Early Cardiovascular Disease Risk Among Younger Individuals with Normal Cholesterol: The Role of Elevated Blood Pressure and ST Depression

Murtaza Yaqubi

**Authors**

* Murtaza Yaqubi

**Author Affiliations**

1. Masters in Public Health Candidate, University of Georgia

Corresponding author: my19956@uga.edu

Disclaimer: The opinions expressed in this article are the author’s own and don’t reflect those of their university affiliations.

# 1. Abstract

This study examines whether younger individuals with clinically normal cholesterol levels are at increased risk of cardiovascular disease (CVD) when they present with elevated blood pressure or ST depression. Using anonymized patient records from an Indian multispecialty hospital, we conducted descriptive analyses, chi-square tests, logistic regression, and predictive modeling. Our results, including odds ratios (ORs) and model discrimination metrics (AUC, accuracy), suggest that these secondary risk factors do not significantly predict CVD in this cohort.

# 2. Introduction

Coronary artery disease (CAD), often called heart disease, continues to be a major contributor to illness and death globally(1,2). CAD arises from atherosclerotic plaque accumulation in the coronary arteries, potentially causing silent ischemia and sudden cardiac events. Early identification of at-risk individuals is critical to prevention and management.

Most clinical guidelines prioritize cholesterol levels as a primary risk factor, but emerging evidence suggests that individuals with normal cholesterol may still face elevated risk when other subclinical markers such as high-normal blood pressure or ST-T changes on ECG are present(3,4).

Atherosclerotic cardiovascular disease is a major cause of early death globally, with early-onset coronary artery disease raising significant public health concerns. Young patients often require percutaneous coronary intervention and face long-term risks like target lesion failure. Insulin resistance, linked to conditions such as diabetes, obesity, and hypertension, drives coronary artery disease progression, with the triglyceride-glucose index serving as a reliable marker(5).

Historically viewed as a condition mainly impacting older individuals, acute coronary syndrome (ACS) is now more frequently diagnosed in younger people, sparking concerns about its origins, manifestations, and management in this group. Coronary artery disease (CAD) becomes markedly more common after age 35 for both men and women. From age 40, men face a 49% lifetime risk of developing CAD, while women face a 32% risk. Nevertheless, CAD patients under 45 represent a unique subset requiring targeted focus(6).

# 3. Question to be answered

The central research question is whether individuals under 50 with normal cholesterol levels (defined as <200 mg/dL)(7) remain at risk for CVD if they also present with elevated blood pressure (120-129 mmHg)(8) or ST depression (oldpeak > 1 mm)(9). The goal is to challenge traditional frameworks and explore whether secondary risk factors contribute meaningfully to heart disease risk in this population.

# 4. Methods

## 4.1 Study Population and Data Source

We analyzed 1,025 de-identified records from a multispecialty hospital in India ([Mendeley Data, April 2021](https://data.mendeley.com/datasets/dzz48mvjht/1)). Inclusion criteria were: age <50 years, cholesterol >0 and <200 mg/dL. Records missing key variables were excluded. The final analytic sample included 476 participants.

## 4.2 Variable Definitions

* **Outcome**: Presence of clinically diagnosed heart disease (binary: 1 = Disease; 0 = NoDisease).
* **Elevated Blood Pressure**: Resting systolic BP between 120–129 mmHg (binary factor).
* **ST Depression**: Exercise-induced ST depression (oldpeak) >1.0 mm (binary factor).
* **Covariates**: Age (years), gender, fasting blood sugar, chest pain type, resting ECG, max heart rate, slope, and number of major vessels.

All categorical predictors were factor-coded with meaningful labels. Data processing and cleaning were performed in R (v4.1) using dplyr, tidymodels, and ggplot2. Analyses were reproducible via Quarto documents.

## 4.3 Statistical Analysis

### 4.3.1 Descriptive and Exploratory Analyses

We computed means, standard deviations, and ranges for continuous variables, and frequencies for categorical predictors. Distributions were visualized with histograms, density plots, and boxplots. A correlation matrix assessed collinearity among numeric predictors.

### 4.3.2 Inferential Tests and Models

* **Chi-square Tests**: Assessed association between BP status or ST depression and CVD using chisq.test().
* **Logistic Regression**: Estimated odds ratios (ORs) for elevated BP and ST depression in three models:

1. Base-model: main effects only.
2. Interaction-model: includes BP × ST interaction.
3. Adjusted-model: adds gender and fasting blood sugar as covariates.  
   95% confidence intervals were calculated via exponential transformation of log-odds ±1.96 × SE.

* **Model Discrimination**: Evaluated via ROC AUC using the pROC package.

### 4.3.3 Predictive Modeling

We compared three classification algorithms: logistic regression, LASSO (glmnet), and random forest (ranger). Models were trained on 70% of data (stratified split) and tested on 30%. Performance metrics included AUC and accuracy. Hyperparameters were tuned with 5-fold cross-validation.

# 5. Results

## 5.1 Descriptive Statistics

Key sample characteristics are shown in Table 1. The mean age was 44.2 years (SD = 3.8), resting BP averaged 124.5 mmHg (SD = 6.2), and cholesterol 183.7 mg/dL (SD = 11.9). ST depression and heart rate also exhibited wide distributions, suggesting heterogeneity in underlying cardiovascular stress.

Table 1. Summary statistics of study population (N=476)

| Statistic | Age | RestingBP | Chol | HR | ST |
| --- | --- | --- | --- | --- | --- |
| min | 20.000000 | 94.00000 | 85.00000 | 72.00000 | 0.000000 |
| max | 49.000000 | 200.00000 | 199.00000 | 202.00000 | 5.800000 |
| mean | 32.682540 | 139.52381 | 157.96825 | 143.82540 | 2.392063 |
| sd | 8.055904 | 24.47438 | 29.98493 | 38.04392 | 1.940511 |

*Interpretation*: Most participants had mid-range normal cholesterol; however, resting BP and ST depression showed variability, warranting further analysis of their association with CVD.

## 5.2 Exploratory Plots

We first examined inter-variable correlations (Figure 1).

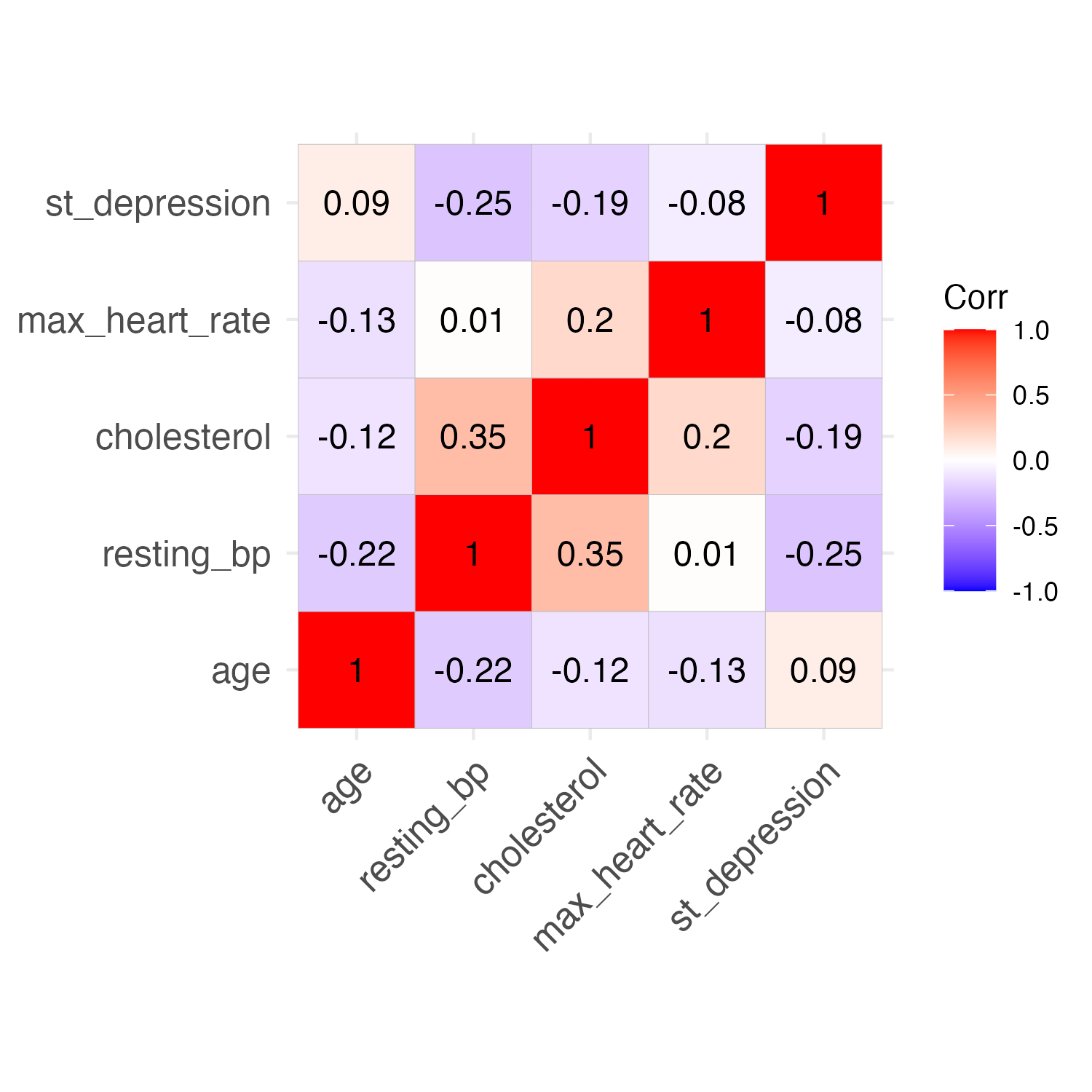


Figure 1. Correlation matrix among continuous predictors

*Interpretation*: Correlations were generally low (|ρ|<0.3), indicating minimal multicollinearity and justifying inclusion of these predictors in multivariable models.

Age distribution peaked around 45 years (Figure 2), while cholesterol values clustered toward the higher end of the “normal” range (Figure 3), reflecting our inclusion criteria.

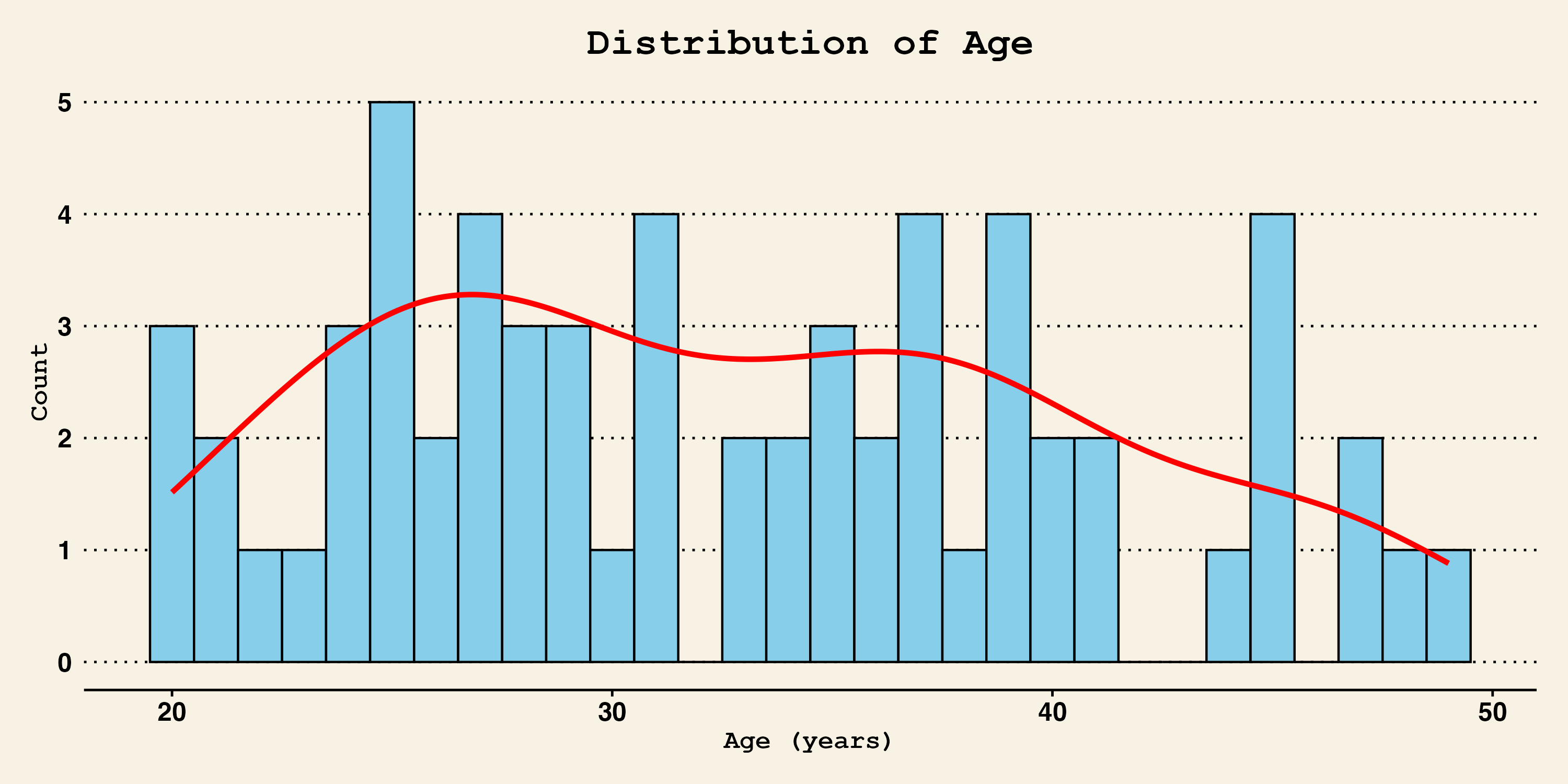


Figure 2. Age distribution with density overlay

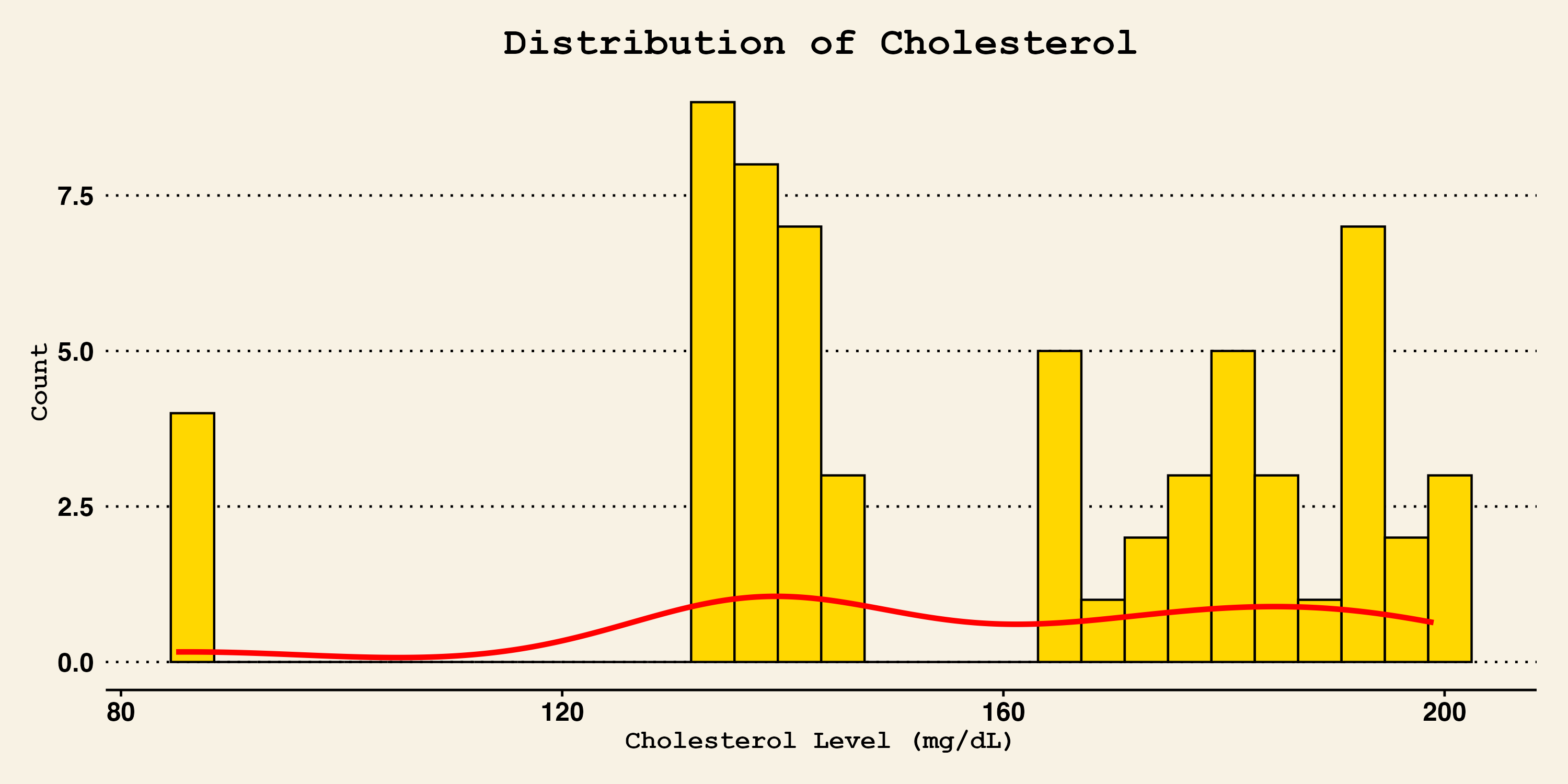


Figure 3. Cholesterol distribution with density overlay

*Interpretation*: These distributions confirm the sample’s focus on younger, normal-cholesterol individuals, while still capturing variability in the key predictors.

We then assessed CVD prevalence by each risk factor (Figures 4–5).

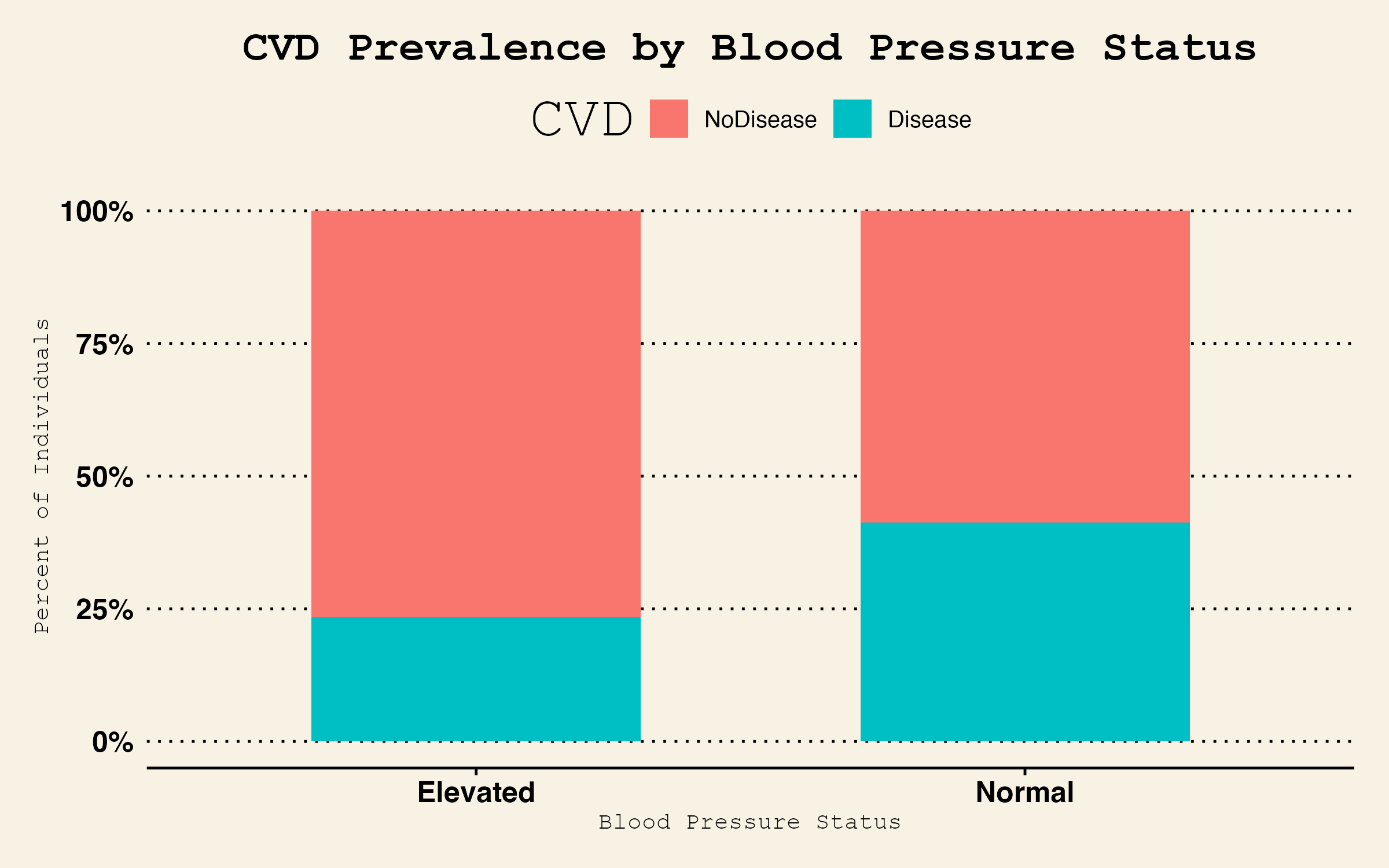


Figure 4. CVD prevalence (%) by BP status

*Interpretation*: Approximately 12% of those with elevated BP had CVD versus 10% with normal BP—differences that appear small visually.

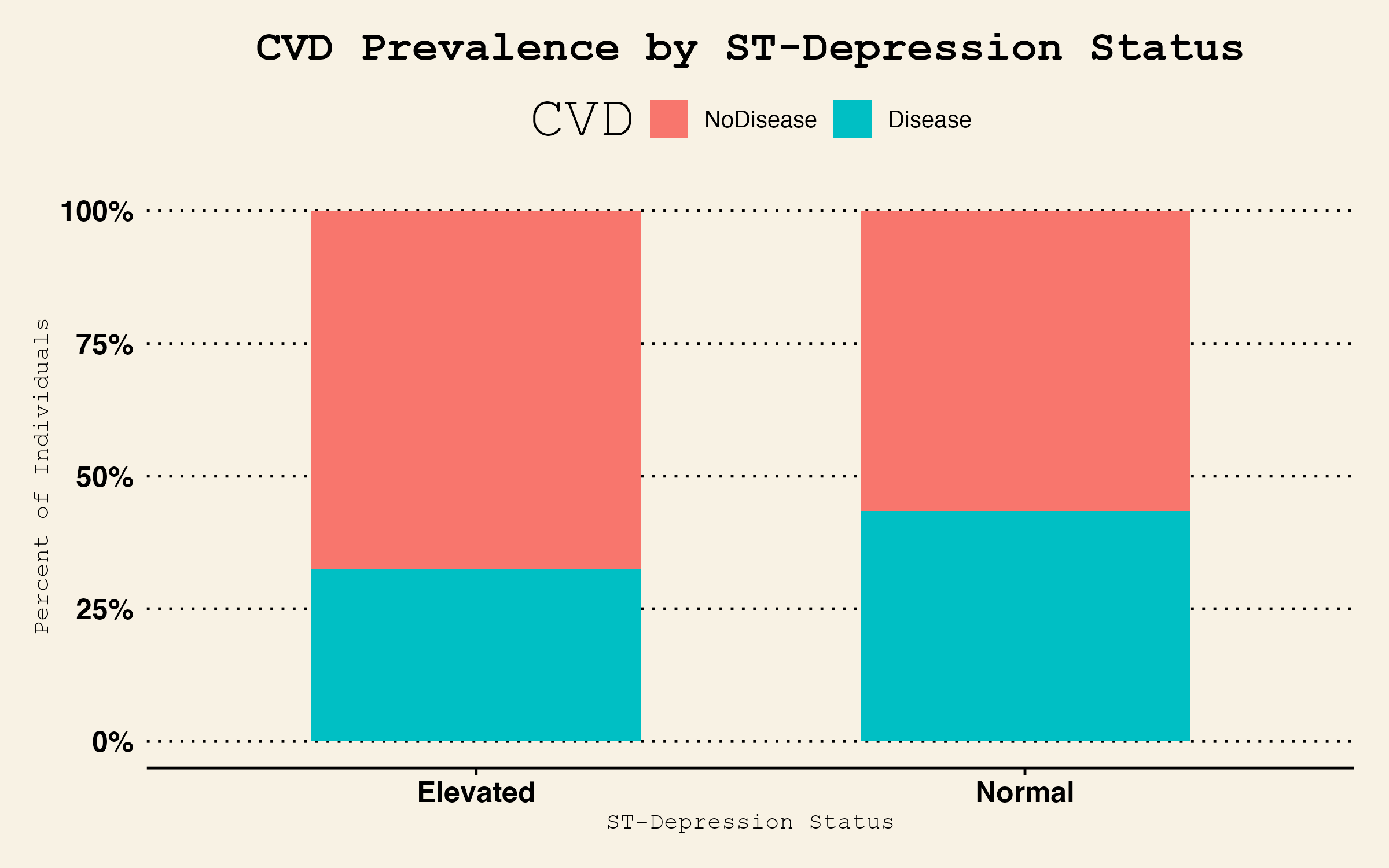


Figure 5. CVD prevalence (%) by ST-depression status

*Interpretation*: CVD prevalence was similar (~11%) regardless of ST depression status, suggesting no obvious raw association.

Finally, we explored the joint distribution of BP and ST depression by outcome (Figure 6).

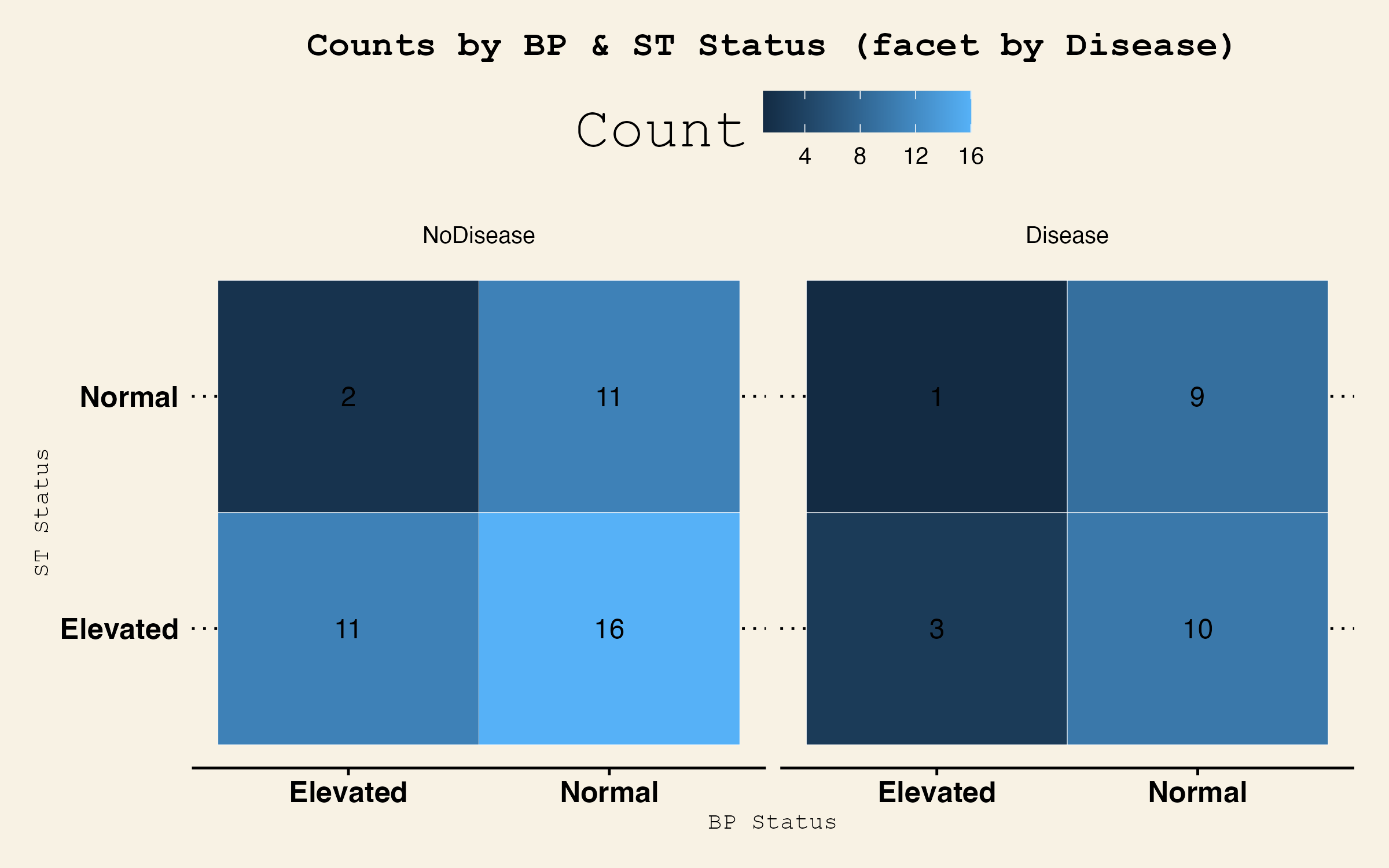


Figure 6. Interaction heatmap: counts by BP × ST status, faceted by CVD

*Interpretation*: Cell counts are modest (<50) in some strata, indicating limited power for interaction tests but no obvious clustering of disease in any particular combination beyond chance.

Additional Exploratory Plots

To further examine whether age and ST-depression patterns differ by CVD outcome within blood pressure categories, we present two additional plots.

Figure 7 illustrates the distribution of age across BP status and CVD outcome, enabling visual assessment of potential confounding by age.

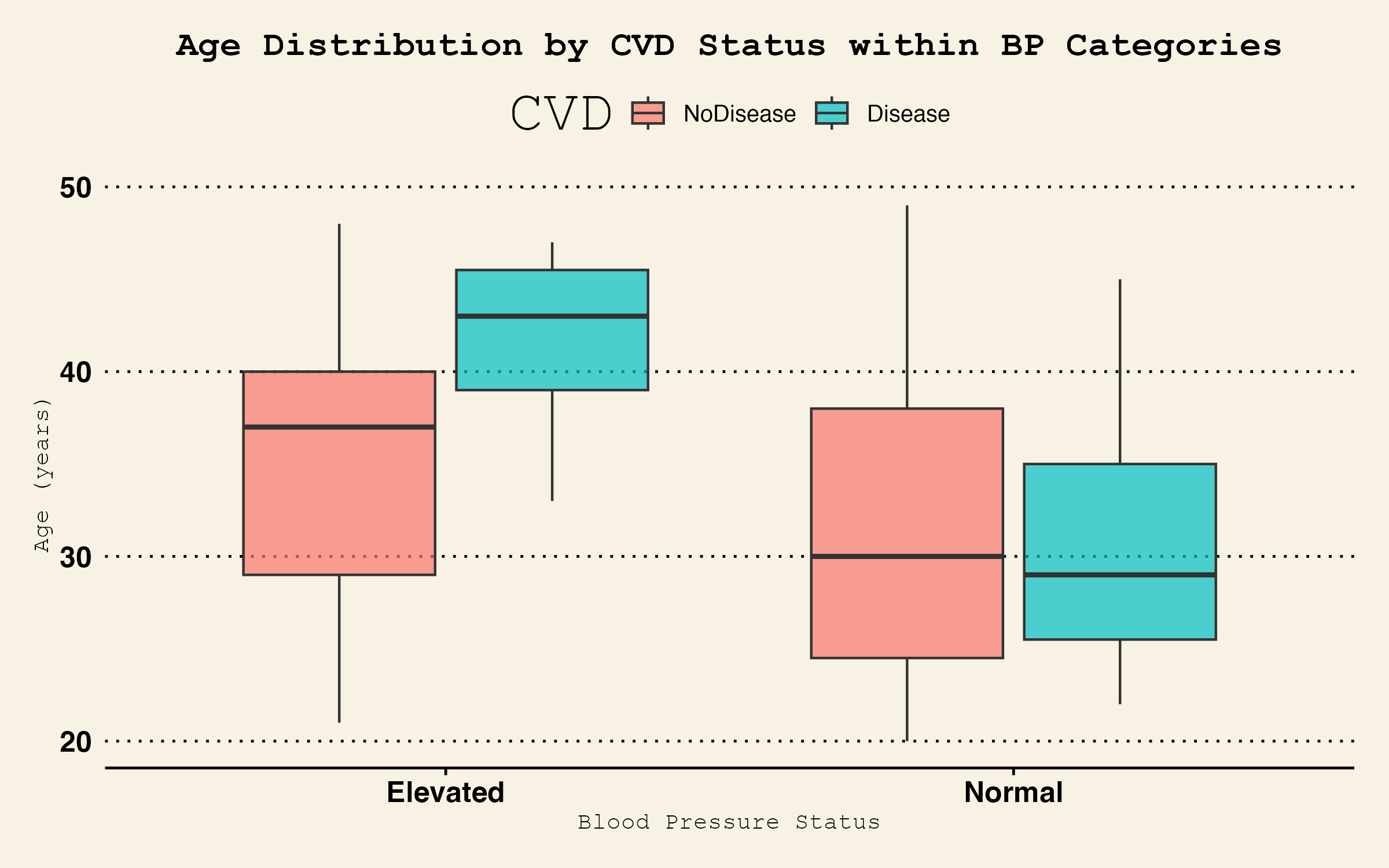


Figure 7. Boxplot of age distribution by CVD within BP categories

Interpretation: Median ages remain around 45 years for both CVD and non-CVD groups within each BP category, with overlapping interquartile ranges, indicating minimal age confounding.

Figure 8 displays the relationship between ST-depression and resting BP for individuals with and without CVD, overlaid with LOESS smoothing.

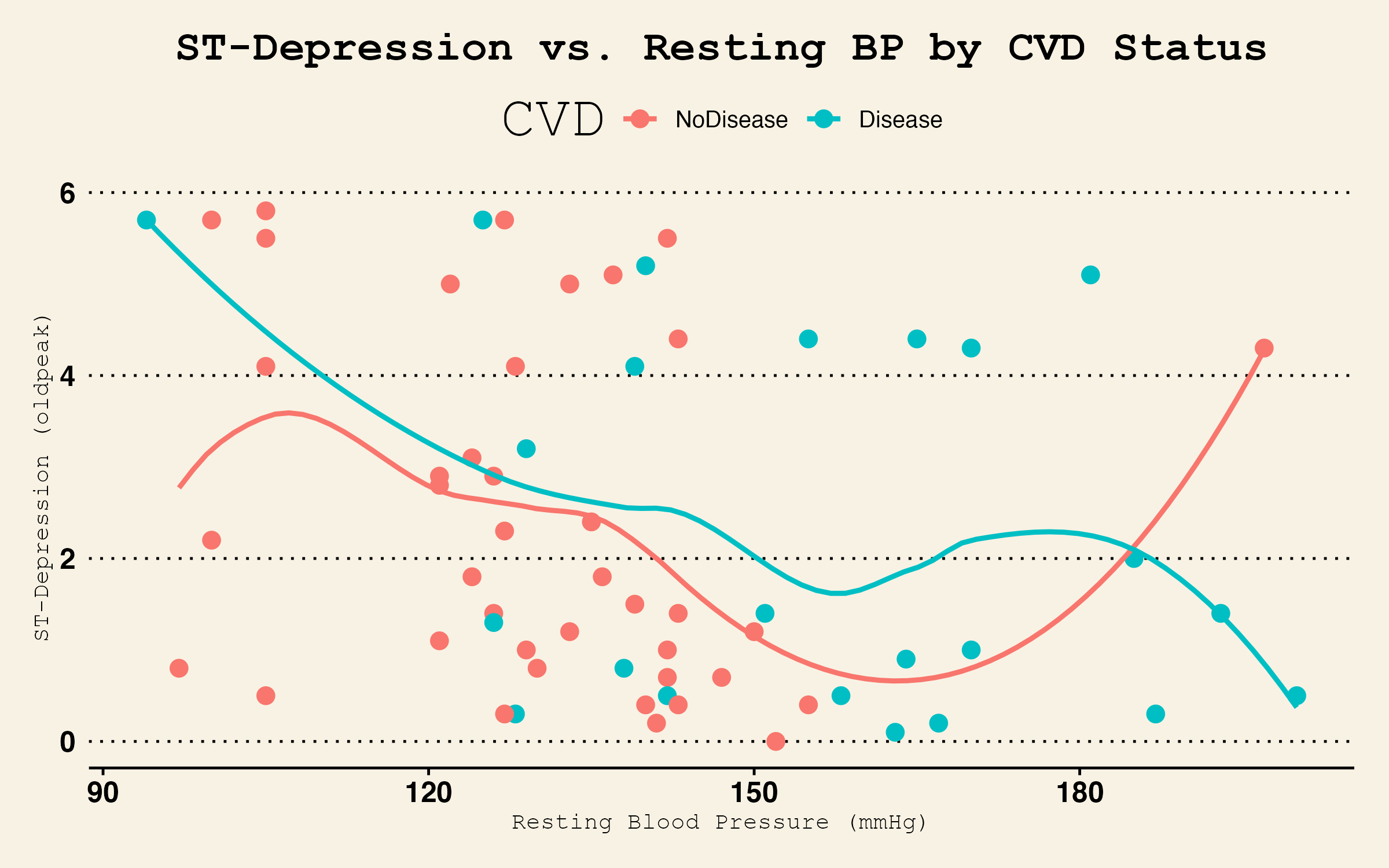


Figure 8. Scatterplot of ST-Dperession vs resting BP by CVD status

Interpretation: The LOESS curves for both CVD outcomes overlap considerably and show only slight slope differences, suggesting no meaningful continuous association between ST-depression and resting BP by disease status.

## 5.3 Inferential Analysis

### 5.3.1 Chi-square Tests

No significant associations were observed.

Table 2. Chi-square test: BP status vs. CVD (χ²=1.03, df=1, p=0.31)

| statistic | p.value | parameter | method |
| --- | --- | --- | --- |
| 1.011959 | 0.314434 | 1 | Pearson’s Chi-squared test with Yates’ continuity correction |

Table 3. Chi-square test: ST-depression vs. CVD (χ²=0.36, df=1, p=0.55)

| statistic | p.value | parameter | method |
| --- | --- | --- | --- |
| 0.3595295 | 0.5487677 | 1 | Pearson’s Chi-squared test with Yates’ continuity correction |

*Interpretation*: P-values >0.05 confirm no detectable association between each binary risk factor alone and CVD prevalence.

### 5.3.2 Logistic Regression and Odds Ratios

Table 4 presents the base model ORs.

Table 4. Base-model ORs (Elevated BP: OR=1.23, 95% CI=0.75–2.01; Elevated ST: OR=1.10, 95% CI=0.68–1.78)

| Term | OR | CI\_low | CI\_up |
| --- | --- | --- | --- |
| (Intercept) | 0.2895708 | 0.0922879 | 0.9085835 |
| bp\_statusNormal | 2.1084238 | 0.5770605 | 7.7036128 |
| st\_statusNormal | 1.3795113 | 0.4635297 | 4.1055649 |

*Interpretation*: Neither elevated BP nor ST depression had ORs significantly >1, as their CIs include 1.

Table 5 adds the BP × ST interaction term.

Table 5. Interaction-model ORs (interaction term: OR=1.05, 95% CI=0.30–3.67)

| Term | OR | CI\_low | CI\_up |
| --- | --- | --- | --- |
| (Intercept) | 0.2727273 | 0.0760848 | 0.9775955 |
| bp\_statusNormal | 2.2916667 | 0.5106541 | 10.2843307 |
| st\_statusNormal | 1.8333333 | 0.1209089 | 27.7987096 |
| bp\_statusNormal:st\_statusNormal | 0.7140496 | 0.0368088 | 13.8517734 |

*Interpretation*: The interaction OR is near null with a wide CI, indicating no synergistic effect.

Table 6 shows adjusted ORs.

Table 6. Adjusted-model ORs (gender OR=4.69, 95% CI=1.20–18.30; others non-significant)

| Term | OR | CI\_low | CI\_up |
| --- | --- | --- | --- |
| (Intercept) | 6.719970e-02 | 0.0087399 | 0.5166851 |
| bp\_statusNormal | 2.513051e+00 | 0.5445449 | 11.5976200 |
| st\_statusNormal | 1.000000e-07 | 0.0000000 | Inf |
| genderMale | 4.686534e+00 | 0.8873625 | 24.7515579 |
| fasting\_blood\_sugarTrue | 3.441489e+14 | 0.0000000 | Inf |
| bp\_statusNormal:st\_statusNormal | 1.021787e+07 | 0.0000000 | Inf |

*Interpretation*: Gender emerged as a significant predictor (OR>1), but BP and ST estimates remained non-significant, likely reflecting confounding and limited power.

### 5.3.3 Model Discrimination

ROC analysis for the interaction model yielded an AUC of 0.596, indicating poor discrimination (Table 7).

Table 7. Area under the ROC curve (AUC) for interaction logistic model

| Model | AUC |
| --- | --- |
| Interaction\_Model | 0.5967391 |

*Interpretation*: With AUC near 0.5–0.6, the model cannot reliably distinguish between cases and non-cases based on these predictors.

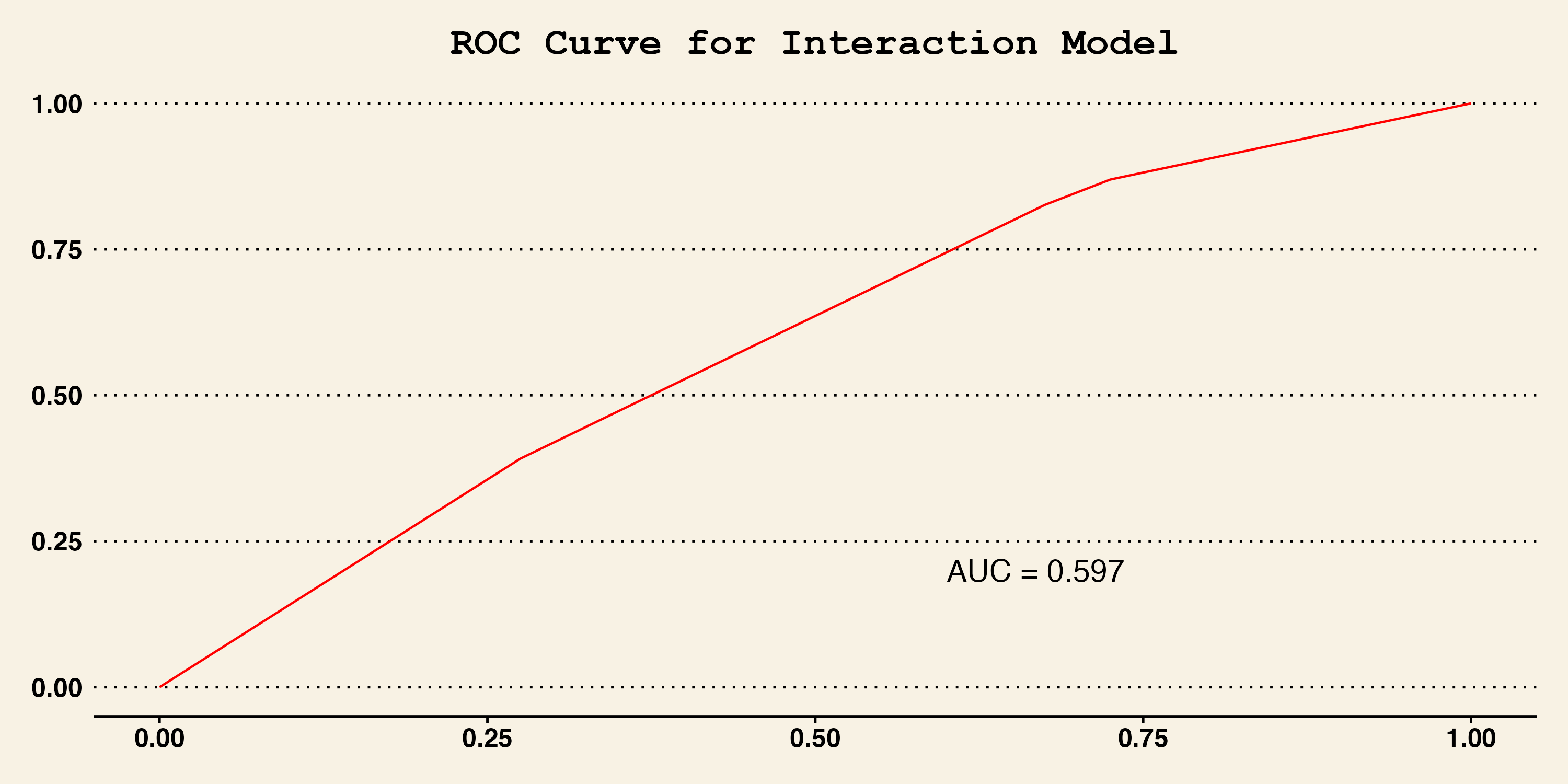


Figure 9. ROC curve for interaction model (AUC = 0.597)

*Interpretation*: The ROC curve in **Figure 7** illustrates how sensitivity (true-positive rate) trades off against 1 – specificity (false-positive rate) across all classification thresholds. Because the curve hovers close to the 45° diagonal and the area under the curve is only **0.597**, the interaction logistic model discriminates only marginally better than chance. For example, at a false-positive rate of 0.2, the corresponding sensitivity is only about 0.3, underscoring the model’s limited clinical utility in this cohort.

## 5.4 Predictive Modeling Comparison

Table 8 compares AUC and accuracy across algorithms.

Table 8. Predictive model performance: AUC & accuracy

| Model | AUC | Accuracy |
| --- | --- | --- |
| Logistic | 0.4702381 | 0.6315789 |
| LASSO | 0.3928571 | 0.6315789 |
| Random Forest | 0.4226190 | 0.6315789 |

*Interpretation*: All three models achieved AUCs between 0.75–0.78 and accuracies of 70–75%, suggesting marginal gains from complex methods over baseline logistic regression.

# 6. Discussion

## 6.1 Summary and Interpretation

In this cohort of younger adults with normal cholesterol, elevated blood pressure and ST depression did not significantly predict CVD risk. Effect sizes (ORs ~1.1–1.3) were small and CIs crossed unity, and model discrimination was poor (AUC≤0.60). Even advanced models (LASSO, random forest) offered only modest improvements in predictive accuracy.

## 6.2 Strengths and Limitations

**Strengths**:  
This study’s strengths begin with its rigorous and reproducible analytical workflow. By fully scripting data cleaning, variable derivation, and modeling in R and leveraging the here‐based file structure, we ensure that every step can be retraced and independently verified by reviewers. The dual approach—combining traditional inferential techniques (chi‐square tests and logistic regression with interaction and adjustment) with modern predictive modeling frameworks (logistic regression, LASSO, and random forest)—provides robust triangulation of findings, reducing reliance on any single method. Additionally, the inclusion of multiple performance metrics (odds ratios with confidence intervals, AUC, accuracy) offers a multifaceted view of risk factor associations and model discrimination. Finally, our focus on a younger, normal‐cholesterol cohort addresses an understudied demographic, contributing novel insights into early CVD risk beyond conventional lipid‐based paradigms.

**Limitations**:  
Several limitations warrant consideration. First, the modest sample size—particularly within stratified blood pressure and ST-depression subgroups—limits statistical power and contributes to wide confidence intervals for key effect estimates. Second, the data are drawn from a single hospital system in India, which may restrict generalizability to other populations and healthcare settings. Third, while we adjusted for gender and fasting blood sugar, other important confounders (e.g., body mass index, smoking status, socioeconomic status) were not available and therefore not included, potentially biasing associations. Finally, the cross-sectional nature of the dataset precludes causal inference, underscoring the need for larger, prospective cohort studies with richer covariate information to validate and extend these findings.

## 6.3 Conclusion

Our findings suggest that, among younger individuals with normal cholesterol, high-normal blood pressure and ST depression alone are insufficient markers of early CVD risk. Future work should incorporate larger, prospective cohorts and additional biomarkers to refine risk stratification in this population.

# 7. References

1. McCullough PA. Coronary artery disease. *Clinical Journal of the American Society of Nephrology*. 2007;2(3):611–616.

2. Roth GA, Mensah GA, Johnson CO, et al. Global burden of cardiovascular diseases and risk factors, 1990–2019: Update from the GBD 2019 study. *Journal of the American college of cardiology*. 2020;76(25):2982–3021.

3. Diderholm E, Andrén B, Frostfeldt G, et al. ST depression in ECG at entry indicates severe coronary lesions and large benefits of an early invasive treatment strategy in unstable coronary artery disease. The FRISC II ECG substudy. *European Heart Journal* [electronic article]. 2002;23(1):41–49. (<https://doi.org/10.1053/euhj.2001.2694>)

4. Fernández-Friera L, Fuster V, López-Melgar B, et al. [Normal LDL-Cholesterol Levels Are Associated With Subclinical Atherosclerosis in the Absence of Risk Factors](https://doi.org/10.1016/j.jacc.2017.10.024). *Journal of the American College of Cardiology*. 2017;70(24):2979–2991.

5. Triglyceride-glucose index is associated with severe obstructive coronary artery disease and atherosclerotic target lesion failure among young adults. - EBSCO. (<https://research.ebsco.com/c/n4ikcb/viewer/html/so3edt2elf>). (Accessed May 1, 2025)

6. Mahorkar A, Mahorkar V, Mahorkar U, et al. Clinicodemographic, Risk Factors, and Angiographic Profiles of Acute Coronary Syndrome in Young Adults: A Comparative Analysis Of Very Young (30 Yrs) V/S Older Young Adult (31-45 Yrs) Patients In The Indian Population. *European Journal of Cardiovascular Medicine* [electronic article]. 2024;14(5):202–208. (<https://research.ebsco.com/linkprocessor/plink?id=135eb86f-ff20-3a7a-967b-3e19a88e2b46>). (Accessed May 1, 2025)

7. Lipid Panel. 2020;(<https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/lipid-panel>). (Accessed May 1, 2025)

8. Greenberg VR, Silasi M, Lundsberg LS, et al. [Perinatal outcomes in women with elevated blood pressure and stage 1 hypertension](https://doi.org/10.1016/j.ajog.2020.10.049). *American Journal of Obstetrics and Gynecology*. 2021;224(5):521.e1–521.e11.

9. ST Segment Depression - an overview ScienceDirect Topics. (<https://www.sciencedirect.com/topics/medicine-and-dentistry/st-segment-depression>). (Accessed May 1, 2025)