

Original C1 Dataset
(p= 85,722;
p= 163,146)

Original Prosecutors Office
(p= 74,833;
Causes= 494,929;
Rel.=568,290;
RUC_Vic_Imp=541,595;
n= 568,297)

Data wrangling and normalization process

- Remove duplicated entries
- Overlapping treatments of patients
- Intermediate treatment events (continuous referrals)

Data wrangling and normalization process

- Filter crimes committed after study follow-up period(p= 1,206; RUCs= 1,308;n= 1,413)
- Remove duplicated entries(p= 7; RUCs= 7;n= 7)
- Correct dates (birth, comission of crime, end of judicial proceedings), missing nationality and sex
- Erase entries with missing values in comission of crime, end of judicial proceedings(p= 13; RUCs= 57;n= 61)
- Erase entries with values in comission of crime when minor to 14 years old after imputation(p= 557; RUCs= 895;n= 568)
- Filter crimes committed before study follow-up(p= 17,506; RUCs= 22,776;n= 28,371)
- Filter records with cause of end of the proceedings= administrative annulment(p= 1,053; RUCs= 1,163;n= 1,503)
- Filter records with cause of end of the proceedings= grouped to another case(p= 15,739; RUCs= 24,861;n= 28,024)

C1 Dataset
(p= 85,048;
n= 109,756)

O.P. Dataset
(p= 74,535;
n= 508,236)

27
56

- Discard observations coded as victims rather than ofenders (p= 56,301; RUCs= 155,538;n= 171,203)
- Discard observations depending on the values of the end of the proceedings among offenders (p= 50,013; RUCs= 148,520;n= 167,721)
- Discard TOP applications with duplicated dates and user (p= 2,693; n= 3,407)
- Get the first treatment of each user (p= 17,764; n= 24,708)

Merge records where discharge age is greater than age of offenses

Offenses previous
to the first admission
(p= 85,048; RUCs= 152,180;
n= 209,646)