

To illustrate national incidence and mortality trend analyses as proposed in the paper, two R datasets are provided (dataI.modified and dataM respectively) and corresponding programs. The data stem from female lung cancer data in France, but they were slightly perturbed for incidence, to respect privacy regulations.

■ **Mortality data from 1975 to 2018, by annual age and year (dataM)**

- National level
- Numbers of death (variable named DEATH) are available up to 2015
- DEATH is NA for years 2016 to 2018 but the person-years (PY) are available and will be used to project mortality up to 2018;

■ **Incidence data from 1985 to 2018, by district, annual age and year (dataI.modified)**

- Districts with a cancer registry ^a
- Numbers of incident cases (variable named K) are available up to 2015
- K is NA for years that are outside the registration period of each registry
- The dataset also includes lines with a district named "fr" which corresponding to whole France; for these lines, K is NA but the PY are available and will be used to estimate national incidence

^a Lille-Metropole registry was excluded from dataI.modified, for simplicity (it is an infra-district area so that historical PY are not fully available)

■ **Programs**

1_Model1_Mortality.R: Program to model national mortality and create the predicted dataset

2_Model2_Incidence.R: Program to model national incidence from local registry data and create the predicted dataset

3_Visualization.R: Program to visualize (here on mortality prediction) the prediction obtained, by age, year or cohort, as presented in the paper