P8110 Applied Regression II

Midterm Project Report 1

Survival Analysis of Depression Onset and Substance Abuse in Offspring of Depressed Parents

Group 6

Leah Li (YL5828)
Xuange Liang (XL3493)
Zexuan Yan (ZY2654)
Yiwen Zhang (YZ4994)

November 4, 2025

Contents

1 Data Preparation

1.1 Survival Outcome Definitions

Depression Onset (Pre-pubertal, age ¡ 13):

- Start time: Birth (age = 0); End time: Age of depression onset (`BEDEPON`) if ¡ 13, otherwise censored at `PTAGE`

Depression Onset (Adolescent/early adulthood, age $\geq$ 13):

- Start time: Birth (age = 0); End time: Age of depression onset (`BEDEPON`) if $\geq$ 13, otherwise censored at `PTAGE`

Substance Abuse Onset:

- Start time: Birth (age = 0); End time: Age of substance abuse onset (`BESUBON`), censored at `PTAGE` if no event

1.2 Data Cleaning

Logical consistency checks were performed on all variables:

- Records with status=1 (diagnosed) but missing onset age (-1) were removed as incomplete
- Records with onset age present but status=0 were corrected to status=1 for consistency
- This procedure was applied to both depression (`DSMDEPHR`, `BEDEPON`) and substance abuse variables (`DSMSUBHR`, `BESUBON`)

1.3 Descriptive Statistic Table

Summary Statistics:

- Age at interview: Mean = 16.72 years (SD = 4.73, Range: 5–25)
- Age of depression onset (n=69): Mean = 14.00 years (SD = 4.16, Range: 4–23)
- Age of substance abuse onset (n=17): Mean = 15.12 years (SD = 2.32, Range: 12–21)

2 Descriptive Statistics Analyses

2.1 Sample Characteristics

The study included 220 children (125 with depressed parents, 95 with non-depressed parents) aged 5–25 years (Mean = 16.72, SD = 4.73). Depression onset occurred in 69 children (31.4%, mean age = 14.00 years, SD = 4.16), and substance abuse in 28 children (12.7%, mean age = 15.12 years, SD = 2.32).

2.2 Comparison by Parental Depression Status

Table 1 shows demographic and clinical characteristics stratified by parental depression status. Children of depressed parents had higher rates of depression (37.6% vs 23.2%) and substance abuse (16.8% vs 7.4%). Separation/divorce was more common among depressed parents (28.0% vs 11.6%).

Overall Kaplan-Meier Analysis: Children of depressed parents demonstrated earlier depression onset (median: 20 vs 23 years). Log-rank test: $\chi^2 = 7.69$, p = 0.0056 (statistically significant).

Table 1: Demographic and Clinical Characteristics of the Offspring Stratified by Parental Depression Status

Variable	Depressed Index Parent Group (N=125) n (%)	Normal Index Parent Group (N=95) n (%)
Sex of Child		
Male	63 (50.4%)	42 (44.2%)
Female	62 (49.6%)	53 (55.8%)
Child depression status		
Ever depressed	47 (37.6%)	22 (23.2%)
Never depressed	78 (62.4%)	73 (76.8%)
Child substance abuse status		
Yes	21 (16.8%)	7 (7.4%)
No	104 (83.2%)	88 (92.6%)
Social class of index parent		
1 (Highest)	16 (12.8%)	1 (1.1%)
2	19 (15.2%)	19 (20.0%)
3	31 (24.8%)	16 (16.8%)
4	50 (40.0%)	47 (49.5%)
5 (Lowest)	8 (6.4%)	12 (12.6%)
Marital status of index parent		
Married with spouse	90 (72.0%)	84 (88.4%)
Separated/Divorced	35 (28.0%)	11 (11.6%)

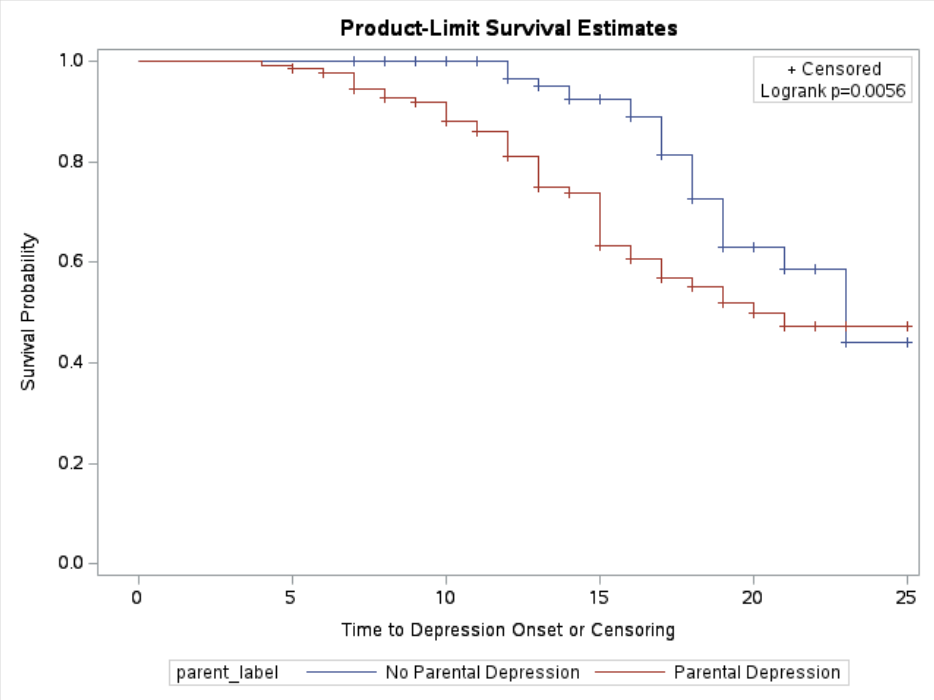

Figure 1: Overall Kaplan–Meier survival curves of depression onset by parental depression status

3 Analyses to Address Hypothesis 1

Hypothesis 1: Offspring of a depressed parent are more likely to have pre-pubertal onset (¡13 years) depression than offspring of a non-depressed parent, but equally likely to have an onset of adolescent/early adulthood depression, given demographic and social characteristics.

3.1 Statistical Modeling Approach

Cox proportional hazards regression models were fit separately for: (a) pre-pubertal onset (age ¡ 13), and (b) adolescent/early adulthood onset (age $\geq$ 13).

Cox Model:

$$h(t|X) = h_0(t) \times \exp(\beta_1 \times \text{PARDEP} + \beta_2 \times \text{SEX} + \beta_3 \times \text{AGE} + \beta_4 \times \text{SESCLASS} + \beta_5 \times \text{MSPARENT}) \quad (1)$$

Where: $h(t|X)$ = hazard at time t; $h_0(t)$ = baseline hazard; PARDEP = parental depression (1=depressed, 0=normal); SEX = child's sex (1=male, 2=female); AGE = age at interview; SESCLASS = social class (1–5); MSPARENT = marital status (1=married, 2=separated/divorced).

Proportional hazards (PH) assumptions were tested using Schoenfeld residuals and supremum tests. Violations were addressed using stratified Cox models.

3.2 Results

3.2.1 Pre-pubertal Onset (Age ¡ 13 years)

Kaplan-Meier: Log-rank $\chi^2 = 10.56$, p = 0.0012. Children of depressed parents had significantly earlier pre-pubertal depression onset.

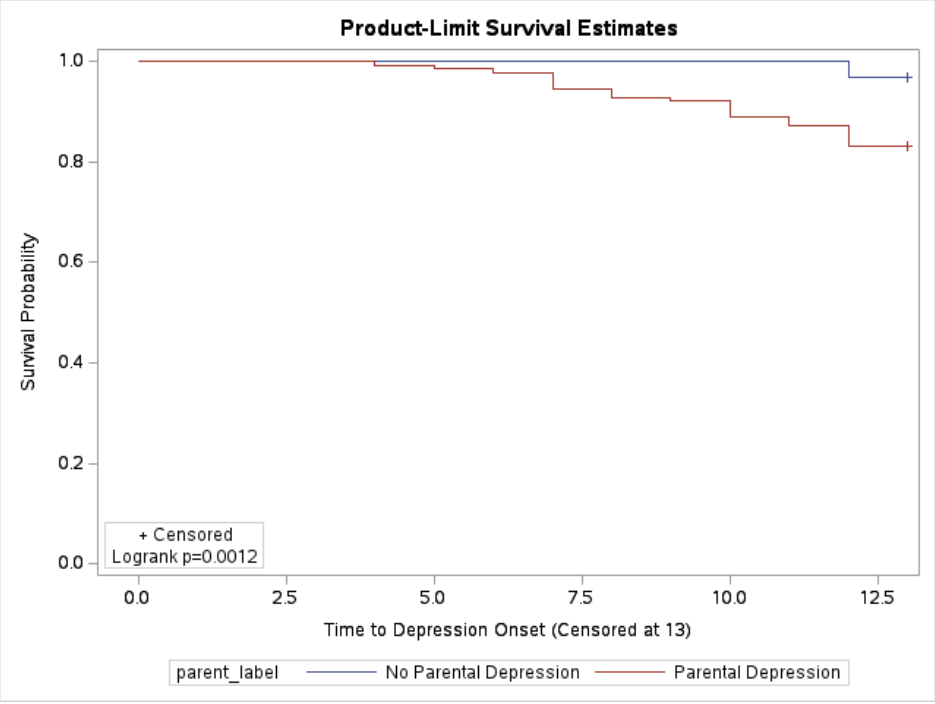

Figure 2: H1a: Kaplan–Meier curves for pre-pubertal depression onset (age < 13) by parental depression

Cox Regression (Table 2a):

Key Findings: Children of depressed parents have 5.86 times the hazard of pre-pubertal depression (p = 0.0048). No other covariates were significant. PH assumptions satisfied (all p ¿

Table 2: Cox Regression Results for Pre-pubertal Depression Onset (Age < 13 years)

Variable	Hazard Ratio	95% CI	P-value
Parental Depression	5.863	(1.717, 20.019)	0.0048**
Sex (Female vs Male)	1.171	(0.522, 2.626)	0.7024
Age at interview	0.988	(0.913, 1.069)	0.7608
Social class	0.822	(0.591, 1.144)	0.2460
Marital status (Sep/Div)	0.585	(0.198, 1.731)	0.3328

** Statistically significant at $\alpha = 0.01$ level

0.05).

3.2.2 Adolescent/Early Adulthood Onset (Age ≥ 13 years)

Kaplan-Meier: Log-rank $\chi^2 = 0.83$, p = 0.3632 (not significant). No difference by parental status, consistent with hypothesis.

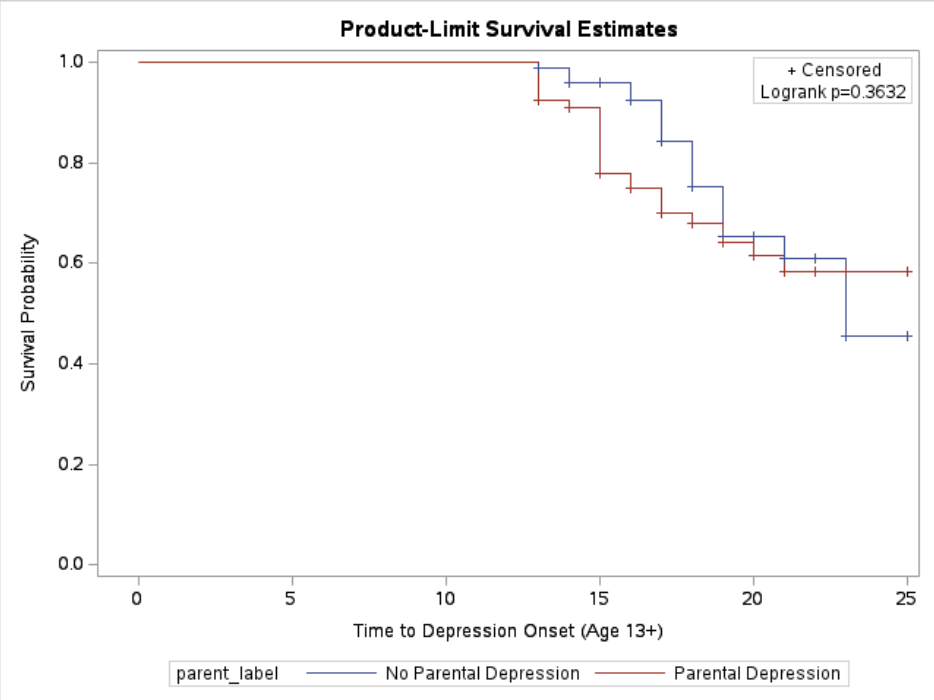

Figure 3: H1b: Kaplan–Meier curves for adolescent/early adulthood depression onset (age ≥ 13) by parental depression

Cox Regression (Table 2b):

Key Findings: Parental depression not significant (HR=1.36, p=0.325). Female sex strongly predicts adolescent depression (HR=2.48, p=0.0055). PH assumption violated for parental depression (p=0.014); stratified model confirmed results.

3.3 Conclusion for Hypothesis 1

Hypothesis 1 is FULLY SUPPORTED. Parental depression strongly increases pre-pubertal depression risk (HR=5.86, p=0.0048) but not adolescent-onset risk (HR=1.36, p=0.325). The findings demonstrate a clear age-specific pattern of familial transmission, with parental depression

Table 3: Cox Regression Results for Adolescent/Early Adulthood Depression Onset (Age $\geq$ 13 years)

Variable	**Hazard Ratio**	**95% CI**	**P-value**
Parental Depression	1.362	(0.736, 2.521)	0.3246
Sex (Female vs Male)	2.481	(1.307, 4.710)	0.0055**
Age at interview	0.984	(0.871, 1.113)	0.8018
Social class	1.140	(0.855, 1.520)	0.3722
Marital status (Sep/Div)	1.622	(0.825, 3.188)	0.1605

** Statistically significant at $\alpha = 0.01$ level

being a significant risk factor only for early-onset depression.

4 Analyses to Address Hypothesis 2

Hypothesis 2: Is there evidence for the effect of (1) prior depression in offspring, as well as (2) the effect of parent's depression status on the age of onset of substance abuse/dependence in offspring, given demographic and social characteristics?

4.1 Statistical Modeling Approach

Three Cox regression models tested: (2a) prior depression effect, (2b) parental depression effect, (2c) joint model. Prior depression defined as depression onset before substance abuse onset.

Model Specifications:

$$\text{Model 2a: } h(t|X) = h_0(t) \times \exp(\beta_1 \times \text{PRIORDEP} + \text{covariates}) \quad (2)$$

$$\text{Model 2b: } h(t|X) = h_0(t) \times \exp(\beta_1 \times \text{PARDEP} + \text{covariates}) \quad (3)$$

$$\text{Model 2c: } h(t|X) = h_0(t) \times \exp(\beta_1 \times \text{PRIORDEP} + \beta_2 \times \text{PARDEP} + \text{covariates}) \quad (4)$$

Where PRIORDEP = prior depression (1=yes, 0=no); PARDEP = parental depression; covariates = sex, age, social class, marital status.

4.2 Results

4.2.1 Model 2a: Effect of Prior Depression

Kaplan-Meier: Log-rank $\chi^2 = 0.07$, p = 0.7988 (not significant). No difference in substance abuse onset by prior depression status.

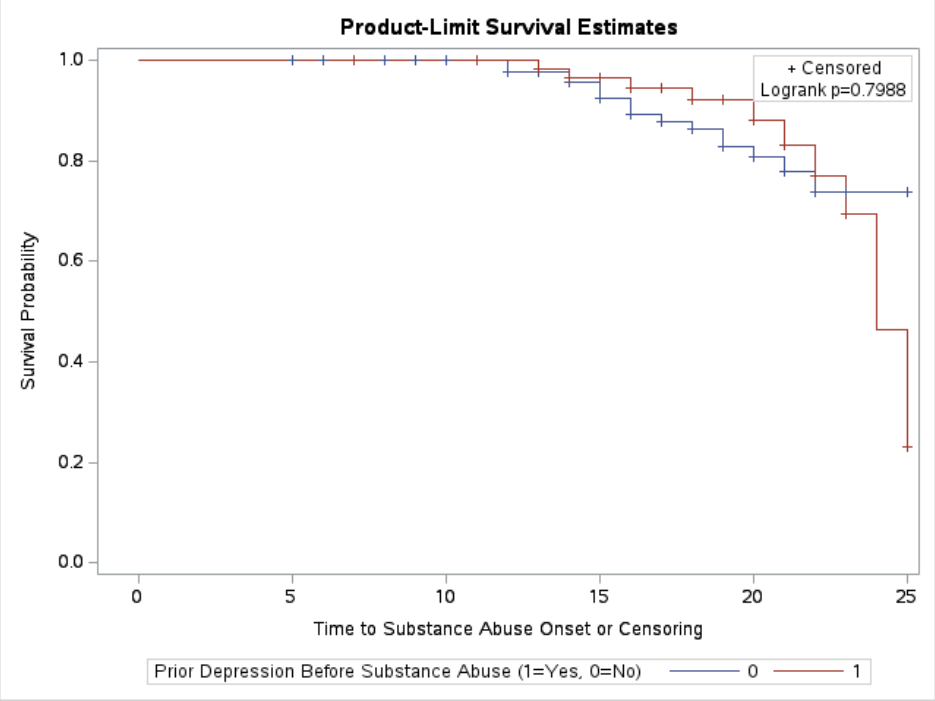

Figure 4: H2a: Kaplan–Meier curves for substance abuse onset by prior depression in offspring

Cox Regression (Table 3a):

Key Findings: Prior depression shows no effect (HR=0.999, p=0.998). No covariates significant.

Table 4: Cox Regression Results for Substance Abuse Onset: Effect of Prior Depression

Variable	Hazard Ratio	95% CI	P-value
Prior Depression	0.999	(0.454, 2.198)	0.9981
Sex (Female vs Male)	0.538	(0.247, 1.170)	0.1155
Age at interview	1.045	(0.889, 1.228)	0.5981
Social class	1.000	(0.692, 1.445)	0.9989
Marital status (Sep/Div)	0.441	(0.149, 1.303)	0.1388

4.2.2 Model 2b: Effect of Parental Depression

Kaplan-Meier: Log-rank $\chi^2 = 3.14$, p = 0.0764 (marginally significant). Trend toward earlier substance abuse in children of depressed parents.

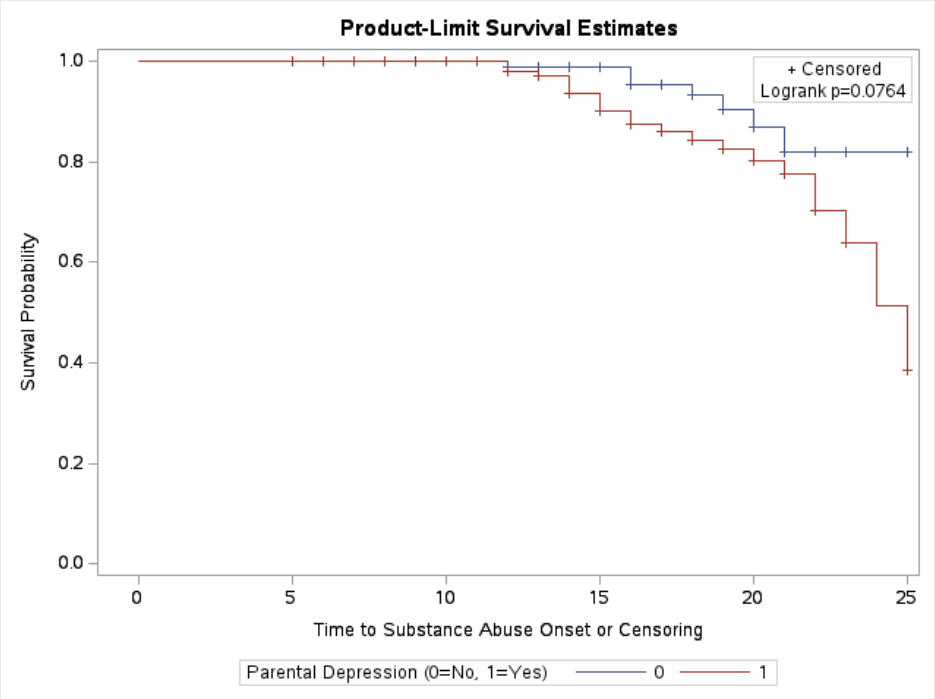

Figure 5: H2b: Kaplan–Meier curves for substance abuse onset by parental depression

Cox Regression (Table 3b):

Table 5: Cox Regression Results for Substance Abuse Onset: Effect of Parental Depression

Variable	Hazard Ratio	95% CI	P-value
Parental Depression	2.466	(1.001, 6.076)	0.0498*
Sex (Female vs Male)	0.592	(0.271, 1.293)	0.1794
Age at interview	1.021	(0.869, 1.200)	0.8018
Social class	1.107	(0.771, 1.591)	0.5728
Marital status (Sep/Div)	0.378	(0.128, 1.117)	0.0801

* Statistically significant at $\alpha = 0.05$ level

Key Findings: Parental depression marginally significant (HR=2.47, p=0.0498). Children of depressed parents have 2.5 times the hazard of substance abuse.

4.2.3 Model 2c: Joint Model

Cox Regression (Table 3c):

Table 6: Cox Regression Results for Substance Abuse Onset: Joint Model (Prior and Parental Depression)

Variable	Hazard Ratio	95% CI	P-value
Prior Depression	0.819	(0.372, 1.802)	0.6273
Parental Depression	2.586	(1.023, 6.539)	0.0430*
Sex (Female vs Male)	0.605	(0.277, 1.322)	0.2008
Age at interview	1.024	(0.871, 1.203)	0.7809
Social class	1.112	(0.770, 1.606)	0.5529
Marital status (Sep/Div)	0.366	(0.121, 1.106)	0.0740

* Statistically significant at $\alpha = 0.05$ level

Key Findings: Parental depression remains significant (HR=2.59, p=0.043) in joint model. Prior depression not significant (HR=0.82, p=0.627). Parental depression is the primary driver of substance abuse risk.

4.3 Conclusion for Hypothesis 2

Hypothesis 2 is PARTIALLY SUPPORTED. Prior depression shows no effect on substance abuse (HR=0.999, p=0.998). Parental depression has a marginally significant effect (HR=2.47-2.59, p=0.043-0.050). In the joint model, parental depression remains significant while prior depression does not, indicating parental depression directly affects substance abuse risk independent of offspring's own depression.

Clinical Implications: Findings suggest familial transmission of substance abuse risk operates through pathways beyond offspring psychopathology. Prevention efforts should target children of depressed parents regardless of child's depression status.

5 Summary

5.1 Main Findings

- **Hypothesis 1 (FULLY SUPPORTED):** Parental depression strongly increases pre-pubertal depression risk (HR=5.86, p¡0.01) but not adolescent-onset risk (HR=1.36, p=0.32). Age-specific familial transmission pattern identified.
- **Hypothesis 2 (PARTIALLY SUPPORTED):** Prior depression shows no effect on substance abuse (p=0.998). Parental depression marginally significant (HR=2.47-2.59, p=0.04-0.05), operating independently of offspring depression.
- **Additional Finding:** Female sex strongly predicts adolescent-onset depression (HR=2.48, p¡0.01).

5.2 Clinical Implications and Limitations

Early intervention critical for children of depressed parents, especially before age 13. Substance abuse prevention should target all children of depressed parents regardless of child's depression status.

Limitations: Retrospective assessment may introduce recall bias; small event counts (n=24 pre-pubertal depression, n=28 substance abuse) limit precision and power; PH assumption violated for parental depression in adolescent analysis (addressed via stratified model).