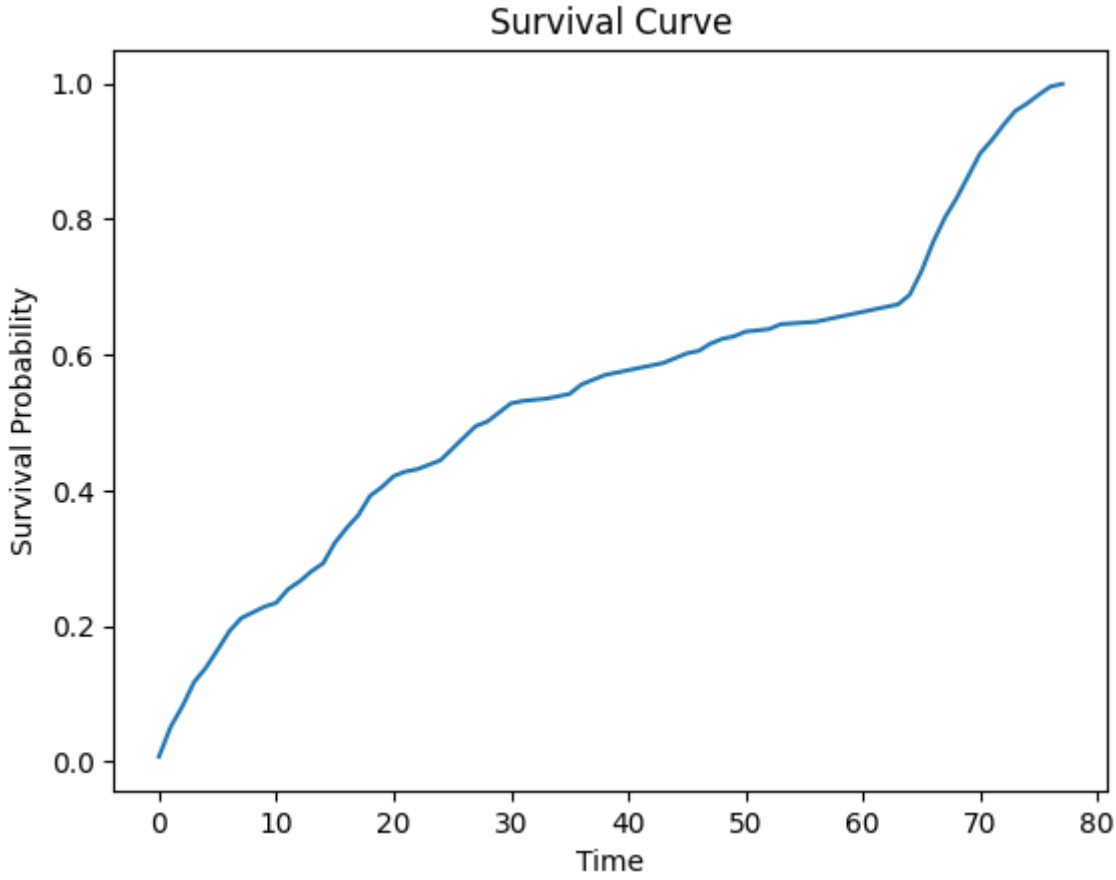
In [ ]:

In [2]:
```
import pandas as pd
from lifelines import CoxPHFitter
from sklearn.impute import SimpleImputer
from sklearn.preprocessing import StandardScaler
import matplotlib.pyplot as plt

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

# Preprocess the data Drop any rows with missing values in the columns of intere
data = data.dropna(subset=['Months', 'DEATH', 'AGE', 'SEX', 'CompositeStage', 'L
# Handle missing values in other columns
imputer = SimpleImputer(strategy='median')
data[['DEATH', 'AGE', 'CompositeStage', 'LNInvolment', 'Comorbidity']] = imputer
# Standardize the covariates
scaler = StandardScaler()
data[['DEATH', 'AGE', 'CompositeStage', 'LNInvolment', 'Comorbidity']] = scaler.
# Create a new DataFrame with the required columns for the Buckley-James estimat
buckley_james_data = data[['Months', 'DEATH', 'AGE', 'SEX', 'CompositeStage', 'L
#print(buckley_james_data.isnull().sum()) #find no of Nan
# Fit the Buckley-James model with custom options
#buckley_james_data = buckley_james_data.drop(['tstart'], axis=1)
#buckley_james_data = buckley_james_data.drop('tstart', axis=1)
cph = CoxPHFitter(penalizer=0.1)  # Set the penalizer parameter to control overf
cph.fit(buckley_james_data, 'Months', 'DEATH', show_progress=True)  # Set the st
# Print the estimated coefficients
print(cph.summary)
# Access other properties of the fitted model (e.g., hazard ratios, p-values)
# For example, to get the hazard ratios:
print(cph.hazard_ratios_)
# Make predictions using the fitted model
# For example, to predict the survival probability at a specific time point for
new_patient_data = pd.DataFrame({'AGE': [90], 'SEX': [1], 'CompositeStage': [2],
partial_hazard = cph.predict_partial_hazard(new_patient_data)
survival_prob = 1 - cph.baseline_survival_
plt.plot(cph.baseline_survival_.index, survival_prob.values)
plt.xlabel('Time')
plt.ylabel('Survival Probability')
plt.title('Survival Curve')
plt.show()
# Perform other analyses or visualizations as needed
```

```
Iteration 1: norm_delta = 0.66384, step_size = 0.9500, log_lik = -1663.17959, new
ton_decrement = 46.04648, seconds_since_start = 0.0
Iteration 2: norm_delta = 0.03630, step_size = 0.9500, log_lik = -1620.53093, new
ton_decrement = 0.19362, seconds_since_start = 0.0
Iteration 3: norm_delta = 0.00176, step_size = 0.9500, log_lik = -1620.33817, new
ton_decrement = 0.00043, seconds_since_start = 0.0
Iteration 4: norm_delta = 0.00000, step_size = 1.0000, log_lik = -1620.33774, new
ton_decrement = 0.00000, seconds_since_start = 0.0
Convergence success after 4 iterations.
                           coef  exp(coef)  se(coef)  coef lower 95%  \
covariate
AGE                     0.019975   1.020175  0.055896       -0.089580
SEX                     0.027013   1.027381  0.106745       -0.182203
CompositeStage          0.531571   1.701603  0.061434        0.411162
LNInvolment            -0.275748   0.759004  0.053051       -0.379725
Comorbidity            -0.034023   0.966549  0.054884       -0.141594
FamiliyHistoryOfCancer  0.003465   1.003471  0.156806       -0.303870

                        coef upper 95%  exp(coef) lower 95%  \
covariate
AGE                           0.129529             0.914315
SEX                           0.236229             0.833432
CompositeStage                0.651980             1.508570
LNInvolment                  -0.171771             0.684049
Comorbidity                   0.073548             0.867974
FamiliyHistoryOfCancer        0.310800             0.737957

                        exp(coef) upper 95%  cmp to         z             p  \
covariate
AGE                                1.138292     0.0  0.357349  7.208303e-01
SEX                                1.266464     0.0  0.253064  8.002191e-01
CompositeStage                     1.919337     0.0  8.652682  5.030319e-18
LNInvolment                        0.842172     0.0 -5.197833  2.016254e-07
Comorbidity                        1.076320     0.0 -0.619903  5.353217e-01
FamiliyHistoryOfCancer             1.364517     0.0  0.022100  9.823684e-01

                         -log2(p)
covariate
AGE                      0.472268
SEX                      0.321533
CompositeStage          57.464056
LNInvolment             22.241820
Comorbidity              0.901522
FamiliyHistoryOfCancer   0.025664
covariate
AGE                       1.020175
SEX                       1.027381
CompositeStage            1.701603
LNInvolment               0.759004
Comorbidity               0.966549
FamiliyHistoryOfCancer    1.003471
Name: exp(coef), dtype: float64
```


Survival Curve
Survival Probability
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
50
60
70
80
Time

In [ ]: