

Background

Acute Myocardial Infarction (MI) (“heart attack”) is caused by the formation of a clot in one of the coronary arteries that supply blood to the heart muscle.

Acute MI is a major public health problem. Mortality is substantial in the period immediately after the event (and during the years after surviving the initial infarction).

Some patients die before reaching the hospital. Patients seen in hospitals are reported to have an average mortality within 30 days around 6–15%, with improvement over time.

The risk of 30-day mortality strongly depends on various prognostic factors. In younger patients, risks are much lower than in older patients. Other patient demographics are also important (gender, height, weight → these latter two are useful to derive the Body Mass Index*), as well as the presence of risk factors (hypertension, diabetes, smoking, family history) and the history of previous cardiovascular events (previous MI, angina, stroke, bypass surgery).

Relevant presenting characteristics include also the location of the infarction and the extent of ECG abnormalities. Very important is the acute state of the patient as reflected by the Killip score.

The data we use here are a sub-sample coming from a trial, a four-arm randomized study that compared the efficacy of different intravenous thrombolytic regimens for acute myocardial infarction.

Aim of this project is to predict the probability of 30-days mortality.

* Body Mass Index is a calculation using a person's height and weight. The formula is

$BMI = \text{kg}/\text{m}^2$ where kg is a person's weight in kilograms and m^2 is their height in metres squared. A BMI of 25.0 or more is overweight, while the healthy range is 18.5 to 24.9.

Data structure:

variable	labels
day30	EVENT 1: death; 0=alive
Sex	0=Male; 1=Female
Age	age in years
a65	age $\geq 65 = 1$; age $< 65 = 0$
killip	killip class (ordinal)
Sho	Shock: Killip class 3/4 versus 1/2 (0/1)
Dia	Diabetes (0/1)
Hyp	Hypotension: systolic BP < 100 (0/1)
Hrt	Heart rate: pulse > 80 (“tachycardia”, 0/1)
Ant	Anterior infarct location (0/1)
pmi	Previous myocardial infarction (0/1)
Hei	Height in cm
Wei	Weight in kg
Htn	Hypertension history (0/1)
smk	Smoking (1 = never; 2 = ex; 3 = current)
Lip	Lipids: hypercholesterolaemia (0/1)
Pan	Previous angina pectoris (0/1)
Fam	Family history of MI (0/1)
Ste	ST elevation on ECG: number of leads
st4	ST elevation on ECG: > 4 leads (0/1)
ttr	Time to relief of chest pain > 1 h (0/1)

Analysis track:

- 1) Start with an initial data analysis, comparing patients dead *versus* patients alive at 30 days. Insert also a column with the total population descriptive statistics. Include also a new variable, the Body Mass Index.
- 2) Perform univariable regression analyses (crude association measures) of all candidate predictors for your prognostic model. [For the continuous variables is the linear effect reasonable?].
- 3) Build a multivariable prognostic model: as you know, various selection procedures of the candidate predictors could be performed (try different approaches and compare them... we don't have here the expert's opinion, therefore the final choice will be based only on statistical evaluations with the aim of maximising predictive accuracy). Pay attention to the multicollinearity between predictors (and to use the same variables but recoded in different ways).
- 4) Evaluate model's performance in terms of discrimination. Evaluate model's performance in terms of calibration.
- 5) Represent to the clinical audience the results from the estimated model.