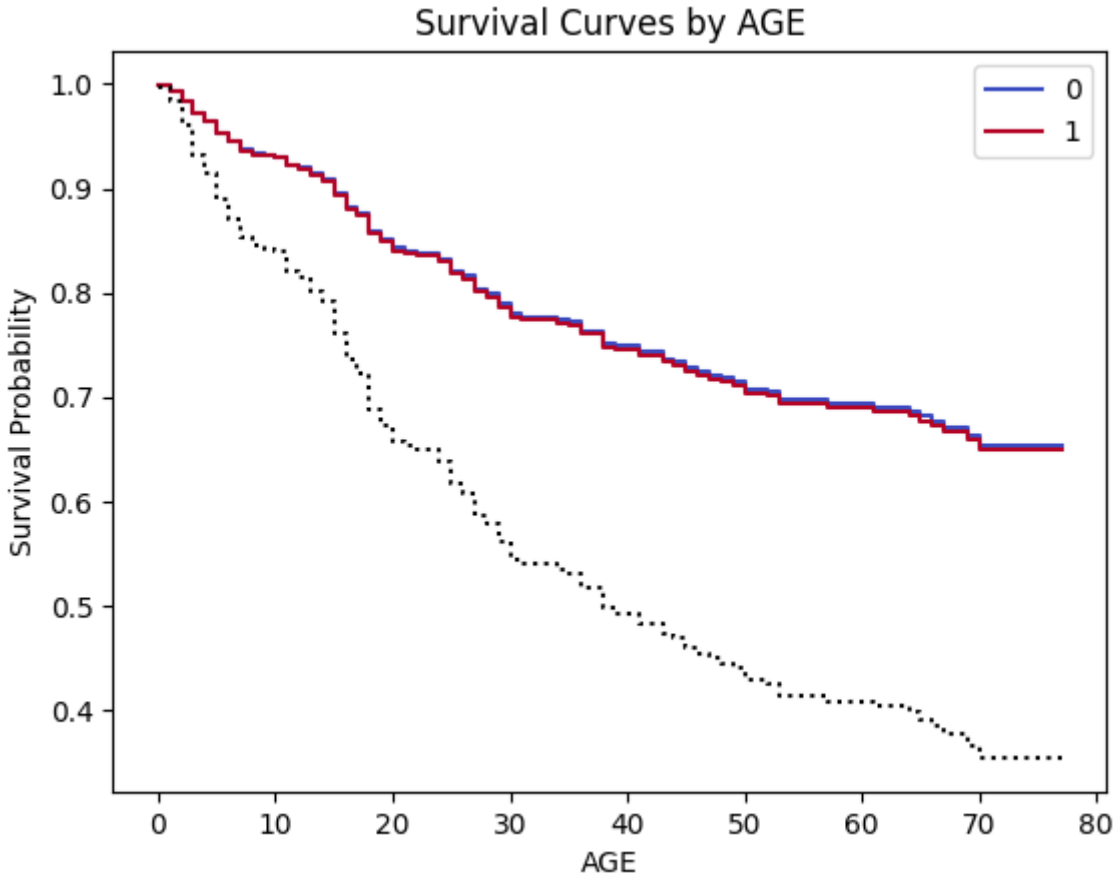
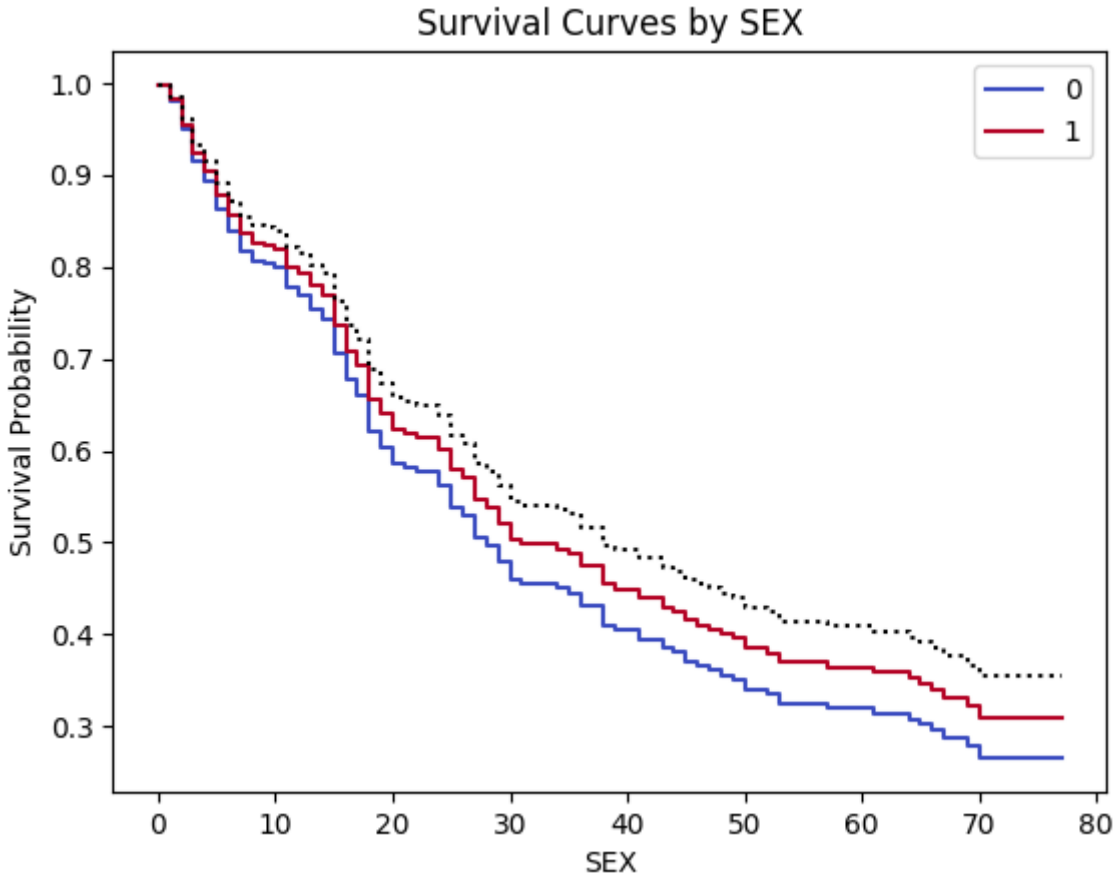
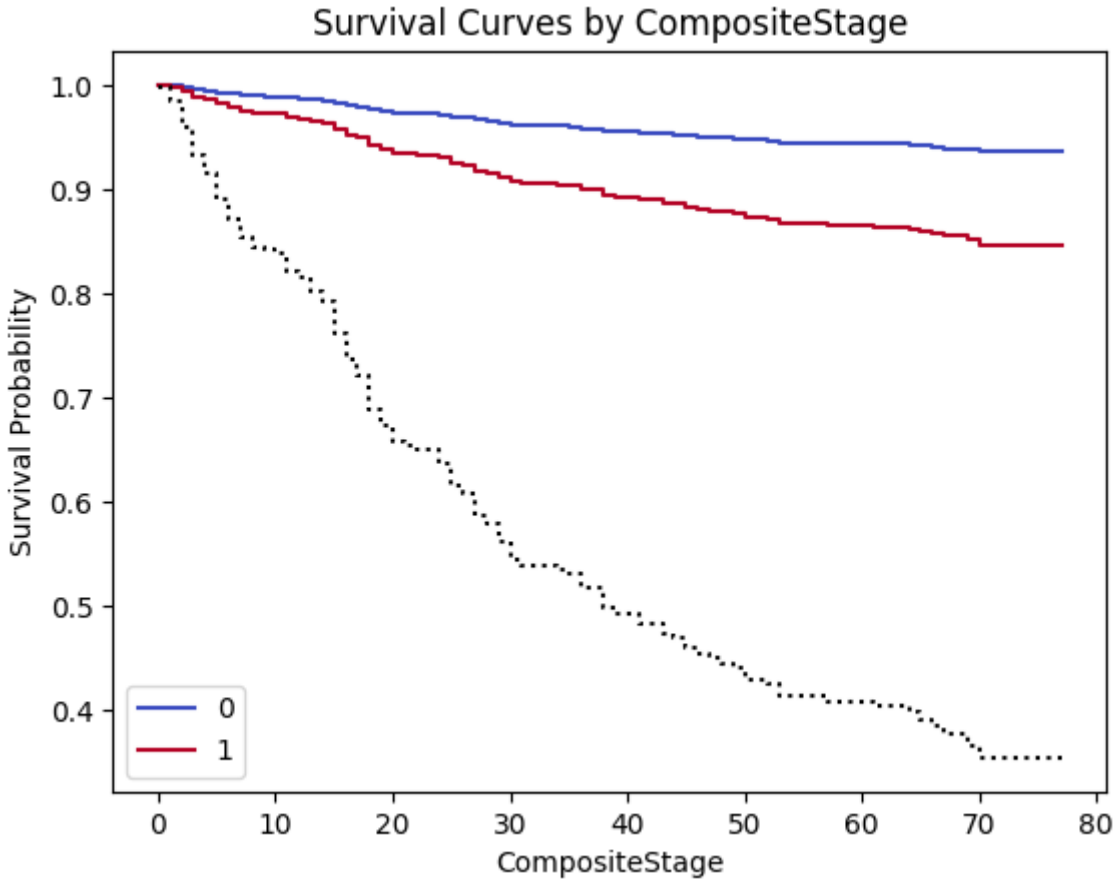
In [7]:
```
import pandas as pd
import numpy as np
from lifelines import CoxPHFitter
import matplotlib.pyplot as plt

# Load the data from the Excel file
data = pd.read_excel('data1.xlsx')

# Create a new instance of the CoxPHFitter class
cph = CoxPHFitter()

# Fit the Cox Proportional Hazard model to the data
cph.fit(data, duration_col='Months', event_col='DEATH')
```

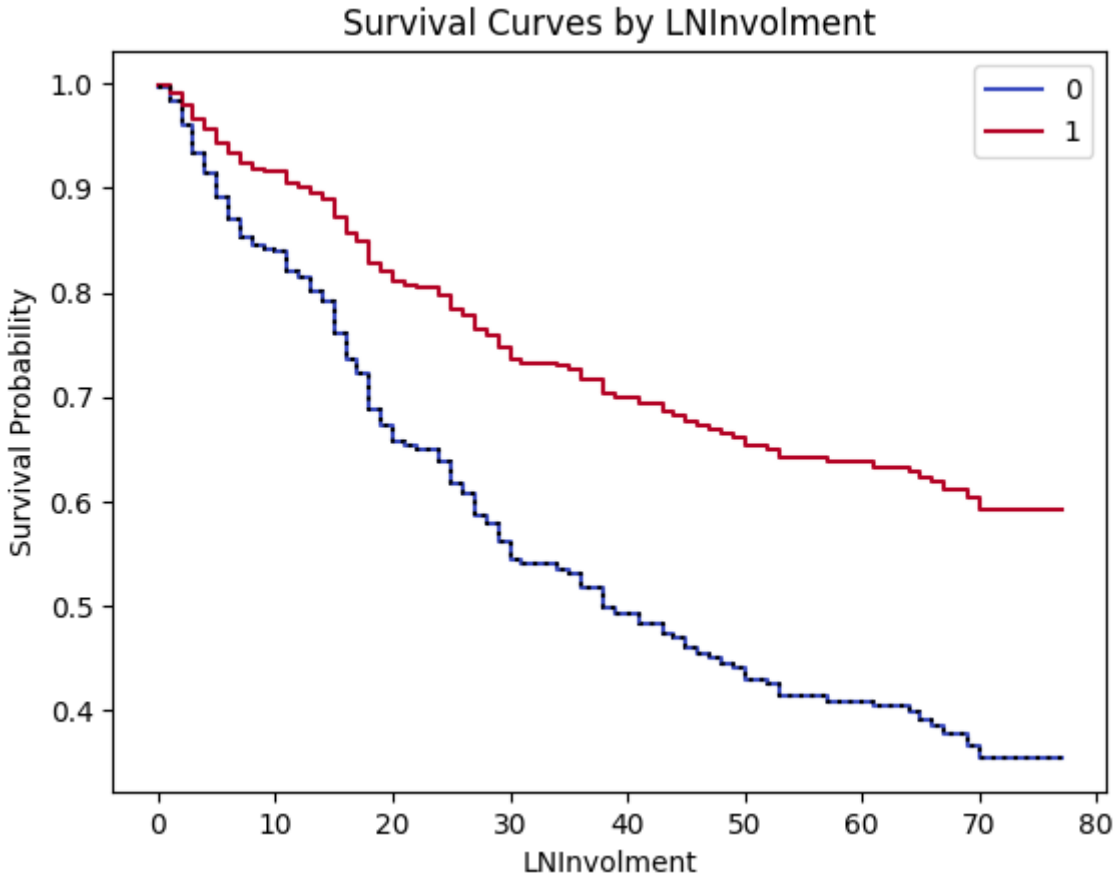
Out[7]:
```
<lifelines.CoxPHFitter: fitted with 343 total observations, 167 right-censored observations>
```

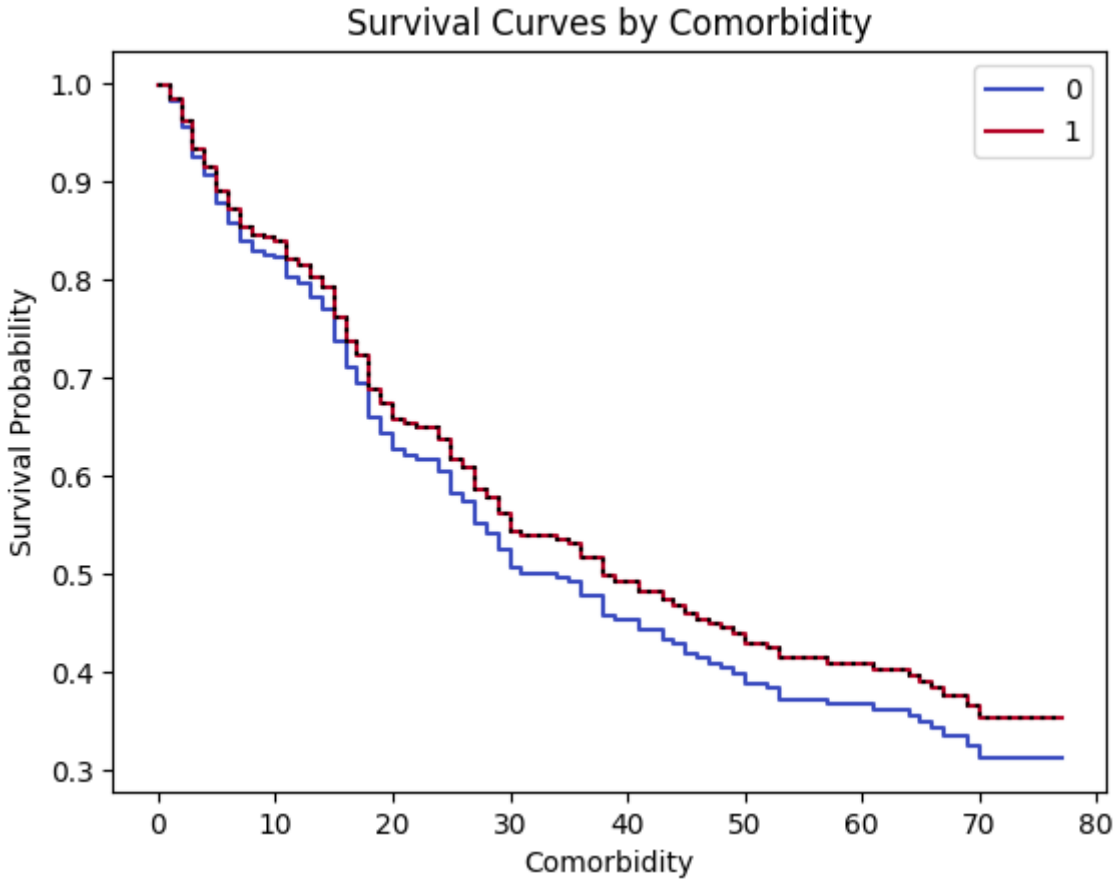
In [2]:
```
covariates_to_plot = ['AGE', 'SEX', 'CompositeStage', 'LNInvolment', 'Comorbidity', 'FamiliyHistoryOfCancer']
for covariate in covariates_to_plot:
    cph.plot_partial_effects_on_outcome(covariates=covariate, values=[0, 1], cmap='coolwarm')
    plt.xlabel(covariate)
    plt.ylabel('Survival Probability')
    plt.title('Survival Curves by ' + covariate)
    plt.legend(['0', '1'])
    plt.show()
```

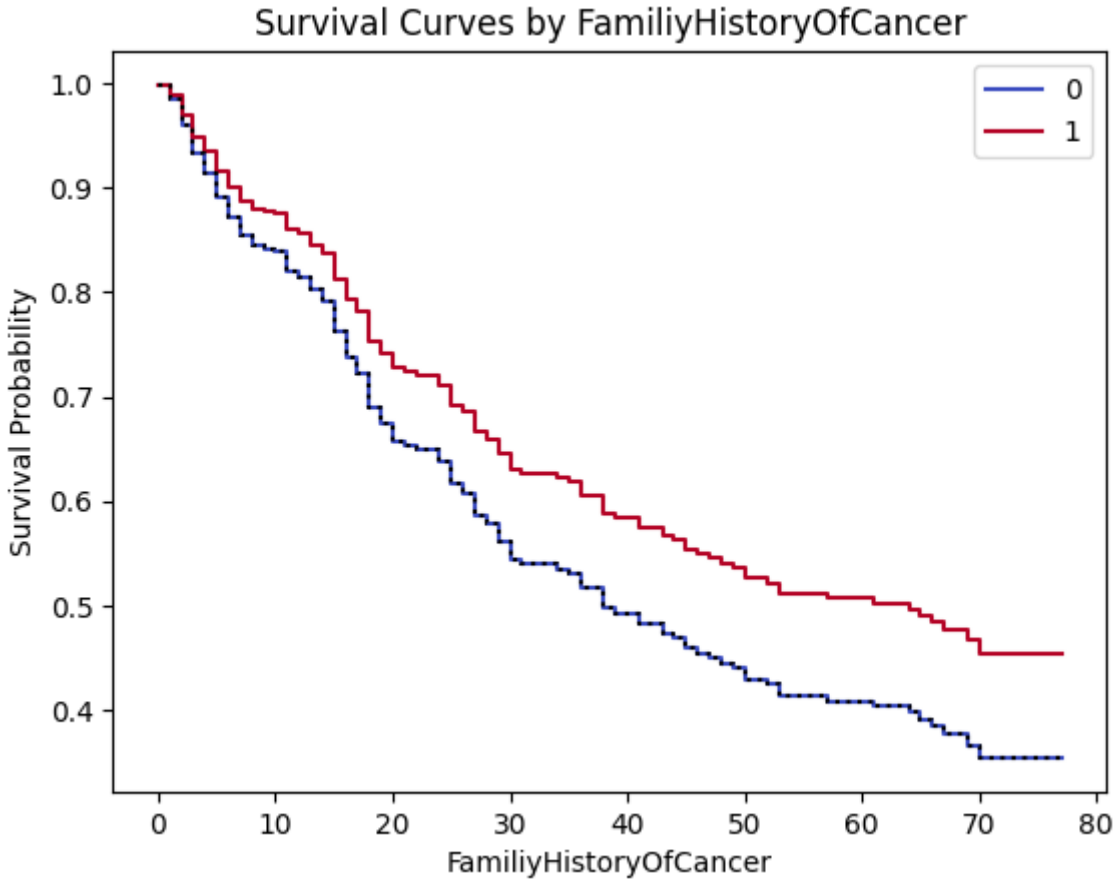

Survival Curves by AGE
Survival Probability
AGE
0
1
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0
10
20
30
40
50
60
70
80


Survival Curves by SEX
Survival Probability
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0
10
20
30
40
50
60
70
80
SEX
0
1

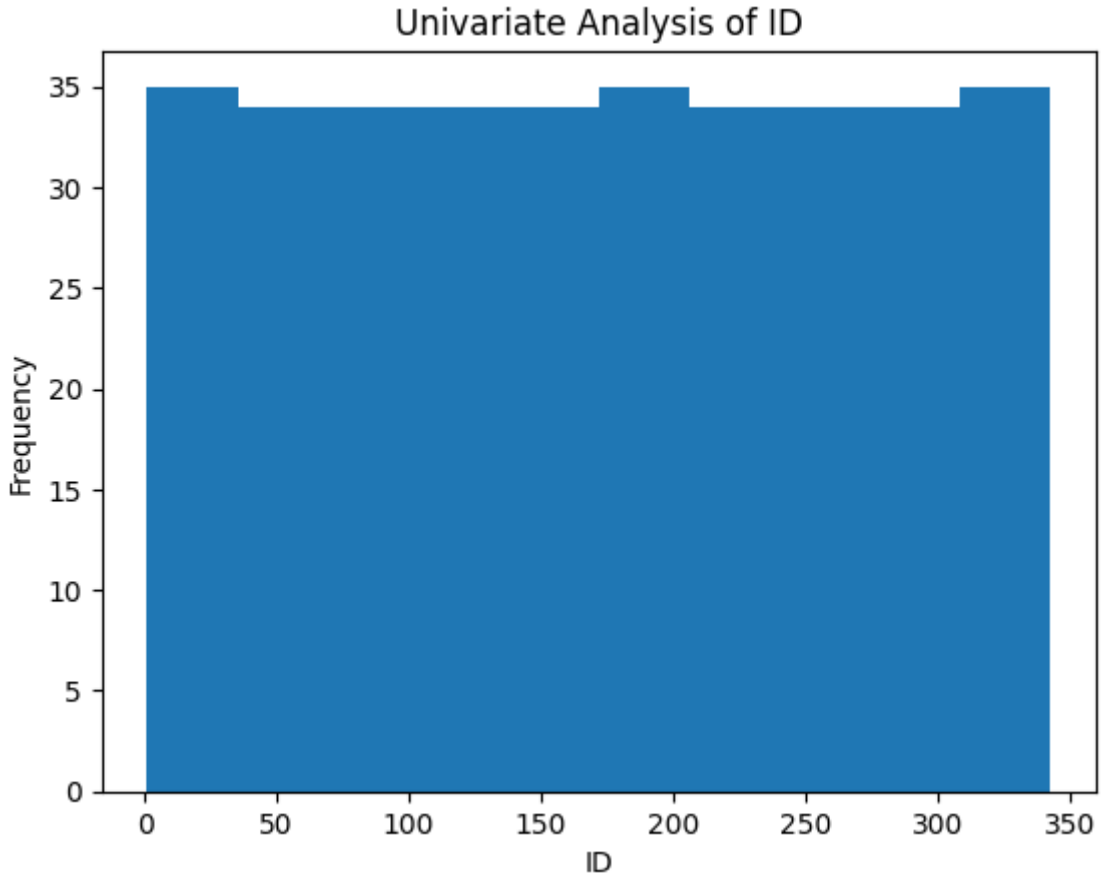

Survival Curves by CompositeStage
Survival Probability
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0
1
0
10
20
30
40
50
60
70
80
CompositeStage

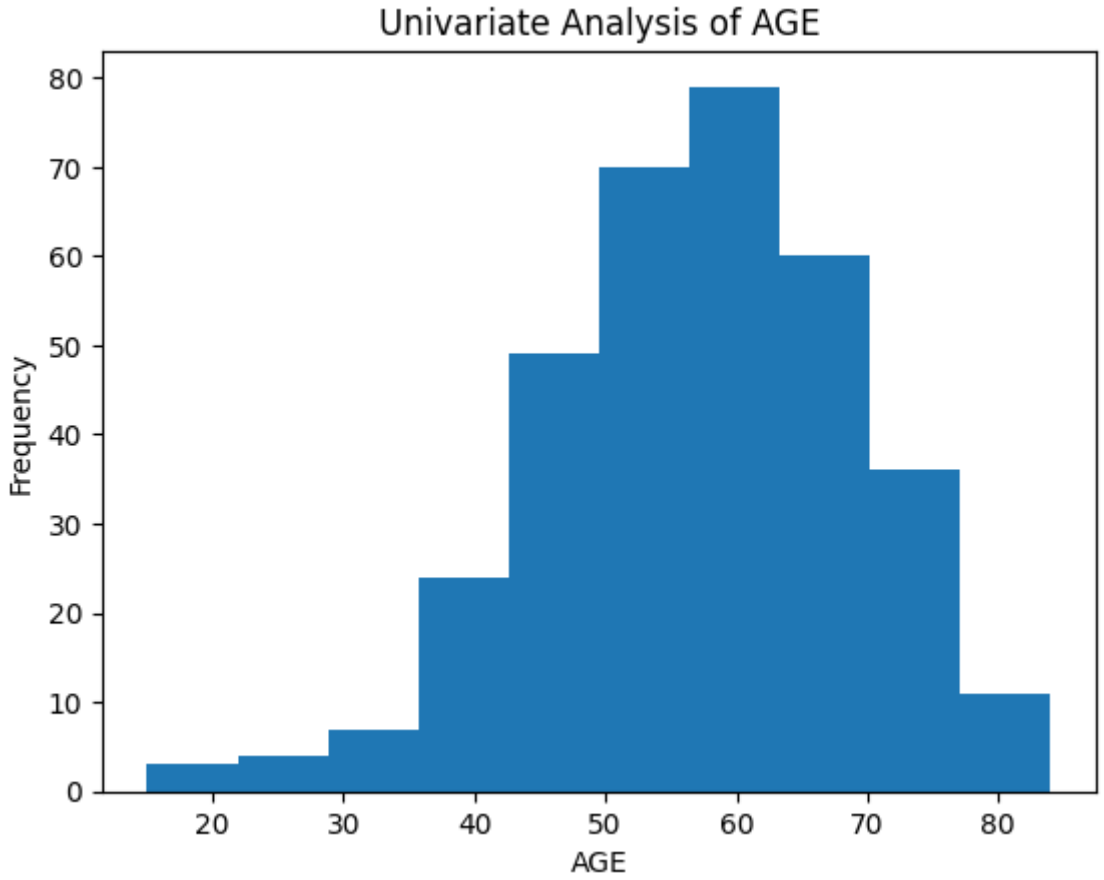

Survival Curves by LNInvolment
Survival Probability
LNInvolment
0
1
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0
10
20
30
40
50
60
70
80

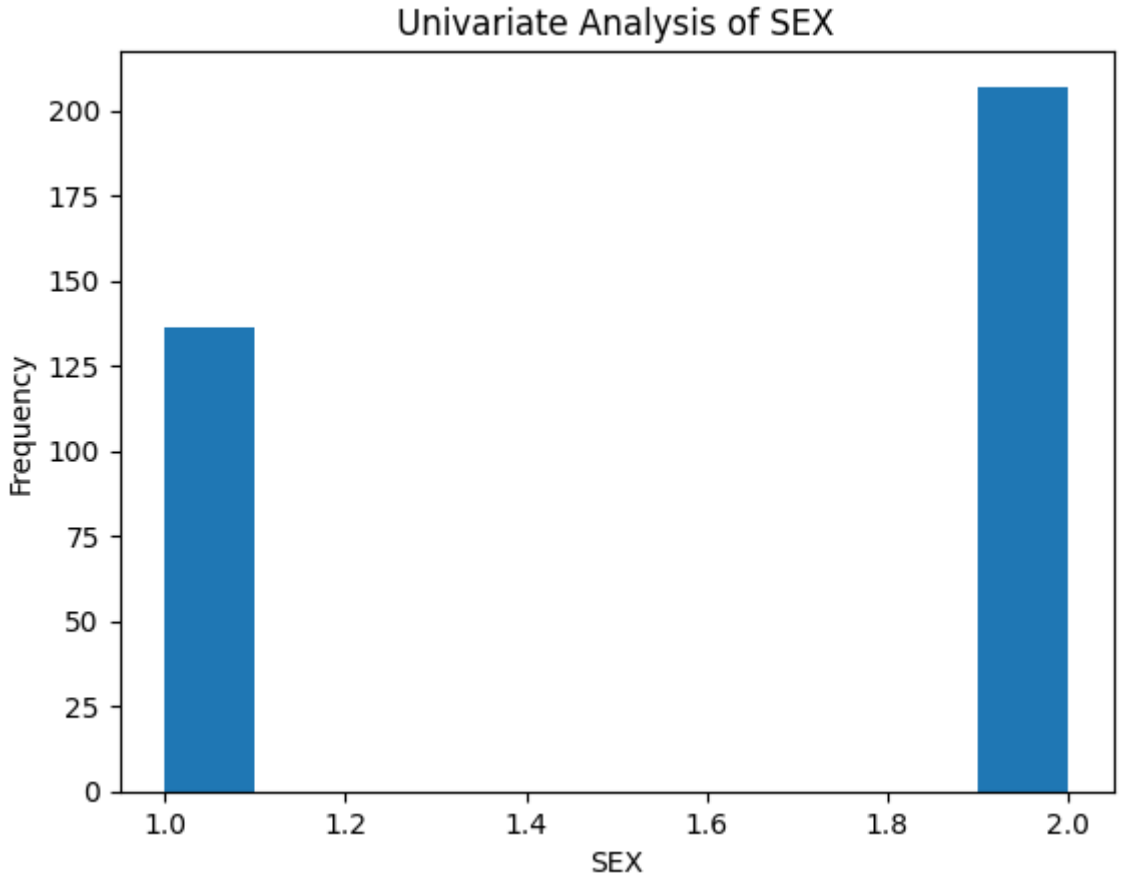

Survival Curves by Comorbidity
Survival Probability
Comorbidity
0
1
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0
10
20
30
40
50
60
70
80

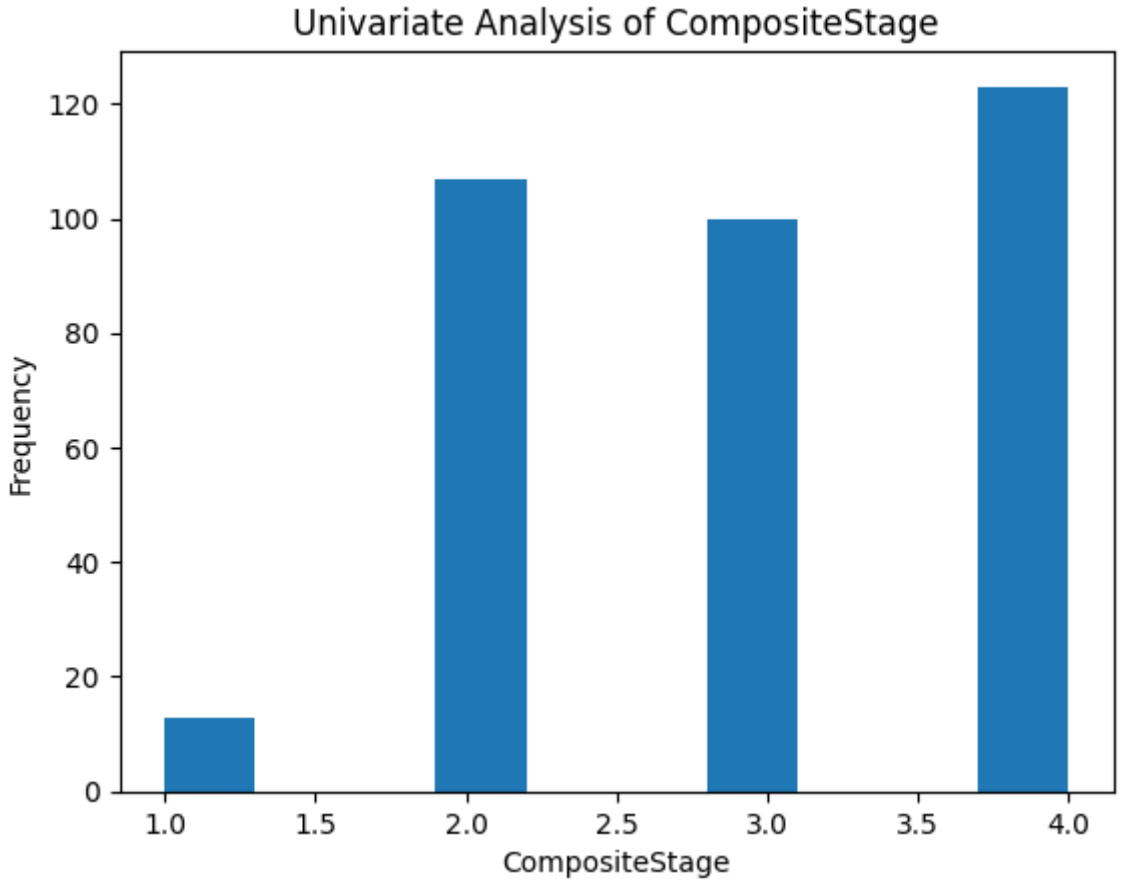


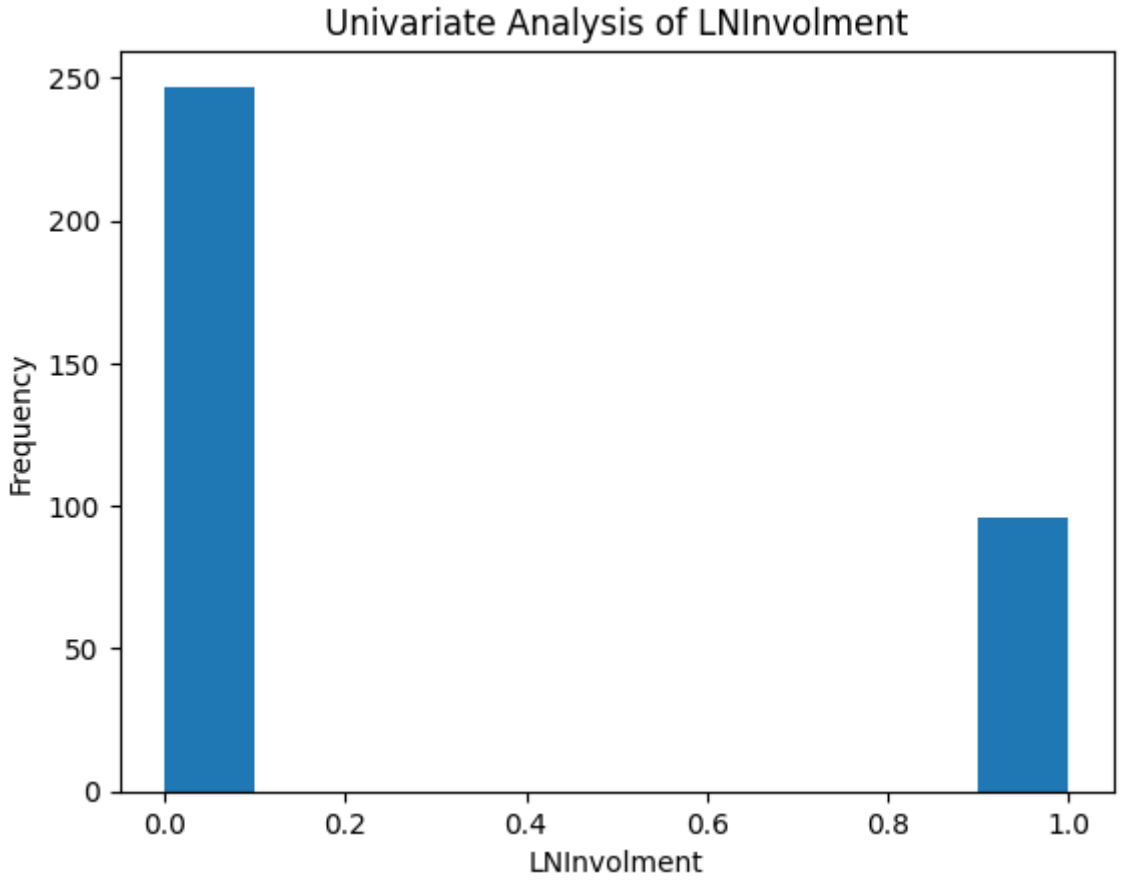
In [3]:
```
for column in data.columns:
    if column not in ['Months', 'DEATH']:
        # Plot the variable
        plt.figure()
        plt.hist(data[column])
        plt.xlabel(column)
        plt.ylabel('Frequency')
        plt.title(f'Univariate Analysis of {column}')
        plt.show()
```

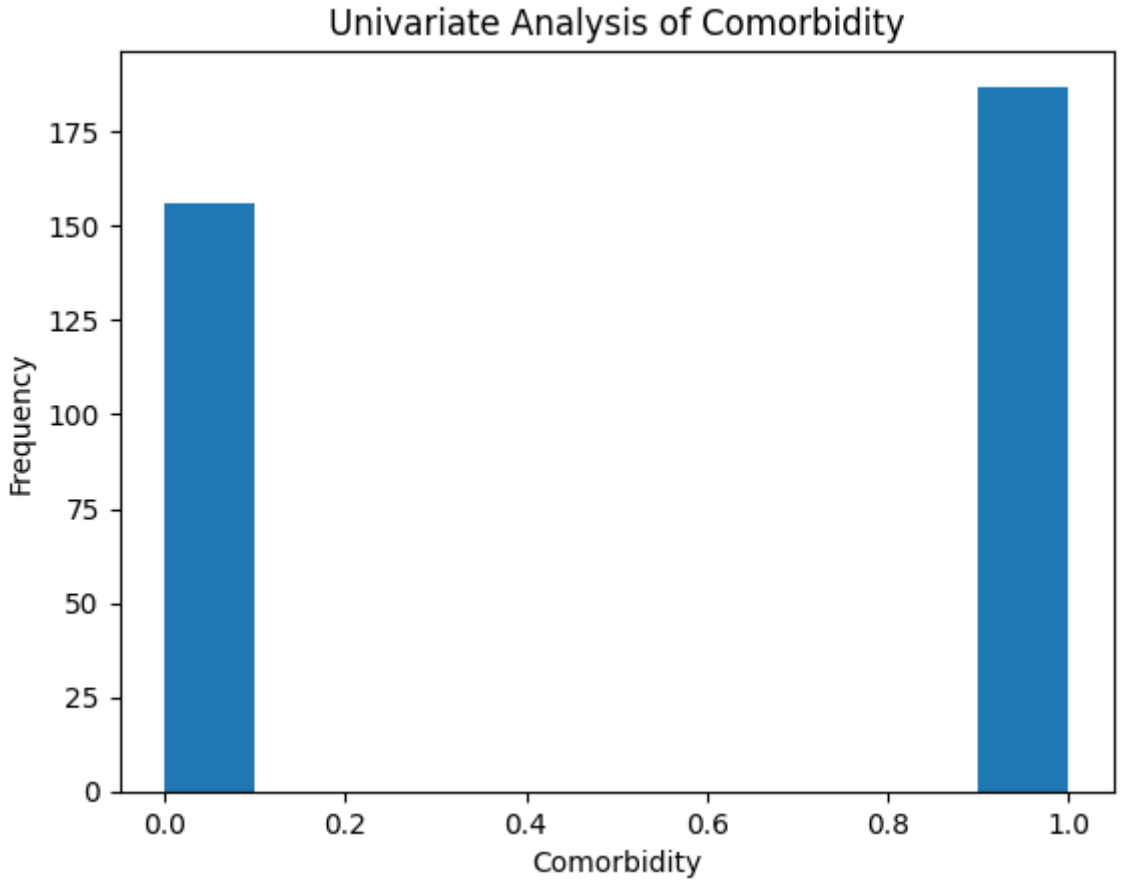

Univariate Analysis of ID
Frequency
ID
0
5
10
15
20
25
30
35
0
50
100
150
200
250
300
350

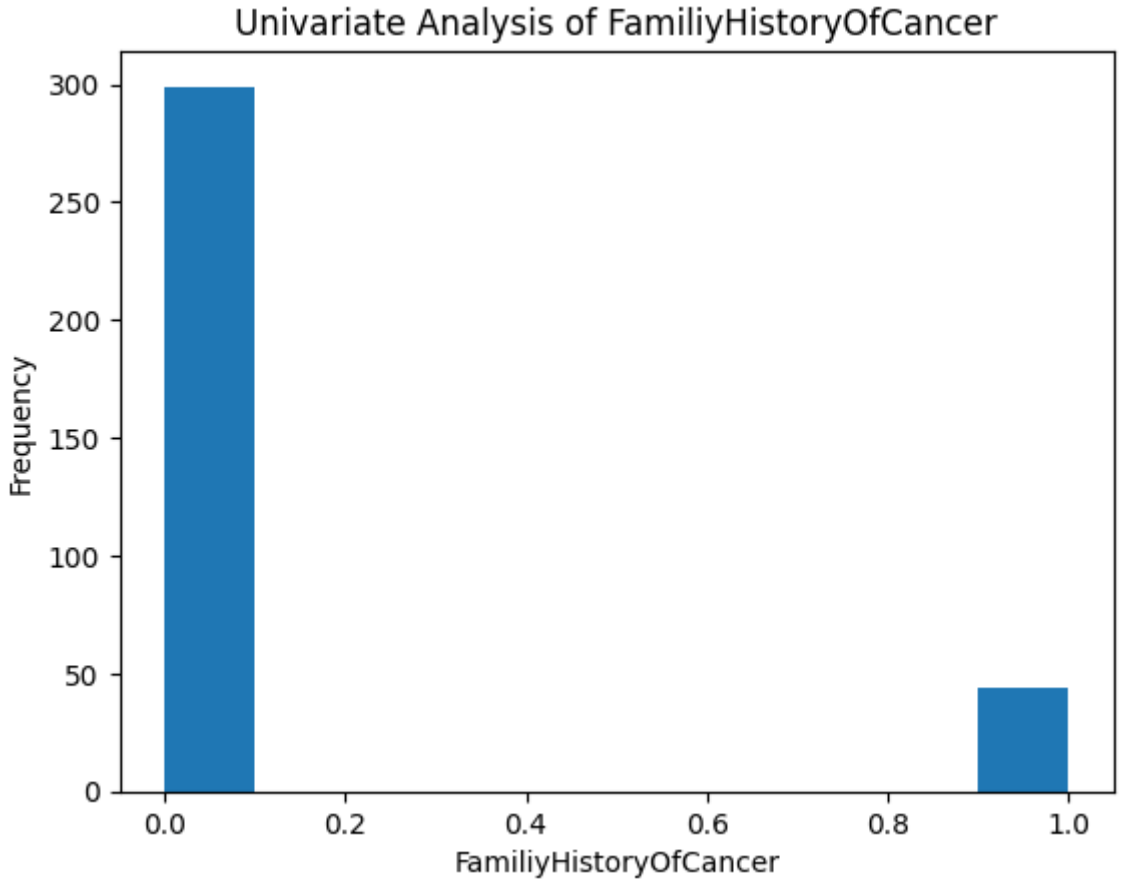

Univariate Analysis of AGE
Frequency
AGE
0
10
20
30
40
50
60
70
80
20
30
40
50
60
70
80


Univariate Analysis of SEX
Frequency
200
175
150
125
100
75
50
25
0
1.0
1.2
1.4
1.6
1.8
2.0
SEX


Univariate Analysis of CompositeStage
Frequency
0
20
40
60
80
100
120
1.0
1.5
2.0
2.5
3.0
3.5
4.0
CompositeStage


Univariate Analysis of LNInvolment
Frequency
250
200
150
100
50
0
0.0
0.2
0.4
0.6
0.8
1.0
LNInvolment


Univariate Analysis of Comorbidity
Frequency
175
150
125
100
75
50
25
0
0.0
0.2
0.4
0.6
0.8
1.0
Comorbidity



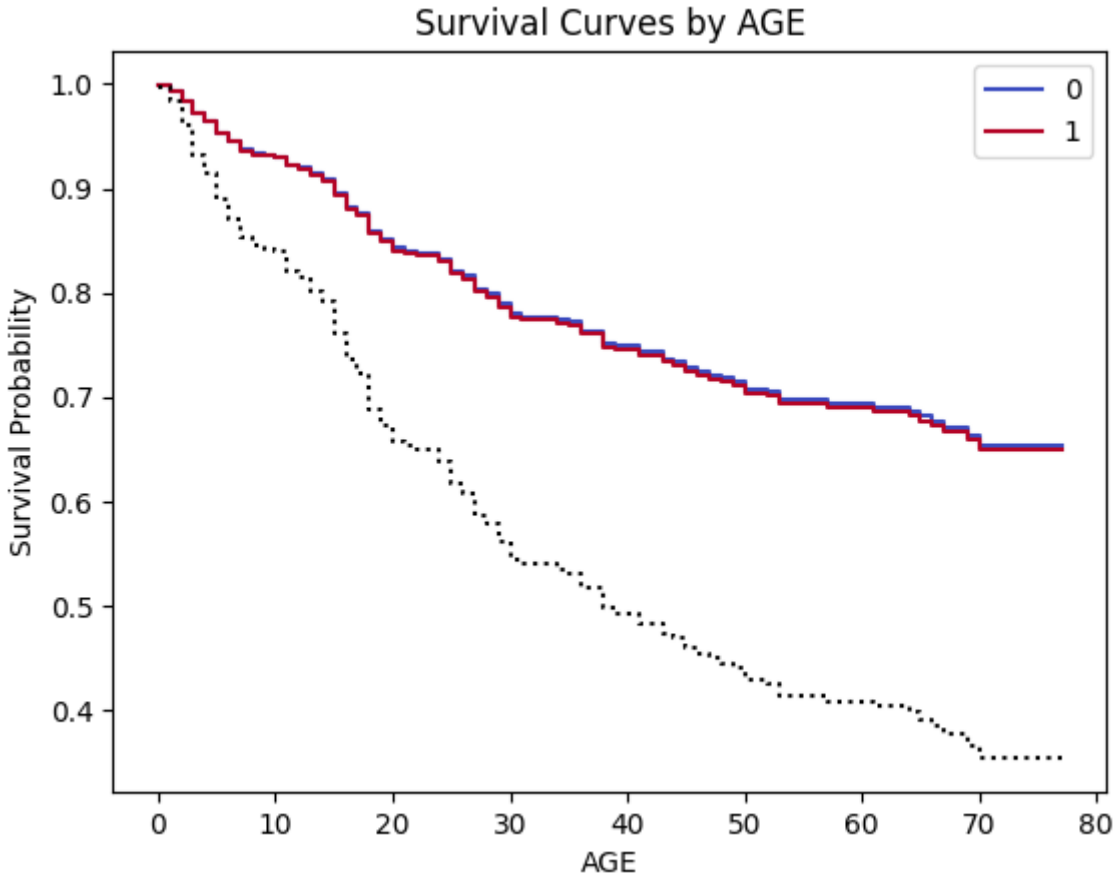
In [5]:
```
n = len(data)
llf = cph.log_likelihood_
k = cph.params_.shape[0]
aic = -2 * llf + 2 * k
bic = -2 * llf + k * np.log(n)
print("AIC of the above given data:", aic)
print("BIC of the above given data:", bic)
```

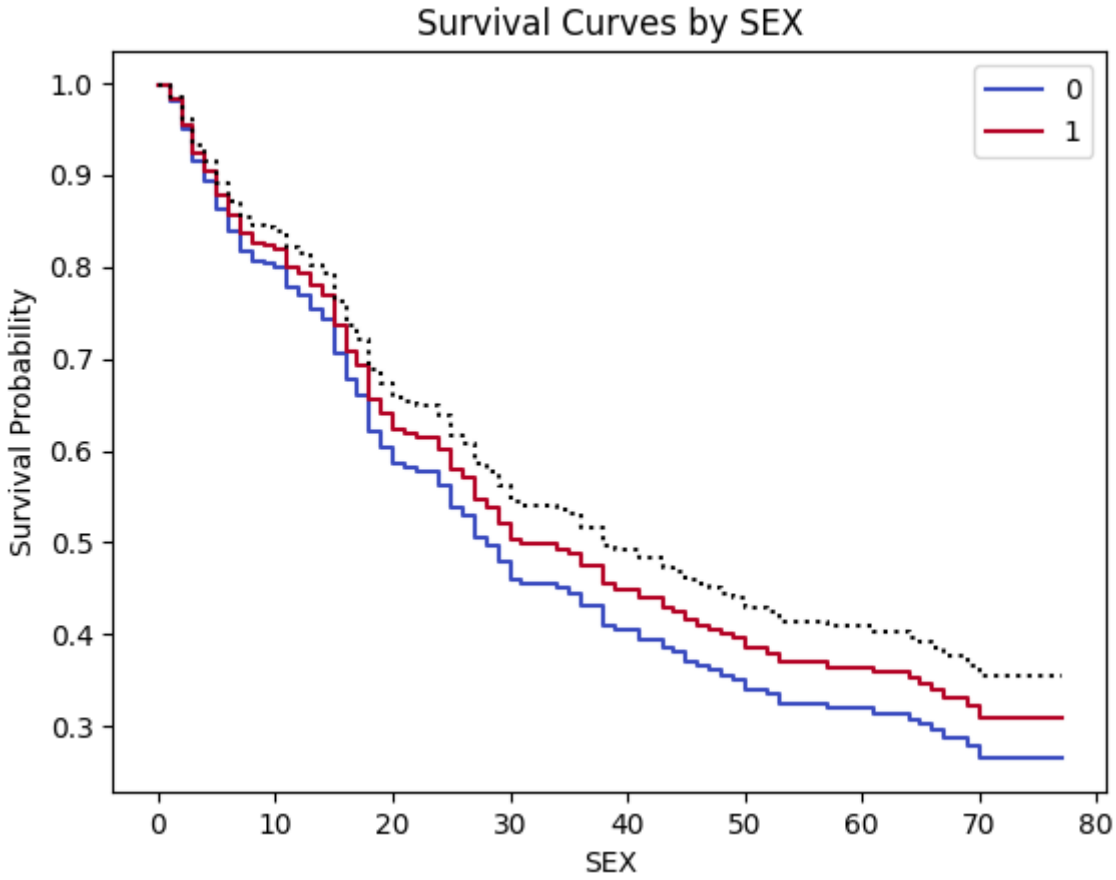
```
AIC of the above given data: 1769.942369337933
BIC of the above given data: 1796.8064824680946
```

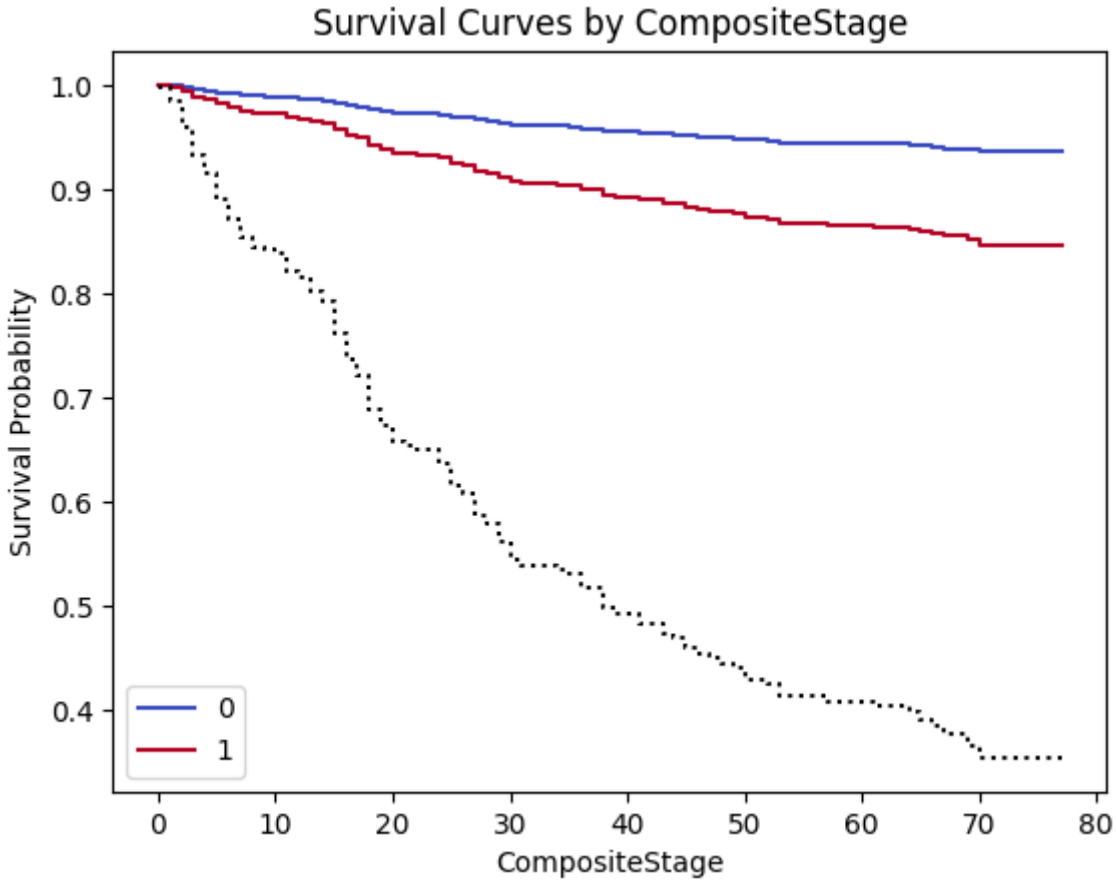
In [12]:
```
import pandas as pd
import numpy as np
```

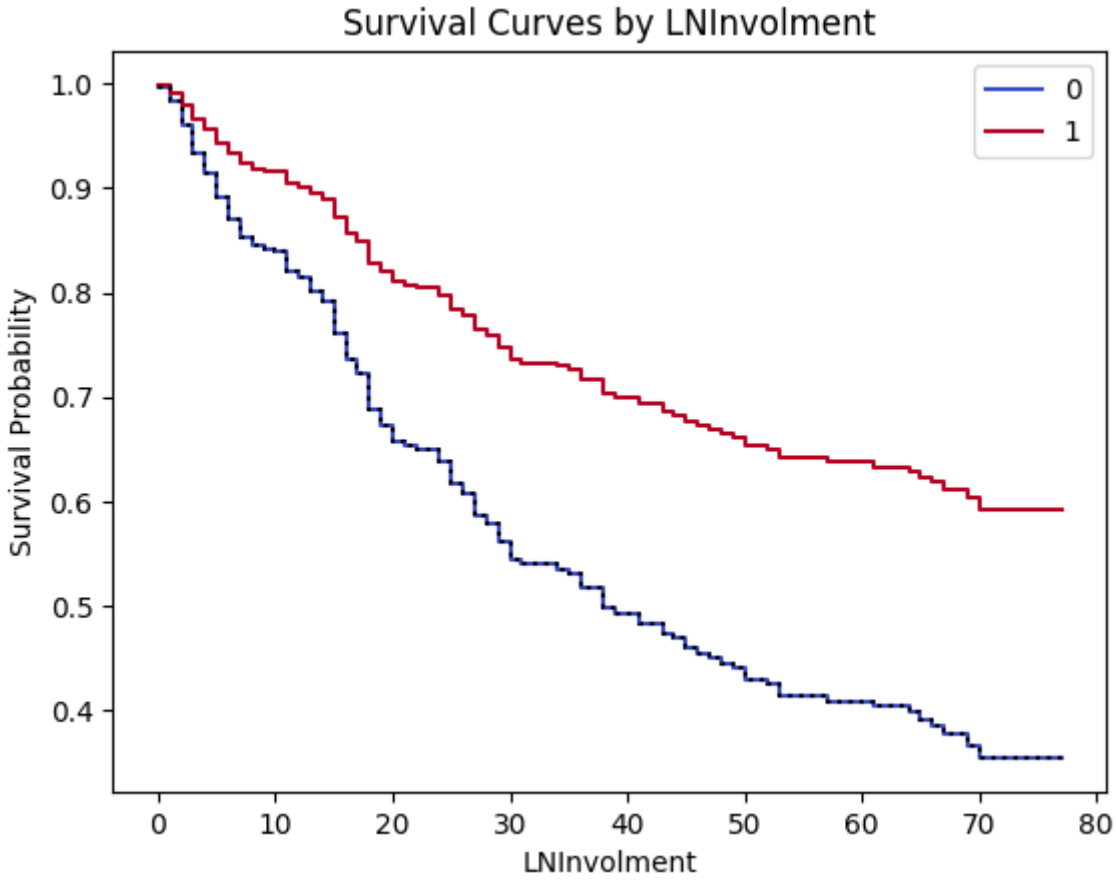
```python
from lifelines import CoxPHFitter
import matplotlib.pyplot as plt

# Load the data from the Excel file
data = pd.read_excel('data1.xlsx')

# Create a new instance of the CoxPHFitter class
cph = CoxPHFitter()

# Fit the Cox Proportional Hazard model to the data
cph.fit(data, duration_col='Months', event_col='DEATH')

covariates_to_plot = ['AGE', 'SEX', 'CompositeStage', 'LNInvolment', 'Comorbidity', 'FamiliyHistoryOfCancer']
significant_covariates = []

for covariate in covariates_to_plot:
    cph.plot_partial_effects_on_outcome(covariates=covariate, values=[0, 1], cmap='coolwarm')
    plt.xlabel(covariate)
    plt.ylabel('Survival Probability')
    plt.title('Survival Curves by ' + covariate)
    plt.legend(['0', '1'])
    plt.show()

    # Prompt user to input whether the covariate is significant or not
    is_significant = input("Is " + covariate + " significant? (y/n): ")

    if is_significant.lower() == 'y':
        significant_covariates.append(covariate)

# Perform multivariate analysis
multivariate_data = data[significant_covariates + ['Months', 'DEATH']]
cph_multivariate = CoxPHFitter()
cph_multivariate.fit(multivariate_data, duration_col='Months', event_col='DEATH')

n = len(data)
llf_univariate = cph.log_likelihood_
k_univariate = cph.params_.shape[0]
aic_univariate = -2 * llf_univariate + 2 * k_univariate
bic_univariate = -2 * llf_univariate + k_univariate * np.log(n)
```

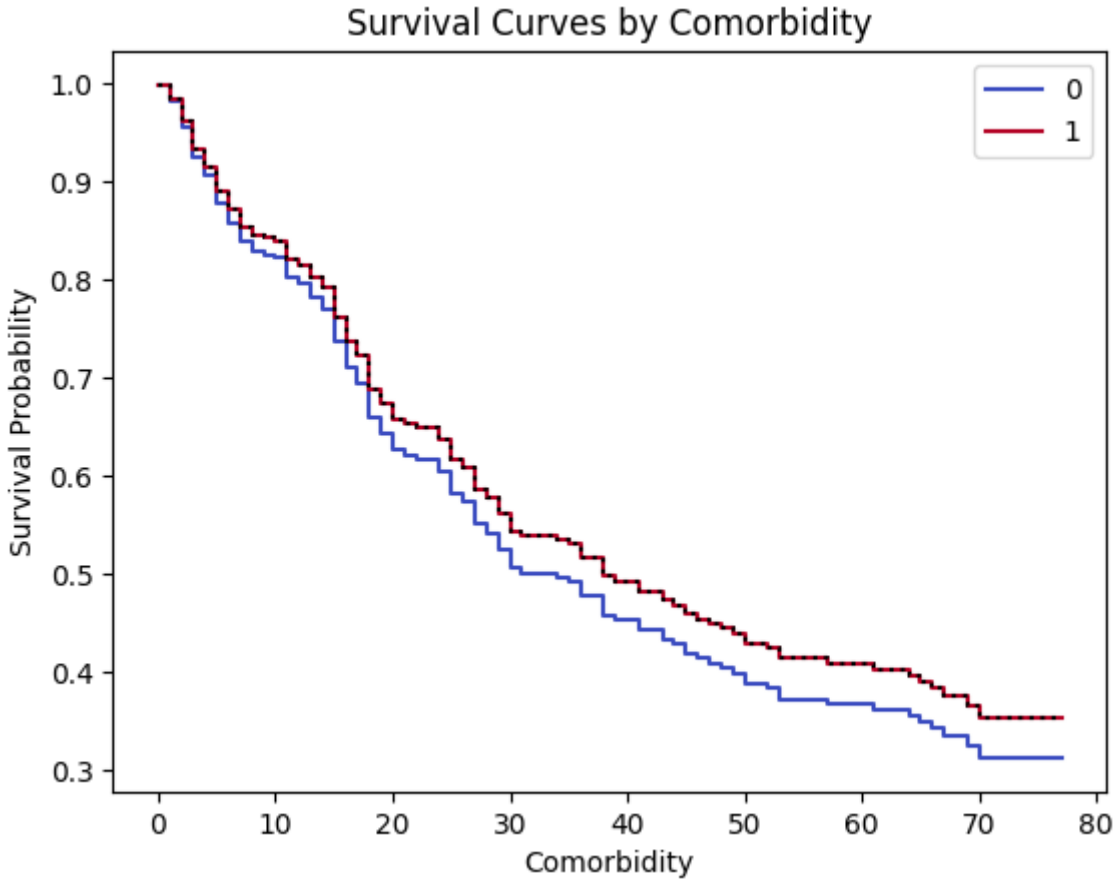
```
llf_multivariate = cph_multivariate.log_likelihood_
k_multivariate = cph_multivariate.params_.shape[0]
aic_multivariate = -2 * llf_multivariate + 2 * k_multivariate
bic_multivariate = -2 * llf_multivariate + k_multivariate * np.log(n)

print("Univariate Analysis:")
print("Significant Covariates:", significant_covariates)
print("AIC of the univariate model:", aic_univariate)
print("BIC of the univariate model:", bic_univariate)

print("\nMultivariate Analysis:")
print(cph_multivariate.summary)
print("AIC of the multivariate model:", aic_multivariate)
print("BIC of the multivariate model:", bic_multivariate)
```

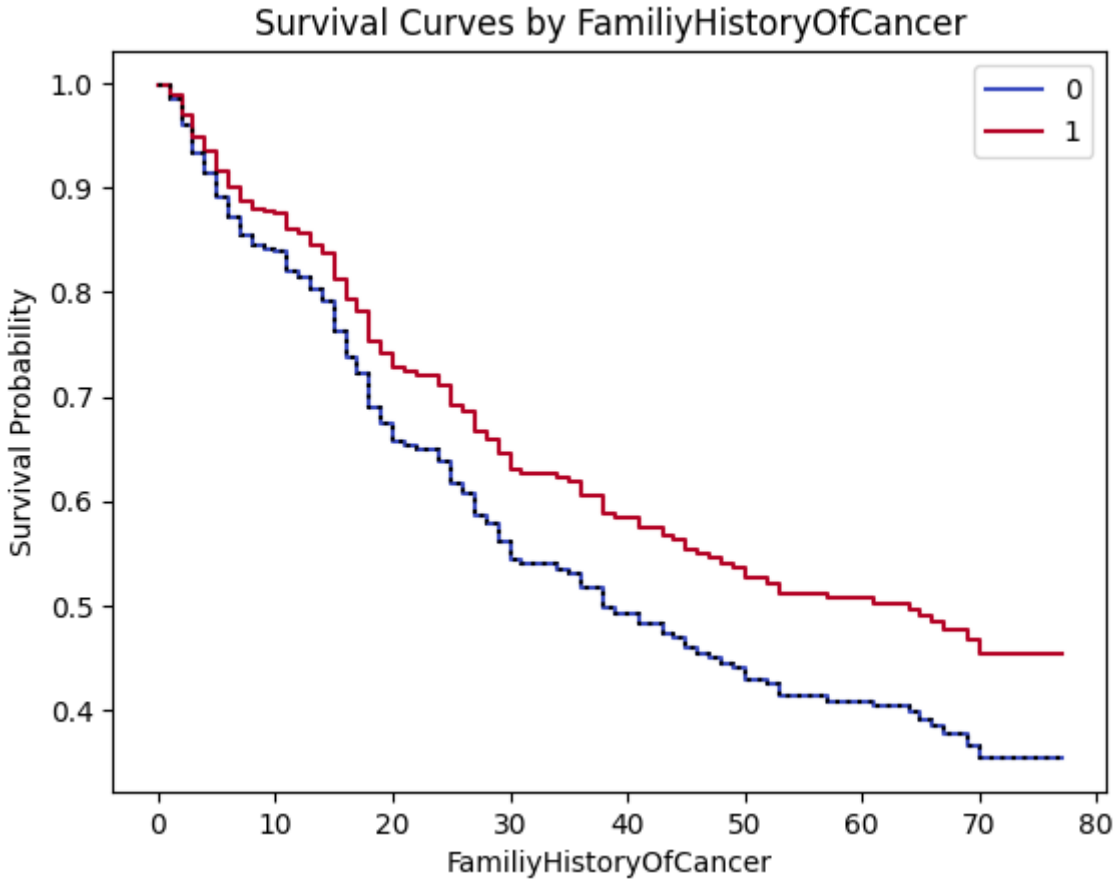

Survival Curves by AGE
Survival Probability
AGE
0
1
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0
10
20
30
40
50
60
70
80


Survival Curves by SEX
Survival Probability
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0
10
20
30
40
50
60
70
80
SEX
0
1


Survival Curves by CompositeStage
Survival Probability
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0
1
0
10
20
30
40
50
60
70
80
CompositeStage


Survival Curves by LNInvolment
Survival Probability
LNInvolment
0
1
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0
10
20
30
40
50
60
70
80


Survival Curves by Comorbidity
Survival Probability
Comorbidity
0
1
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0
10
20
30
40
50
60
70
80


Survival Curves by FamiliyHistoryOfCancer
Survival Probability
FamiliyHistoryOfCancer
0
1
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0
10
20
30
40
50
60
70
80

```
Univariate Analysis:
Significant Covariates: ['AGE']
AIC of the univariate model: 1769.942369337933
BIC of the univariate model: 1796.8064824680946

Multivariate Analysis:
               coef  exp(coef)  se(coef)  coef lower 95%  coef upper 95%  \
covariate
AGE        0.009362   1.009406  0.006565       -0.003506         0.02223

           exp(coef) lower 95%  exp(coef) upper 95%  cmp to        z  \
covariate
AGE                    0.9965             1.022479     0.0  1.42599

                 p  -log2(p)
covariate
AGE        0.153871  2.700206
AIC of the multivariate model: 1887.3267837410287
BIC of the multivariate model: 1891.1645141881947
```

In [ ]: