# En systematisk patientflödeskartläggning med hjälp av rådata från TakeCare

Danilo Romero

eg psykolog, Beroendecentrum Stockholm

Doktorand, Forskargruppen för digital psykiatri, Centrum för psykiatriforskning

Hälsodata i forskning - Lunchseminarium 2024-05-15 Clinicum

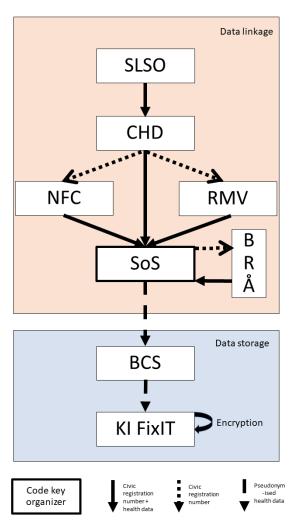
# Agenda

- Kort om studien
- Dataspecifikation
- Förbehandling av rådata ett nedslag i ICD-koder

#### Kort om studien

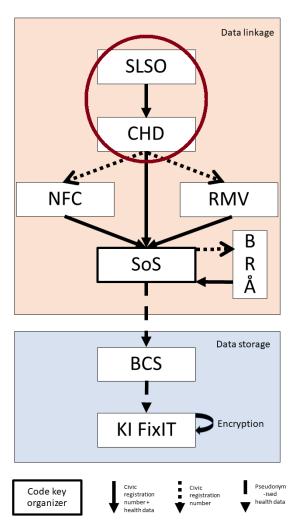
- **Syfte:** Kvantifiera avhopp i vårdkedjan, stratifierat på åldersgrupp och **problemsubstans**
- *n* = 9,771 personer som besökt BAS mellan 2018-02-28 och 2019-06-30

### Datakällor



Abbrevations: BCS = Stockholm Centre for Dependency Disorders; BRÅ = National Council for Crime Prevention; CHD = Centre for Health Data, Stockholm Regional Council, XI = Karolinska Institutet; NFC = National Forensic Centre; SoS = National Board of Health and Welfare; SLSO = Stockholm Health Care Services; RMV = National Board of Forensic Medicine.

## Datakällor



Abbrevations: BCS = Stockholm Centre for Dependency Disorders; BRÅ = National Council for Crime Prevention; CHD = Centre for Health Data, Stockholm Regional Council; K1 = Karolinska Institutet; NFC = National Forensic Stockholm Health Grave Services; RMV = National Board of Forensic Medicine.

	TC (SLSO)	VAL (CHD)
Filtervillkor	Samtliga besökare på BAS	Samtliga vårdhändelser med F- diagnos
Tidsfönster	16M-period, 2018 till 2019	10 år bakåt och 6M framåt
Variabeluppsättning	Tidsstämpel, Diagnos-	och åtgärdskod
Registerspecifikt utdrag	Ankomstsätt (Söker själv, LOB), Alkoholutandingsprov	Besöksstatus (t.ex. "5. Utebliven")

# Ytterligare uppgifter från TC

Studiepopulationens:

- vårdkedjetider
- remisser och egenremisser

#### **Dataformat**

- långt format
- en rad = en vårdhändelse
- Fördelar:
  - Undviker informationsförlust
  - Större flexibilitet i analysen
  - Möjlighet till mer djupgående analys av vårdkonsumtion

pseudo_id	indexdatum	besoksdatum	icd_raw
Tolvan Tolvansson	2019-01-07	2019-06-13	F100,F200,F300
Tolvan Tolvansson	2019-01-07	2019-01-29	F101
Tolvan Tolvansson	2019-01-07	2019-01-07	F100, F200
John Doe	2019-03-04	2019-03-12	F150
John Doe	2019-03-04	2019-03-04	F190

# Utveckling av studiens dataspecifikation

- VAL-handboken
- SLSO
  - Besök på BAS
  - Mejlkorrespondens med SLSO (Sead)
  - Kvalitetssäkringsarbete

- Under kvalitetssäkringsarbete på kliniken upptäcktes diskrepans mellan journal och kassaregister
- Andel uteblivande från vårdkedjetider fördubblades när journalanteckning istället användes som datakälla

# Förbehandling av rådata

#### Ett exempel utifrån ICD-diagnos

- Uppgift: bearbeta ICD-koder till en lämplig detaljnivå.
- I detta fall vill vi kunna avgöra problemsubstans utifrån ICDdiagnos vid indexbesök (från fyr- till treställig diagnos).

Kod	Kodbeskrivning	Funktion
F10	Psykiska störningar [] orsakade av alkohol	Sjukdomskategori
F10.0	Akut intoxikation av alkohol	Subkategori (specifik kod)

```
1 library(tidyverse)
2 data |>
3   filter(indexdatum==besoksdatum) |>
4   mutate(first_icd = str_extract(icd_raw, pattern = "^[^,]+")) |>
5   mutate(icd_tre = str_extract(first_icd, pattern = "^.{3}"))
6
7
8
9
```

pseudo_id	indexdatum	besoksdatum	icd_raw
Tolvan Tolvansson	2019-01-07	2019-06-13	F100,F200,F300
Tolvan Tolvansson	2019-01-07	2019-01-29	F101
Tolvan Tolvansson	2019-01-07	2019-01-07	F100, F200
John Doe	2019-03-04	2019-03-12	F150
John Doe	2019-03-04	2019-03-04	F190

```
1 library(tidyverse)
2 data |>
3    filter(indexdatum==besoksdatum) |>
4    mutate(first_icd = str_extract(icd_raw, pattern = "^[^,]+")) |>
5    mutate(icd_tre = str_extract(first_icd, pattern = "^.{3}"))
6
7
8
9
```

pseudo_id	indexdatum	besoksdatum	icd_raw
Tolvan Tolvansson	2019-01-07	2019-01-07	F100, F200
John Doe	2019-03-04	2019-03-04	F190

```
1 library(tidyverse)
2 data |>
3   filter(indexdatum==besoksdatum) |>
4   mutate(first_icd = str_extract(icd_raw, pattern = "^[^,]+")) |>
5   mutate(icd_tre = str_extract(first_icd, pattern = "^.{3}"))
6
7
8
9
```

pseudo_id	indexdatum	besoksdatum	icd_raw	first_icd
Tolvan Tolvansson	2019-01-07	2019-01-07	F100, F200	F100
John Doe	2019-03-04	2019-03-04	F190	F190

#### stringr::str\_extract

extraherar den **första** instansen av ett mönster specificerat av ett reguljärt uttryck (regex) från en sträng. Om extrahera alla instanser, använd **str\_extract\_all**, så returneras en lista med strängar.

#### regex

Ett universellt verktyg för att söka (och extrahera) ett specificerat textmönster.

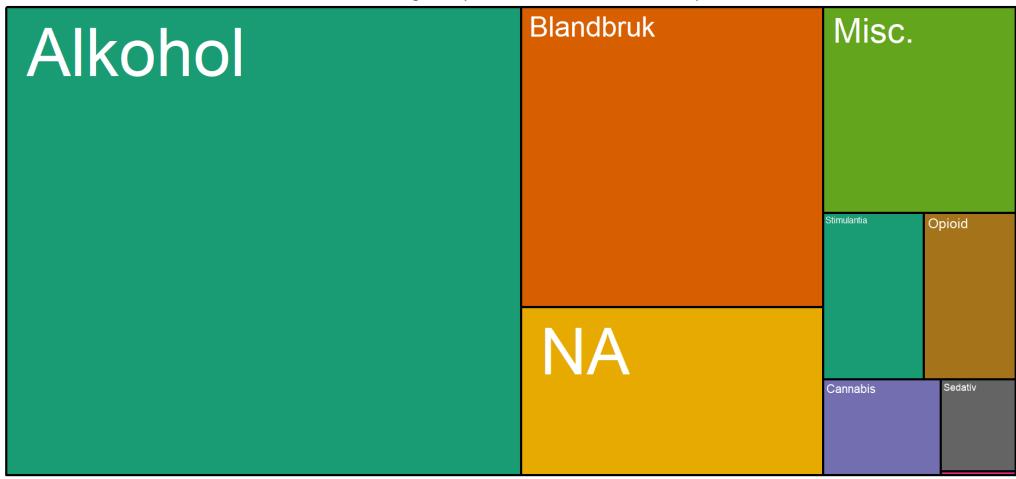
- Speech and language processing, Kap. 2: https://web.stanford.edu/~jurafsky/slp3/2.pdf
- Generativ AI (GPT-4)

```
1 library(tidyverse)
2 data |>
3   filter(indexdatum==besoksdatum) |>
4   mutate(first_icd = str_extract(icd_raw, pattern = "^[^,]+")) |>
5   mutate(icd_tre = str_extract(first_icd, pattern = "^.{3}"))
6
7
8
9
```

pseudo_id	indexdatum	besoksdatum	icd_raw	first_icd	icd_tre
Tolvan Tolvansson	2019-01-07	2019-01-07	F100, F200	F100	F10
John Doe	2019-03-04	2019-03-04	F190	F190	F19

# Från fyr- till treställig ICD-kod

Primär diagnos (vid indexbesök; n = 9,771)



# Sammanfattning/tips

- Besök vårdenheten under utveckling av dataspecifikation
- Sök samarbetsmöjligheter med kvalitetsäkringsarbeten
- Håll utkik efter diskrepans mellan journal och kassaregister

# Tack för er uppmärksamhet!

För frågor eller vidare diskussion, kontakta mig på

☑ danilo.romero@ki.se

Presentationsbilder och tillhörande R-kod finns på

@DaniloCarlosRomero