

Original C1 Dataset
(n = 163,146;
patients: 85,722)

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

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
- Remove duplicated entries
- Correct dates (birth, comission of crime, end of judicial proceedings)
- Define cases that acted as victims & imputed in a cause

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

O.P. Dataset
(n= 565,601;
individuals= 74,745)

<U+22C9>

- Long-to-wide relationships/crimes, end of judicial proceedings, penalty
- Group crimes into violent, drug-related, etc.
- Check overlapping events (e.g., incarcerated while completing a residential treatment)

Merged database