## 1 Preliminary Hypothesis on Relevant Data

One-time historical data pull from the claims data

Claims data for all patients for **10 years** (ideal), 5 years (decent), 3 years (minimum necessary)

* Unique patient IdentifierWe would need to **link claims to each individual** patient to track their medical history and opioid use trajectory over time.

Minimum desired data fields

* Claims data
  + Independent (predictor) variables
    - prescription history,
    - Pain management services,
    - Co-morbidities
    - Chronic conditions
    - Specialist visits,
    - Injuries,
    - PCP visits,
  + Dependent variables (prediction target)
    - Surgeries
    - Ambulance calls
    - Overdose events
    - ER Visits,
    - Hospitalizations
* Socio-demographic; e.g.,
  + Age
  + Race
  + Gender
  + Smoking, drinking, and drug history,
  + History of substance abuse,
  + Family history substance abuse,
  + History of physical/sexual abuse