

# Midterm Project - Report 2

Group 6: Leah Li, Xuange Liang, Zexuan Yan, Yiwen Zhang

P8110 Applied Regression II

November 20, 2025

## Data Analysis

We analyzed data from 220 offspring (aged 6–23 years) to investigate the impact of parental depression on the risk of developing depression and substance abuse. The study sample consisted of 125 children with a depressed index parent and 95 children with a non-depressed (control) parent. The primary outcomes were the age at onset of depression (BEDEPON) and substance abuse (BESUBON). For participants who did not experience an event, survival times were censored at their age at interview (PTAGE). In cases where substance abuse was diagnosed but the age of onset was missing (BESUBON = -1), the onset age was imputed as the interview age minus 0.5 years, per study protocol. A single missing value for socioeconomic class (SESCLASS) was imputed to the modal category (Class 3). Kaplan-Meier survival curves and log-rank tests were used to assess univariate associations between parental depression and the age of onset for both outcomes.

For Hypothesis 1, we fitted a Cox proportional hazards model stratified by SESCLASS to account for confounding by social class. To test the specific hypothesis of differential effects before and after age 13, we included a time-dependent covariate defined as the interaction between parental depression and an indicator of age less than 13 years. We used ‘CONTRAST’ statements to estimate hazard ratios (HRs) for parental depression specifically within the  $< 13$  and  $\geq 13$  age windows.

For Hypothesis 2, we assessed the risk of substance abuse using a non-stratified Cox model. The model adjusted for relevant covariates and a time-dependent indicator for prior depression to determine if parental depression remained a risk factor after adjusting for the child’s own depression history. Ties in survival times were handled using Efron’s method. Model fit was assessed using the Likelihood Ratio (LR) test for the global null hypothesis. All analyses were performed using SAS 9.4, with statistical significance defined at  $\alpha = 0.05$ .

## Results

The study analyzed data from 220 offspring (aged 6–23 years), consisting of 125 children with a depressed index parent and 95 children with a non-depressed (control) parent. The primary outcomes were the age at onset of depression (BEDEPON) and substance abuse (BESUBON). For participants who did not experience an event, survival times were censored at their age at interview (PTAGE). In cases where substance abuse was diagnosed but the age of onset was missing (BESUBON = -1), the onset age was imputed as the interview age minus 0.5 years. A single missing value for socioeconomic class (SESCLASS) was imputed to the modal category (Class 3).

Covariates included parental depression (PARDEP), offspring sex (PTSEX), parental

marital status (MSPARENT), and socioeconomic class (SESCCLASS). Additionally, ‘prepubertyonset’ was defined as the interaction between parental depression and age < 13, and ‘priordep’ indicated whether offspring depression occurred prior to substance abuse onset.

Among the 220 offspring, 69 (31.4%) developed depression and 28 (12.7%) developed substance abuse. Descriptive statistics (Table 1) indicate that offspring of depressed parents had a higher prevalence of depression (37.6% vs. 23.2%) and substance abuse (16.8% vs. 7.4%) compared to controls. They were also more likely to be female (49.6%) and come from households with separated or divorced parents (28.0% vs. 11.6%).

**Hypothesis 1: Depression Onset.** The stratified Cox model provided a good fit to the data (Global Null Hypothesis LR  $\chi^2 = 19.42$ ,  $p=0.0007$ ). The interaction term between parental depression and the pre-puberty indicator was statistically significant ( $p=0.030$ ), providing evidence that the effect of parental depression differs by age. Specifically, the estimated hazard ratio for parental depression was 3.16 (95% CI: 0.56–17.85) for onset before age 13, compared to 0.69 (95% CI: 0.37–1.28) for onset at age 13 or later (Table 2). Although the individual hazard ratios within these time windows did not reach statistical significance due to the small number of early events (only 24 events occurred before age 13), the significant interaction confirms a differential impact. Additionally, female sex was a strong independent predictor, increasing the hazard of depression by 1.78 times (95% CI: 1.05–3.01,  $p=0.028$ ).

**Hypothesis 2: Substance Abuse Onset.** The multivariable Cox model for substance abuse was statistically significant (Global Null Hypothesis LR  $\chi^2 = 16.19$ ,  $p=0.040$ ). After adjusting for covariates, parental depression remained a marginally significant predictor, associated with a 2.33-fold increase in the hazard of substance abuse (95% CI: 0.88–6.17,  $p=0.088$ ) (Table 3). Interestingly, prior depression in the offspring was not significantly associated with substance abuse risk (HR=0.81,  $p=0.622$ ), suggesting that the risk conferred by parental depression operates independently of the child’s depression history. Parental marital status showed a trend where separation/divorce was associated with a lower hazard (HR=0.36,  $p=0.078$ ), though this result should be interpreted with caution given the wide confidence intervals. Socioeconomic class and sex were not significant predictors in this model.

## Conclusion

Our analysis supports the hypothesis that parental depression adversely affects offspring mental health, with effects varying by developmental stage and outcome. For depression, we found a significant interaction ( $p=0.030$ ) indicating that the impact of parental depression is qualitatively different in childhood versus adolescence, with a trend toward a strong pathogenic effect before puberty (HR=3.16). Female offspring are also at significantly higher risk regardless of parental status. For substance abuse, parental depression confers a more than two-fold increase in hazard (HR=2.33), a risk that appears independent of the child’s own history of depression. These findings highlight the importance of early intervention. Clinicians treating depressed adults should be aware of the elevated risks for their children, particularly for early-onset depression and subsequent substance use. Future studies with larger samples are needed to precisely estimate the magnitude of pre-pubertal risks.

**Table 1:** Demographic and Clinical Characteristics by Parental Depression Status

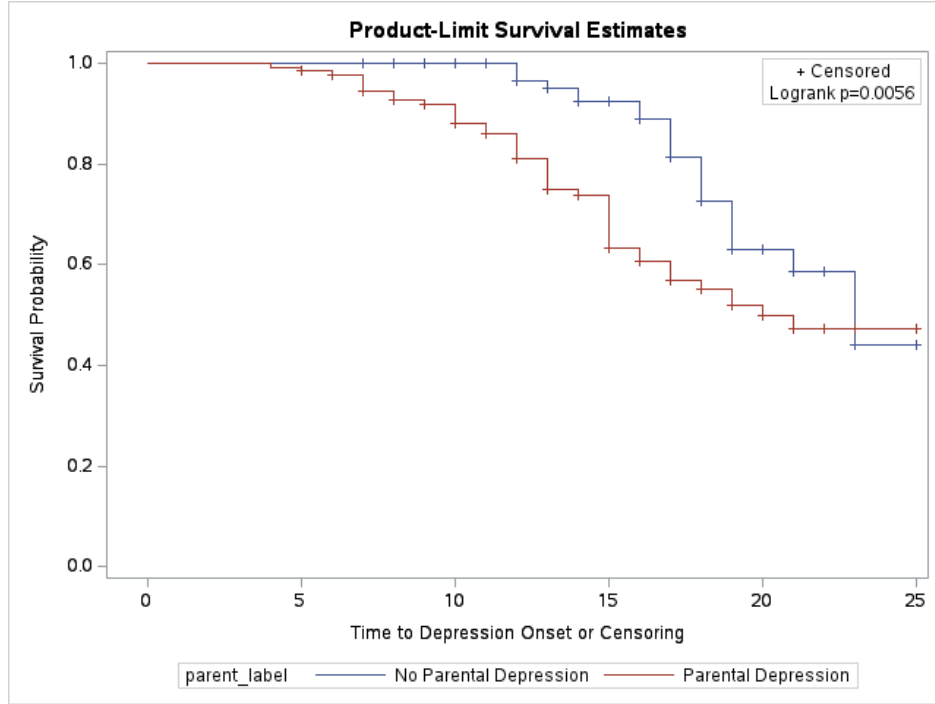
Variable	Depressed Parent (N=125)	Normal Parent (N=95)
<b>Child Sex</b>		
Male	63 (50.4%)	42 (44.2%)
Female	62 (49.6%)	53 (55.8%)
<b>Child Depression</b>		
Ever depressed	47 (37.6%)	22 (23.2%)
Never depressed	78 (62.4%)	73 (76.8%)
<b>Child Substance Abuse</b>		
Yes	21 (16.8%)	7 (7.4%)
No	104 (83.2%)	88 (92.6%)
<b>Parental Marital Status</b>		
Married	90 (72.0%)	84 (88.4%)
Separated/Divorced	35 (28.0%)	11 (11.6%)

**Table 2:** Cox Regression Results for Depression Onset: Interaction Model Testing Age-Specific Effects of Parental Depression

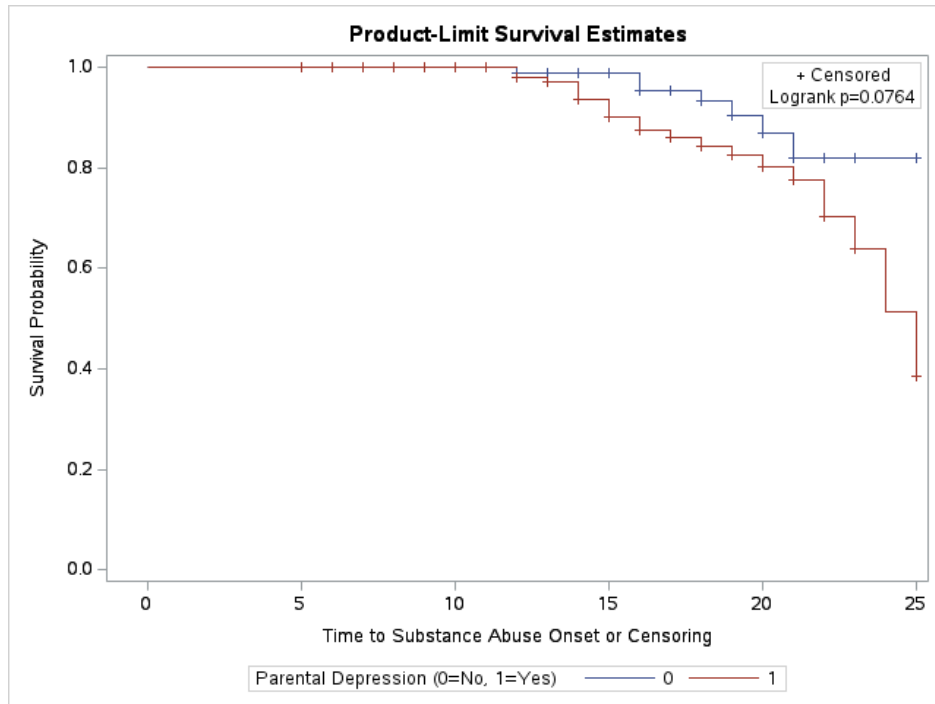
Parameter	HR	95% CI	P-value
<b>Parental Depression Effects:</b>			
Before age 13	3.16	(0.56, 17.85)	0.194
After age 13	0.69	(0.37, 1.28)	0.243
<b>Other Predictors:</b>			
Female sex	1.78	(1.05, 3.01)	0.032*
Marital status (Sep/Div)	0.96	(0.53, 1.75)	0.894
Model stratified by social class; * p < 0.05			
N=220, events=69			

**Table 3:** Cox Regression Results for Substance Abuse Onset: Joint Model of Parental and Prior Depression Effects

Variable	HR	95% CI	P-value
Parental Depression	2.33	(0.88, 6.17)	0.088
Prior Depression in Offspring	0.81	(0.34, 1.90)	0.622
Female Sex	0.77	(0.35, 1.72)	0.526
Marital Status (Sep/Div)	0.36	(0.12, 1.12)	0.078
Social class included in model (not shown)			
N=220, events=28			



**Figure 1:** Kaplan-Meier survival curves for depression onset by parental depression status (Log-rank  $\chi^2=7.69$ ,  $p=0.006$ )



**Figure 2:** Kaplan-Meier survival curves for substance abuse onset by parental depression status (Log-rank  $\chi^2=3.14$ ,  $p=0.076$ )