Real-World Evidence of Association between Autoimmune Diseases

Shah Lab

Observational Health Data Science
Weekly Meeting
Hypothesis

Association amongst diseases seen in medical claims data can guide molecular understanding of diseases
Summary of Analysis

Build Clinical Profiles: each disease, each State across the United States

Nation-Wide Medical Insurance Claims: over 150 million patients

Patient Feature Matrix

Match
Age|Gender
Record Length

Similar patient without disease

Disease-1
disease-2
disease-n

Patient with disease

Autoimmune Diseases
Other disease types

AD, AD2, AD3
p1 1 1 1 1 1
p2 1 1 1 1 1
p3 1 1 1 1 1
p4 1 1 1 1 1

Disease of Interest as outcome

Nation-Wide Medical Insurance Claims: over 150 million patients

Meta Analysis: each disease across the United States account for heterogeneity and improve reproducibility

Systems Immunology: understanding shared genetic signature for informed medical decision making

- BioPlex: analysis of biophysical interactions at genome scale
- ImmunoState: analysis of blood gene expression across GEO
- UKBiobank: genetic correlation and polygenic

Not studied: Connected if there is no association inferred from PubMed

Understudied: Connected if there is less than 50% chance of association inferred from PubMed

Association: between autoimmune diseases representing medical practice

Nigam Shah Lab
Rohit Vashisht: rohitv@stanford.edu
Nigam Shah: nigam@stanford.edu
Clinical profile of a disease D is a vector of length $i$, where each feature is another disease $D_i$ and the values are the strength of association between D and $D_i$ obtained from a matched cohort of patient with and without disease D.
Meta analysis of associations among diseases
Associations seen in practice, but not in literature

140 Disease Pairs with strong real-world evidence.

52 Understudied Disease Pairs

9 Novel Disease Pairs

a. Evidence from Nation-Wide Medical Claims
   Connected if evidence is strong
   Medically Connected

b. Evidence from Literature
   Connected if under reported in the literature
   Under-Explored

c. Evidence from Literature
   Connected if not reported in the literature
   Not Explored

d. Systems Immunology Analysis
   - Shared Genetic Signature
     Biophysical interactions
   - Blood Gene Signatures
     MultiCohort blood gene expression analysis
   - Genetic correlation & Polygenicity
     UkBiobank analysis of genome wide associations
Preliminary results

• We identified 140 pairs of autoimmune diseases that have a strong relation based on nationwide medical claims of over 150 million patients.

• Of the 140 pairs, 52 had very poor evidence of being associated in the literature and 9 pairs have never been studied together based on PubMed.

• For each of the 61 pairs we performed a systematic genome-wide analysis using the UK BioBank and gene expression data from GEO to uncover genetic correlation, polygenicity and co expressed genes underlying a disease pair to understand shared genetic signatures.
Related Work

• A Non-degenerate code of deleterious variants in Mendelian loci contribute to complex disease risk Blair et al. Cell 2013 Sep26;155(1):70-80

Thank You

let's see a demo before we get to questions

http://autoimmunedb.stanford.edu