

Using indication embeddings to represent patient health for drug safety studies

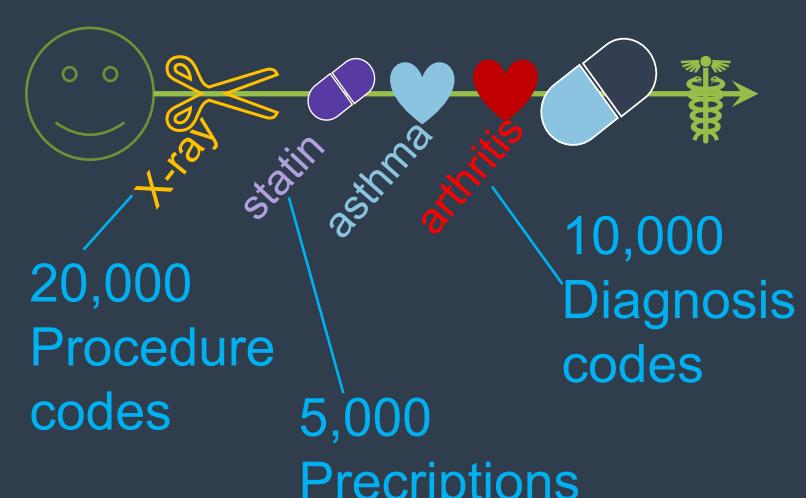
Rachel Melamed
Biomedical Data Science @University of Chicago & Biology @ UMass Lowell
melamed@uchicago.edu | @RDMelamed

Drug safety

Does taking this drug change your risk of some health outcome?



Health data



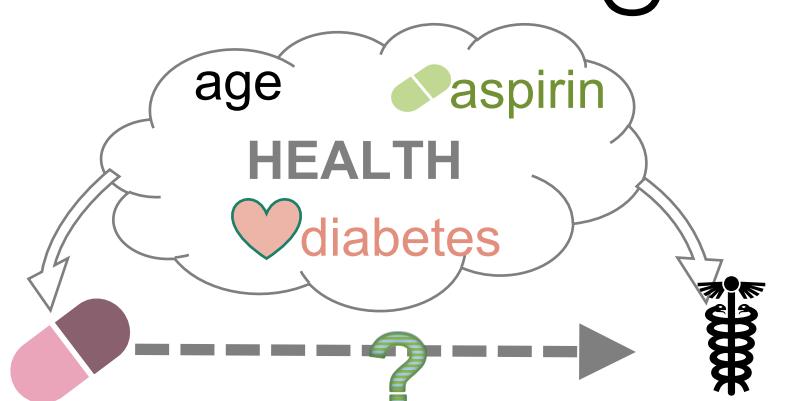
High-throughput cohort studies

Currently, cohort study design relies on domain experts:

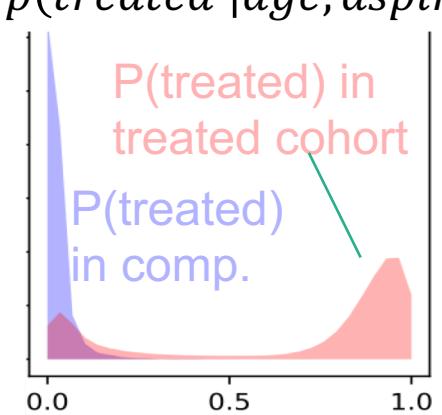
Expert Task 1 Find suitable comparator drug

Expert Task 2 Design matching—identify confounders

The challenge: confounding



Evaluate with propensity score:
 $p(\text{treated} \mid \text{age}, \text{aspirin}, \dots)$

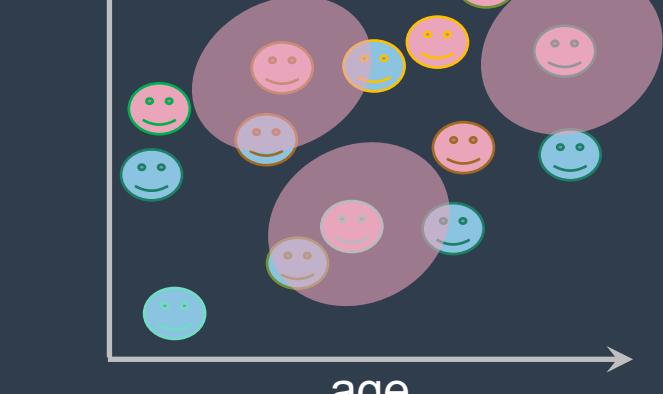


The solution: matching

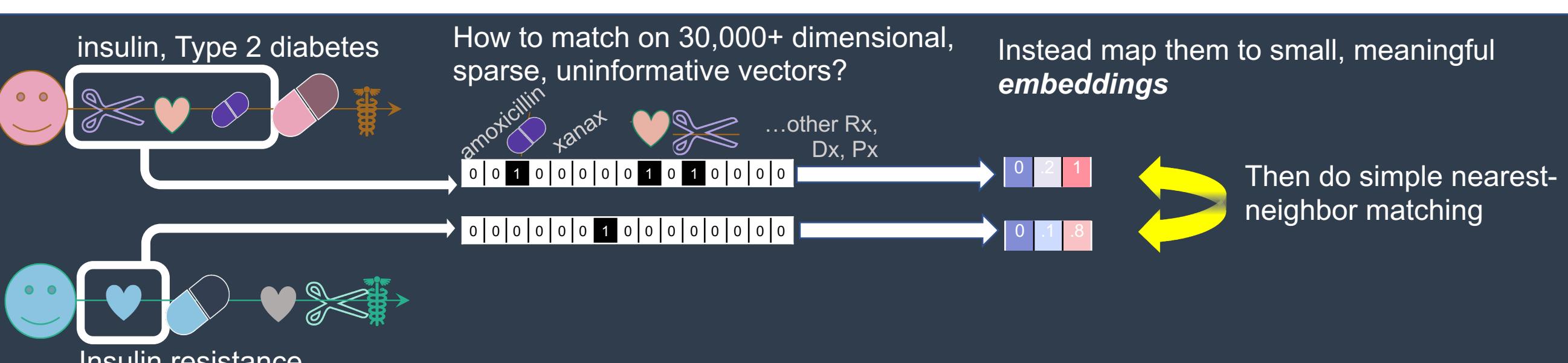
Propensity score match



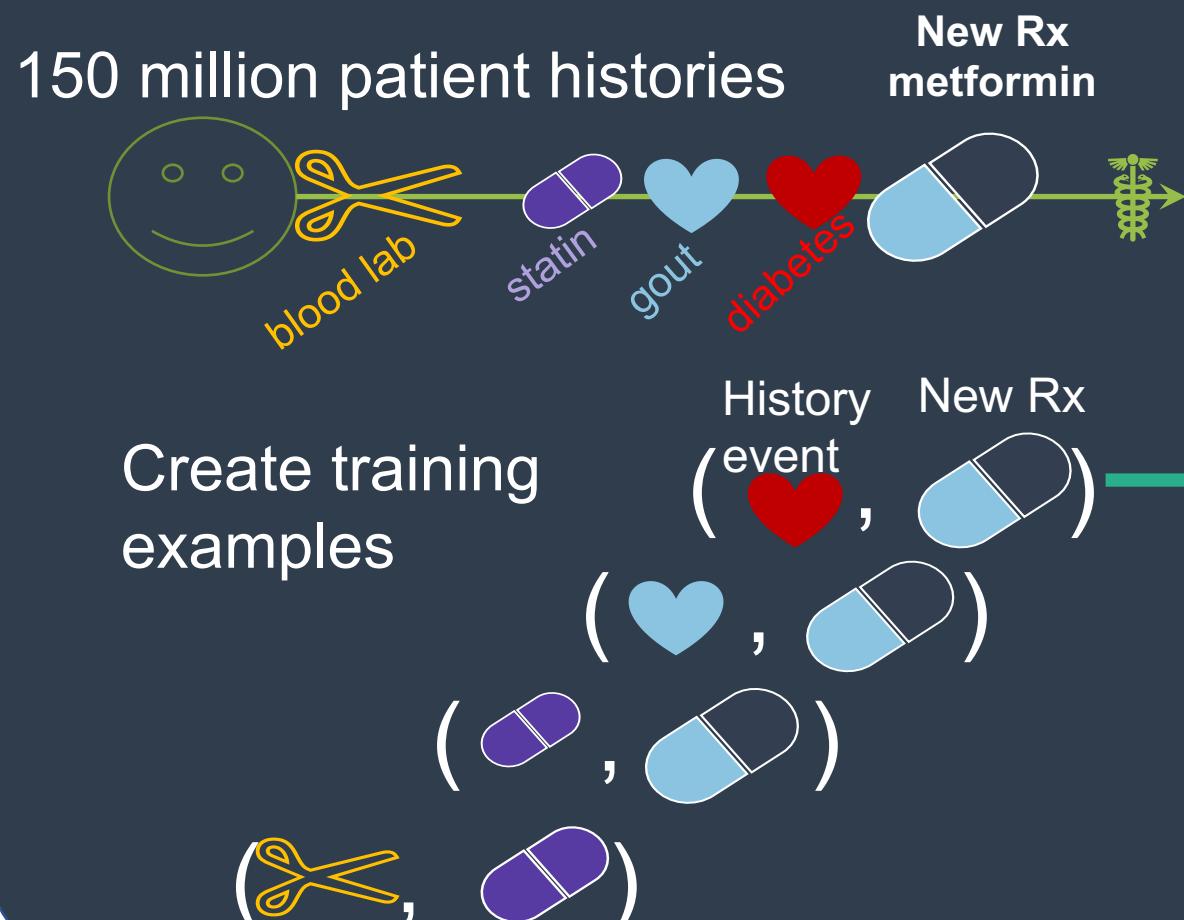
Match on confounders



Can we match without expert design?

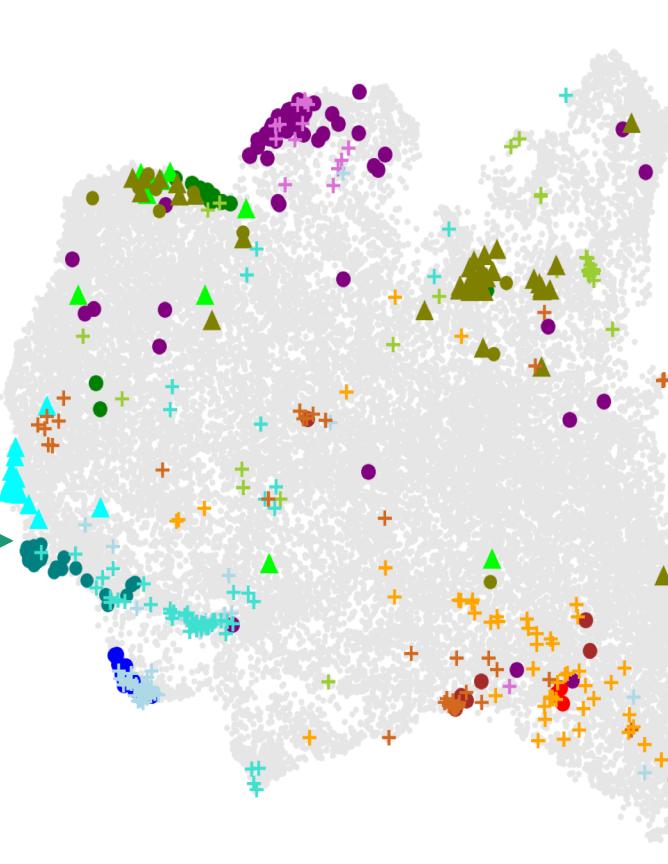


Creating indication embeddings



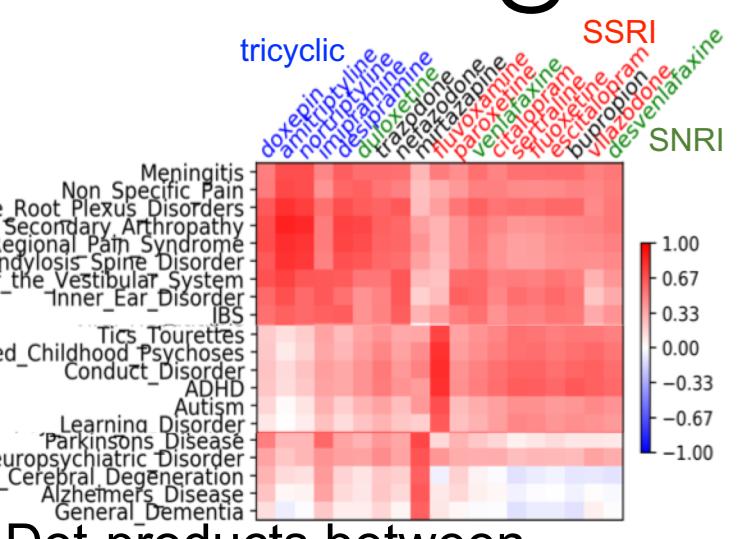
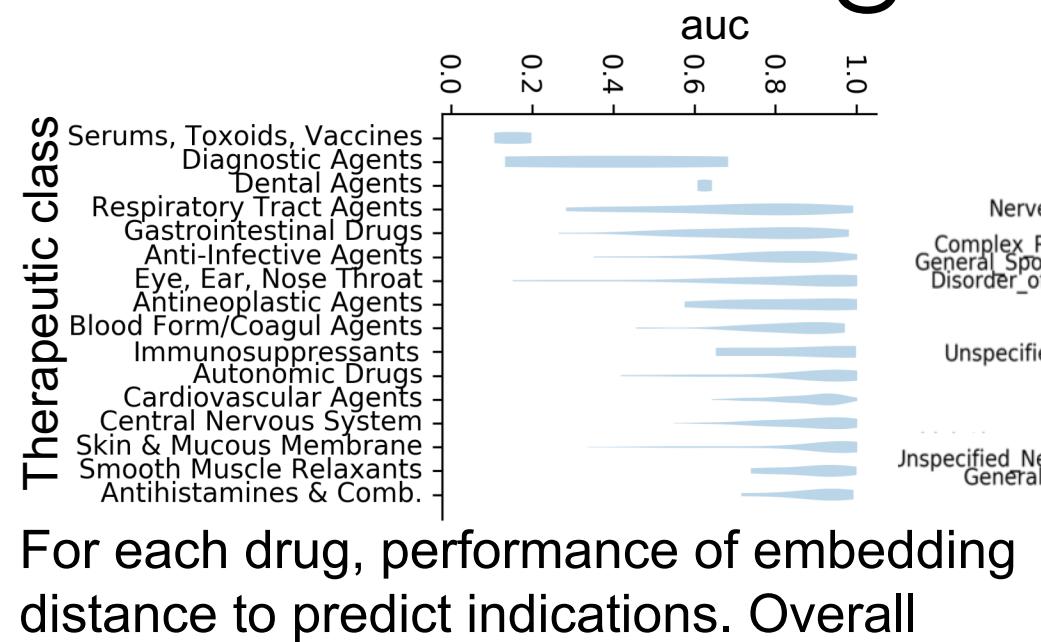
Embeddings relate codes to health needs

Drug embedding = drugs given in most similar health contexts
Indication embedding = health context for prescription of a new drug

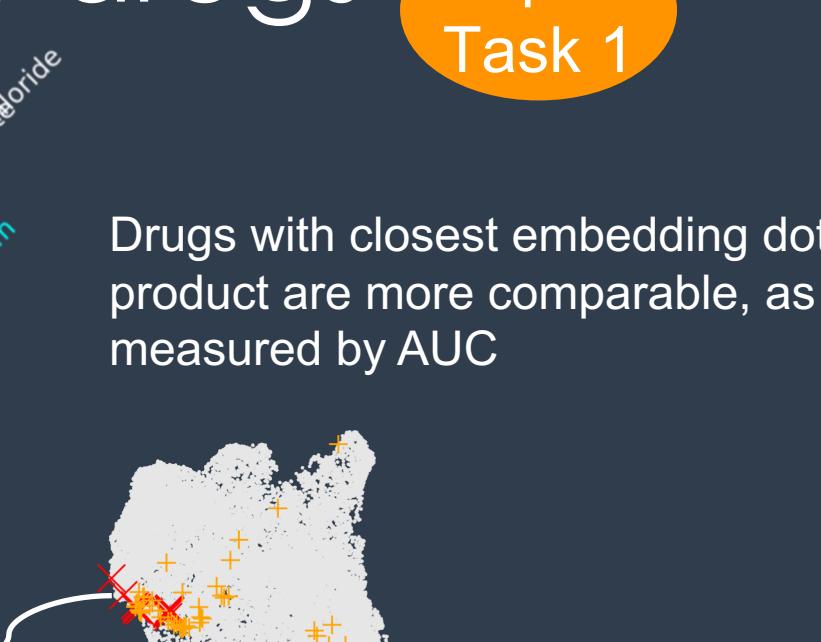
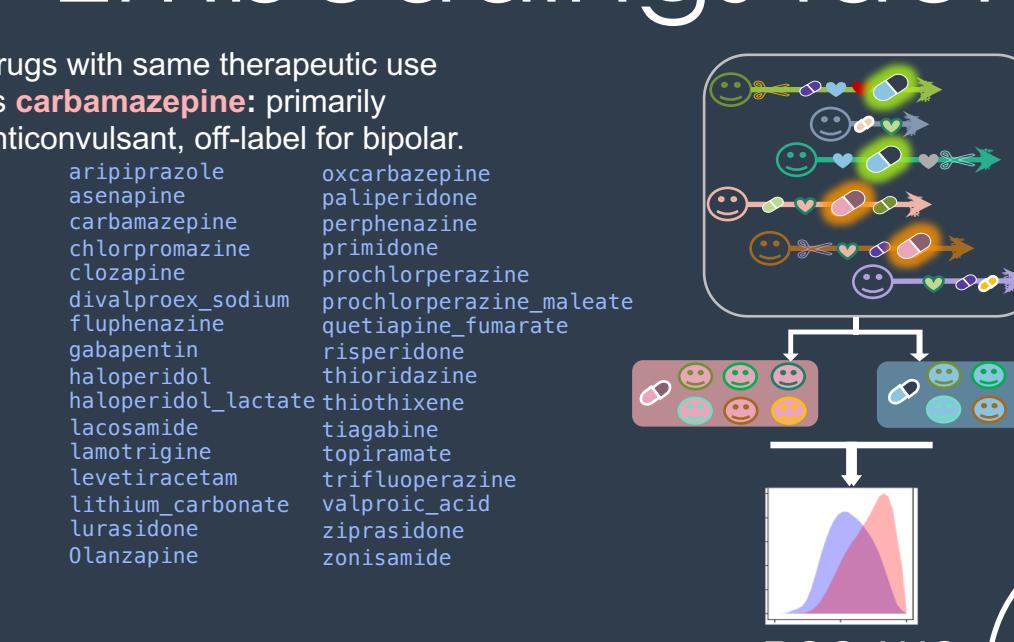


- Schizophrenia_Related_Psychosis
- psychosis related medications
- Type_II_Diabetes_Mellitus
- diabetes related medications
- Breast_Cancer
- Breast cancer medications
- mastectomy
- Ovarian_Cancer
- hysterectomy
- Myocardial_Infarction
- MI medications
- Coronary bypass graft
- Psoriasis_Related_Disorders
- psoriasis medications
- Allergy
- allergy medications

Evaluating embeddings



Embeddings identify comparator drugs

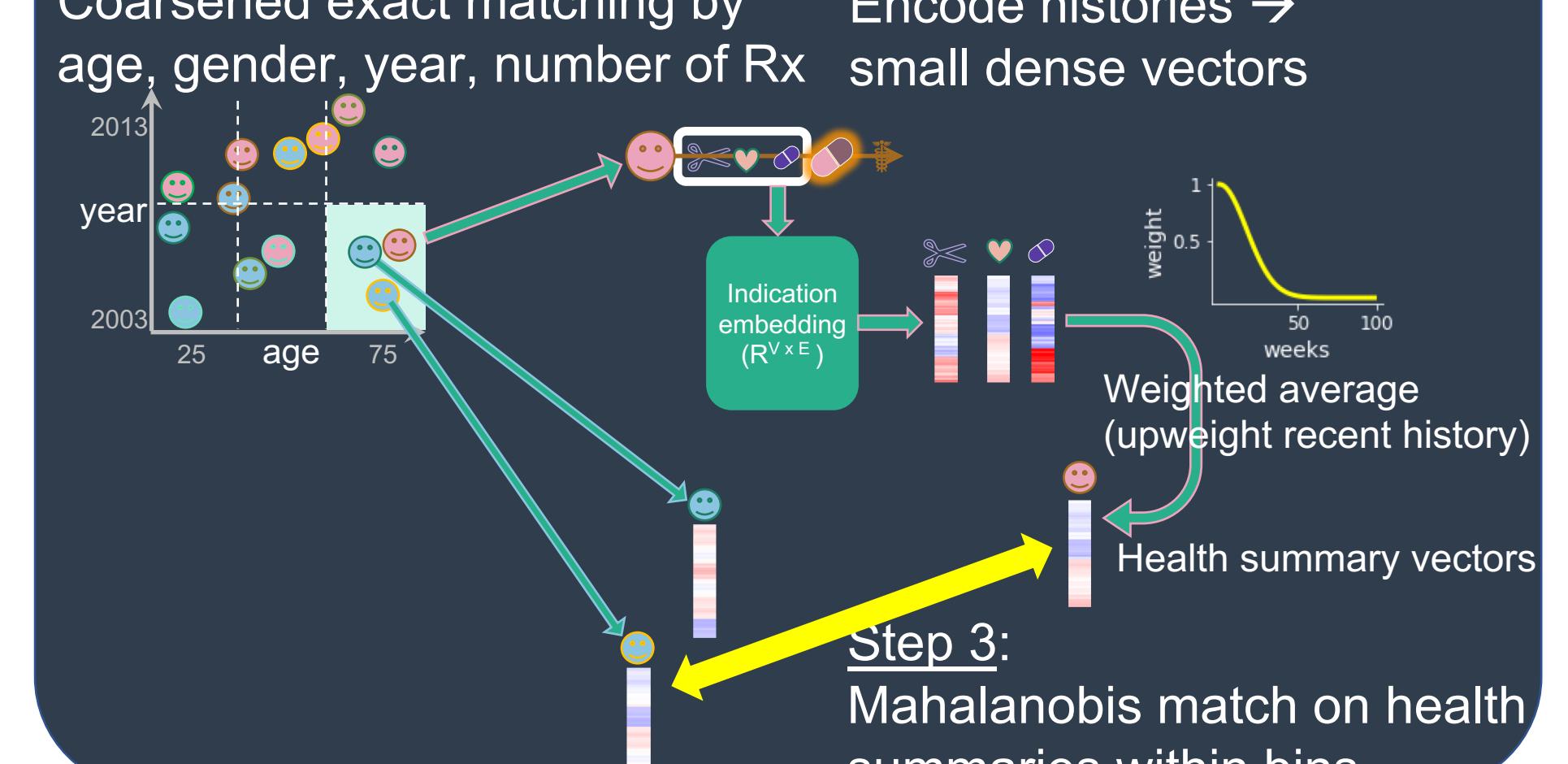


Match with embeddings

Expert Task 2

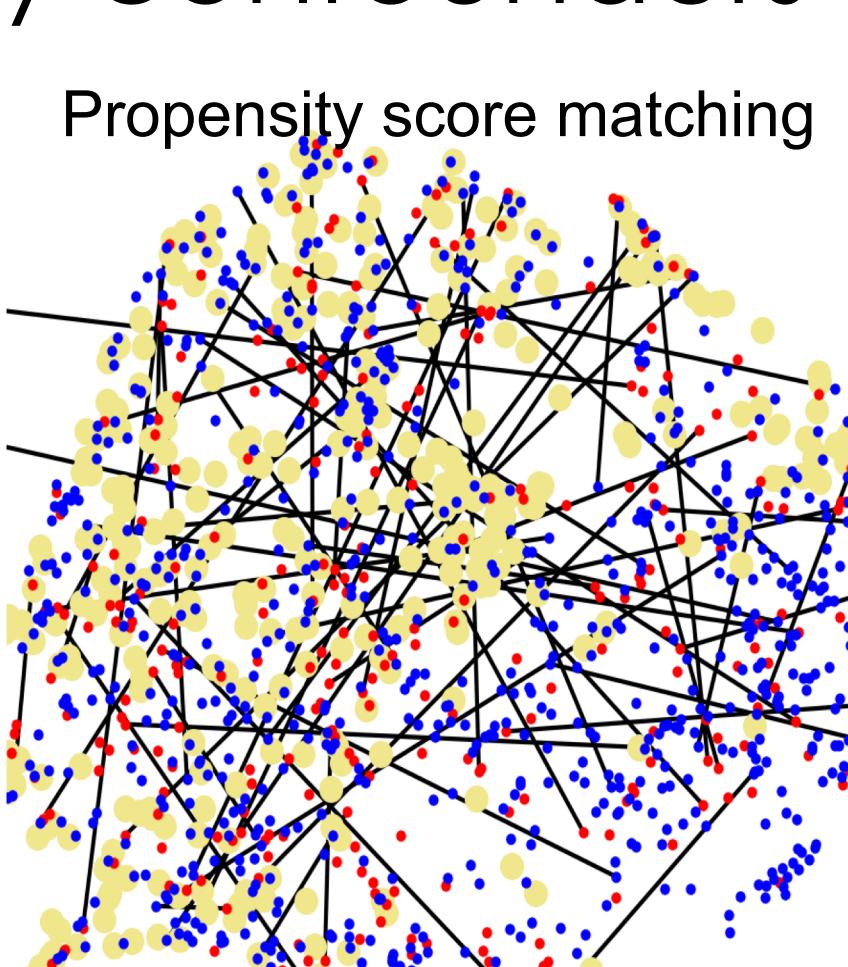
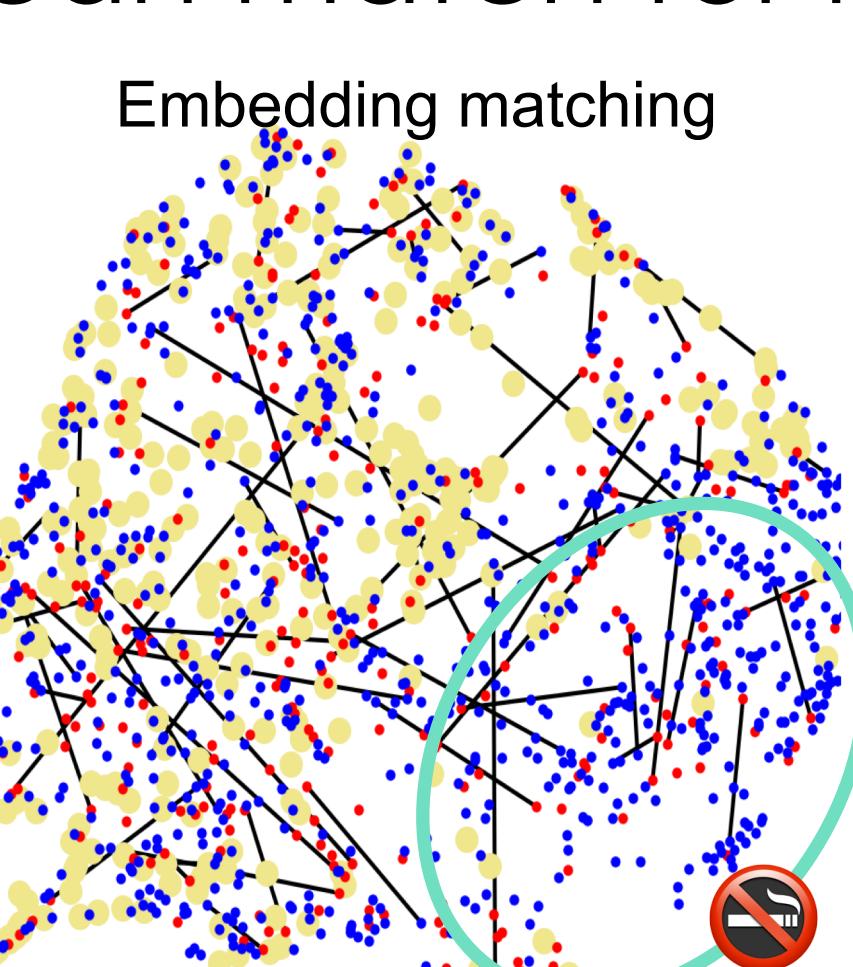
Step 1: Coarsened exact matching by age, gender, year, number of Rx

Step 2: Encode histories → small dense vectors



Embedding can match for key confounders

Matching people on bupropion to trazodone is complicated by alternate indication of bupropion for smoking cessation. Each point is one person on bupropion, trazodone, or varenicline.



Embedding better matches nonsmokers to nonsmokers