In [1]:

```
import pandas as pd
import numpy as np
from lifelines import CoxPHFitter
import matplotlib.pyplot as plt

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

# Perform univariate analysis
univariate_results = []
univariate_aic_bic = []
for covariate in data.columns:
    if covariate not in ['Months', 'DEATH']:
        cph_univariate = CoxPHFitter()
        cph_univariate.fit(data[[covariate, 'Months', 'DEATH']], duration_col='M

        # Calculate log-likelihood
        llf = cph_univariate.log_likelihood_

        # Calculate number of parameters
        k = cph_univariate.params_.shape[0]

        # Calculate AIC and BIC
        n = len(data)
        aic = -2 * llf + 2 * k
        bic = -2 * llf + k * np.log(n)

        univariate_results.append((covariate, cph_univariate.summary['p'][covari
        univariate_aic_bic.append((covariate, aic, bic))

# Sort the univariate models based on AIC
univariate_aic_bic.sort(key=lambda x: x[1])

# Select significant variables based on p-values
significant_variables = [(var, p_value) for var, p_value in univariate_results i

# Fit the multivariate model
multivariate_data = data[[var for var, _ in significant_variables] + ['Months',
cph_multivariate = CoxPHFitter()
cph_multivariate.fit(multivariate_data, duration_col='Months', event_col='DEATH'

# Calculate log-likelihood for multivariate model
llf_multivariate = cph_multivariate.log_likelihood_

# Calculate number of parameters for multivariate model
k_multivariate = cph_multivariate.params_.shape[0]

# Calculate AIC and BIC for multivariate model
n = len(data)
aic_multivariate = -2 * llf_multivariate + 2 * k_multivariate
bic_multivariate = -2 * llf_multivariate + k_multivariate * np.log(n)

# Print AIC and BIC for univariate models
print("AIC and BIC for univariate models:")
for covariate, aic, bic in univariate_aic_bic:
    print(f"{covariate}: AIC={aic}, BIC={bic}")

# Print significant variables from univariate analysis
```

```
print("\nSignificant variables from univariate analysis:")
for var, p_value in significant_variables:
    print(f"{var}: p-value={p_value}")

# Print AIC and BIC for multivariate model
print("\nAIC and BIC for multivariate model:")
print("AIC of the multivariate model:", aic_multivariate)
print("BIC of the multivariate model:", bic_multivariate)
```

```
AIC and BIC for univariate models:
CompositeStage: AIC=1801.0425436924681, BIC=1804.8802741396341
ID: AIC=1854.1291924816765, BIC=1857.9669229288425
LNInvolment: AIC=1877.7266674202021, BIC=1881.5643978673681
AGE: AIC=1887.3267837410287, BIC=1891.1645141881947
Comorbidity: AIC=1888.099387799303, BIC=1891.937118246469
FamiliyHistoryOfCancer: AIC=1889.2113357577587, BIC=1893.0490662049247
SEX: AIC=1889.2852518349302, BIC=1893.1229822820962

Significant variables from univariate analysis:
ID: p-value=5.656542610612734e-09
CompositeStage: p-value=2.724672855669092e-18
LNInvolment: p-value=0.0011578069966356516

AIC and BIC for multivariate model:
AIC of the multivariate model: 1768.6487964813414
BIC of the multivariate model: 1780.1619878228391
```

In [ ]: