = uc_df_prepost, aes(x = cotimean_34wk, y = cotimean_pp6mo_baby)) +
  geom_point() +
  stat_cor(method = "pearson", label.x = -5, label.y = 40)
#### Interrelations Among Self-Regulation Data
df1 <- read.csv("project1.csv")
df$swan_hyperactive <- df1$swan_hyperactive
df$swan_inattentive <- df1$swan_inattentive

scores_data <- df %>%
  select(all_of(sr_child), all_of(ip_child), all_of(ep_child), all_of(ap_child), all_of
(adhd))

scores_data <- scores_data[complete.cases(scores_data), ]

var_label(df) <- list(
  erq_cog = "Average CR Score",
  erq_exp = "Average ES Score",
  bpm_att_p = "Parent-Reported AP Score",
  bpm_ext_p = "Parent-Reported EP Score",
  bpm_int_p = "Parent-Reported IP Score",
  bpm_att = "Self-Reported AP Score",
  bpm_ext = "Self-Reported EP Score",
  bpm_int = "Self-Reported IP Score",
  swan_inattentive = "SWAN Score - Inattentive",
  swan_hyperactive = "SWAN Score - Hyperactive"
)
colnames(scores_data) <- lapply(colnames(scores_data), label_func)

correlation_matrix <- cor(scores_data)
corrplot(correlation_matrix, method = "number", number.cex = 0.5, title = "Figure 3. Cor
relations Between Self-Regulation Variables", tl.cex = 0.5, mar=c(0,0,2,0))
```