Predicting patient sensitivity to pharmacogenomic drugs, using EHR data

Adrien Coulet, LORIA & Inria Nancy-Grand Est
Nigam H. Shah, BMIR, Stanford University
Michel Dumontier, BMIR Stanford University & Snowflake Associate Team
The Snowflake Inria associate team

• Shah lab@Stanford
  – Analysis of EHR data

• Dumontier lab@Stanford
  – Semantic Web applications in life sciences

• Orpailleur@Inria NGE
  – Symbolic data mining & knowledge guided knowledge discovery
Objectives of Snowflake

• Studying *variability in drug response* using EHRs

• And how *symbolic data mining* approaches can help
Identifying co-occurring ADEs in EHRs
Co-occurring ADEs in EHRs

- ADEs are Adverse Drug Events
- Aim: Finding several ADEs that are frequently co-occurring in patients

Support($\{ADE_1, ADE_2\}$) = $\frac{3}{4}$
Why this is hard?

- Manifestation of ADE may differ
- Reporting of ADE may differ
- ADE are not that frequent
Why this is hard?

• Manifestation of ADE may differ
• Reporting of ADE may differ
• ADE are not frequent

=> Generalization may help!

Knowledge engineering  Formal Concept Analysis

(warfarin, bleeding ulcer)
(anticoagulant, bleeding ulcer)
(warfarin, bleeding)
Why this is hard?

• Manifestation of ADE may differ
• Reporting of ADE may differ
• ADE are not frequent

=> Generalization may help!

Knowledge engineering  Formal Concept Analysis

Drug ontology

(drug)

(drug, anticoagulant)

(drug, anticoagulant, bleeding ulcer)

(drug, warfarin)

(drug, anticoagulant, warfarin)

(drug, anticoagulant, bleeding ulcer, warfarin)

(drug, bleeding ulcer)

(drug, bleeding)

(drug, bleeding ulcer, bleeding)

(drug, bleeding ulcer, bleeding, warfarin)
Why this is hard?

• Manifestation of ADE may differ
• Reporting of ADE may differ
• ADE are not frequent

=> Generalization may help!

Knowledge engineering    Formal Concept Analysis

Drug ontology

Drug ontology

Phenotype ontology

Phenotype ontology

(warfarin, bleeding ulcer)  (anticoagulant, bleeding ulcer)
(warfarin, bleeding)
Why this is hard?

- Manifestation of ADE may differ
- Reporting of ADE may differ
- ADE are not frequent

=> Generalization may help!

Knowledge engineering  Formal Concept Analysis

- (warfarin, bleeding ulcer)
- (anticoagulant, bleeding ulcer)
- (warfarin, bleeding)
- (drug, phenotype)
• Co-occurring ADEs project
  – Lead by Gabin Personeni, 3rd year PhD student
  – 3 month stay of Gabin@Stanford
  – Funded by the France-Stanford Center for Indisciplinary Studies & Inria
  – Preliminary results: Phenoday@ISMB 2016
Many perspectives for the Snowflake AT

What?

• More data
  – Experiment with EHRs that include genetic data
  – and lab results
• Implementation of personalized medicine
• Experiment with French EHRs
• Include more sophisticated learning

How?

• ANR PractiKPharma (2016-2019) with French Hospitals
  – Enlarging the consortium
  – involving new students
• Renewal of the Snowflake Inria AT (2017-2019) this year
Thanks to

- Inria@Silicon Valley

- Snowflake AT
  - Michel Dumontier
  - Nicolas Jay
  - Yen Low
  - Paea Lependu
  - Mark Musen
  - Amedeo Napoli
  - Gabin Personeni
  - Nigam Shah
  - Malika Smaïl-Tabbone
  - Yannick Toussaint

- Other collaborators
  - Anita Burgun
  - Marie-Dominique Devignes
  - Clément Jonquet
  - Bastien Rance