

Disease comorbidities networks

Xavier Duran
GCAT Genomes for Life

September 20th, 2016

Population-based disease associations

- ▶ Diseases have multiple causes and are related through several dimensions
- ▶ Diseases can be linked based on comorbidities associations
- ▶ Calculate pairwise comorbidity correlations

Distance between two diseases: Relative Risk

$$RR_{ij} = \frac{C_{ij}N}{P_i P_j}$$

- ▶ C_{ij} is the number of the patients affected by both diseases, N is the total number of patients and P_i and P_j are the prevalences of diseases i and j
- ▶ $RR_{ij} > 1$ comorbidity is larger than expected by chance
- ▶ RR **overestimates** relationships involving **rare diseases** and **underestimates** comorbidity between **highly prevalent** ones
- ▶ Most widespread use in the medical literature (Hidalgo et al. (2009))

Distance between two diseases: ϕ -correlation

$$\phi_{ij} = \frac{C_{ij}N - P_iP_j}{\sqrt{P_iP_j(N - P_i)(N - P_j)}}$$

- ▶ Pearson's correlation for binary variables
- ▶ **Accurately discriminates** comorbidities between pairs of diseases of **similar prevalence**
- ▶ **Underestimates** the comorbidity between rare and common diseases

GCAT comorbidities

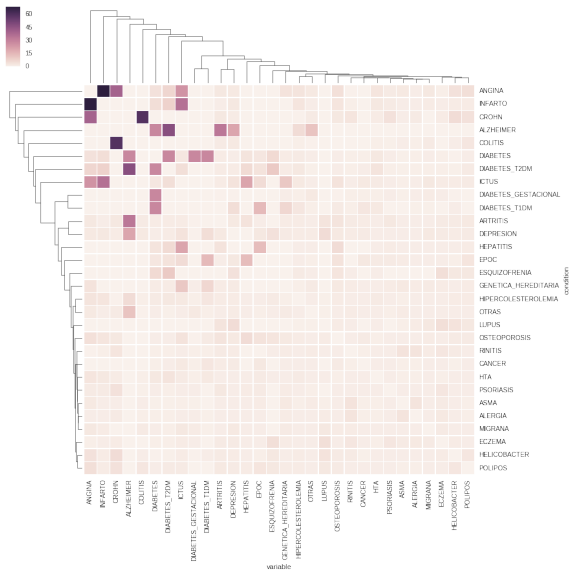


Figure 1: Comorbidity matrix

GCAT comorbidities

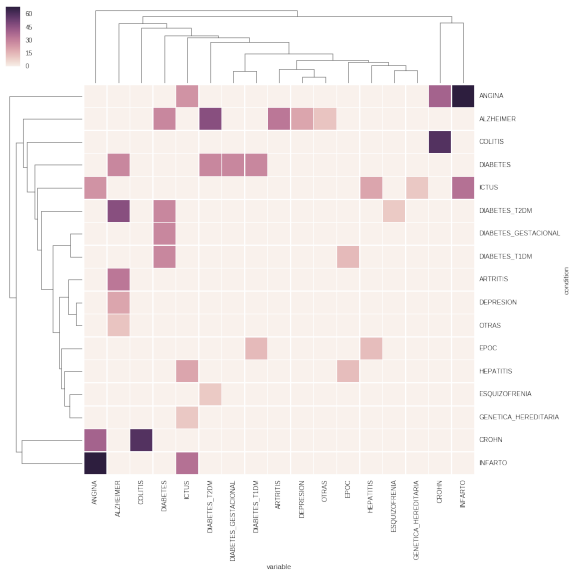


Figure 2: Comorbidity matrix (RR > 10)

GCAT comorbidities

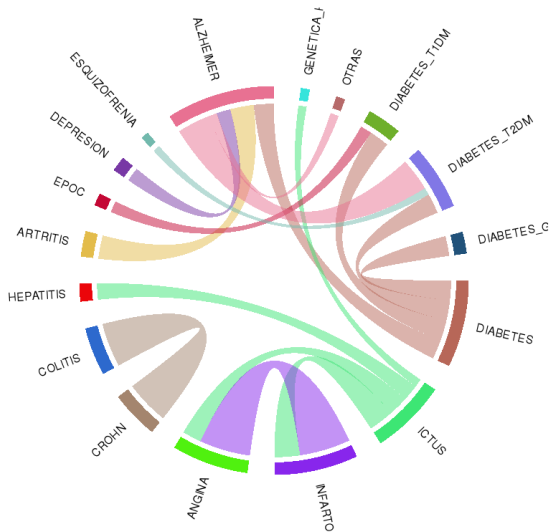


Figure 3: Circle plot ($RR > 10$)

Human disease network

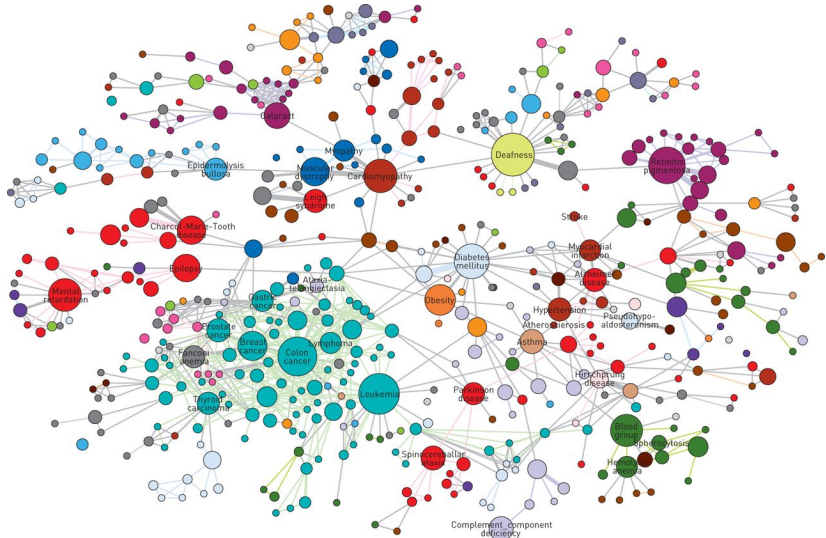


Figure 4: Human Disease Network (Goh et al. (2007))

Temporal comorbidity analysis

Temporal disease trajectories condensed from population-wide registry data covering 6.2 million patients (A. B. Jensen et al. (2014))

- ▶ EHR as medical event sequences
- ▶ Temporal disease progression
- ▶ Comprehensive set of disease trajectories
- ▶ $D1 \rightarrow D2$ pairs of diagnosis
- ▶ Relative Risk (RR) > 1
- ▶ Directional pairs were then combined into longer trajectories consisting of three or more diagnoses
- ▶ 1171 significant trajectories of four diagnosis that forms 15 major clusters

Temporal comorbidity analysis

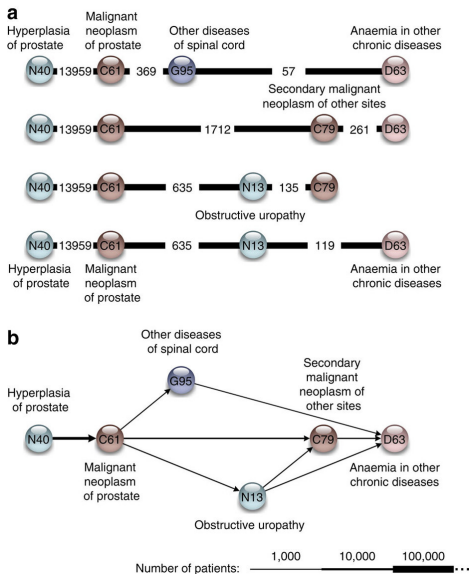


Figure 5: Disease trajectories and trajectory-cluster for prostate cancer

Temporal comorbidity analysis

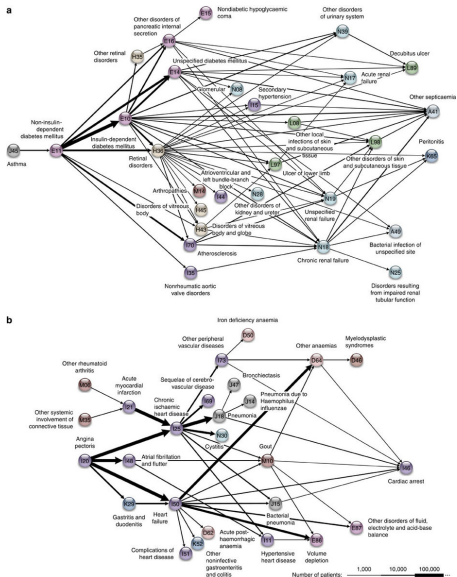


Figure 6: Diabetes and cardiovascular disease trajectory clusters

References

- Goh, Kwang-il, Michael E Cusick, David Valle, Barton Childs, and Marc Vidal. 2007. "The human disease network." doi:10.1073/pnas.0701361104.
- Hidalgo, César a., Nicholas Blumm, Albert-László Barabási, and Nicholas a. Christakis. 2009. "A Dynamic Network Approach for the Study of Human Phenotypes." *PLoS Computational Biology* 5 (4): e1000353. doi:10.1371/journal.pcbi.1000353.
- Jensen, Anders Boeck, Pope L Moseley, Tudor I Oprea, Sabrina Gade Ellesøe, Robert Eriksson, Henriette Schmock, Peter Bjødstrup Jensen, Lars Juhl Jensen, and Søren Brunak. 2014. "Temporal disease trajectories condensed from population-wide registry data covering 6.2 million patients." *Nature Communications* 5 (May): 4022. doi:10.1038/ncomms5022.