AACT Database - An Open Reference for Subjects Safety in Clinical Trials

Balamurugan SD

Gowtham M

Symbiance Pharma Solutions Pvt ltd.

Disclaimer

The views and opinions expressed in the following PowerPoint slides are those of the individual presenter and should not be attributed to Symbiance, its directors, employees, members, or affiliates, or any organization with which the presenter is employed or affiliated.

Overview

- Objective
- Data Dictionary
- Challenges
- Data insights
- Pattern Mining
- Limitation
- Conclusion

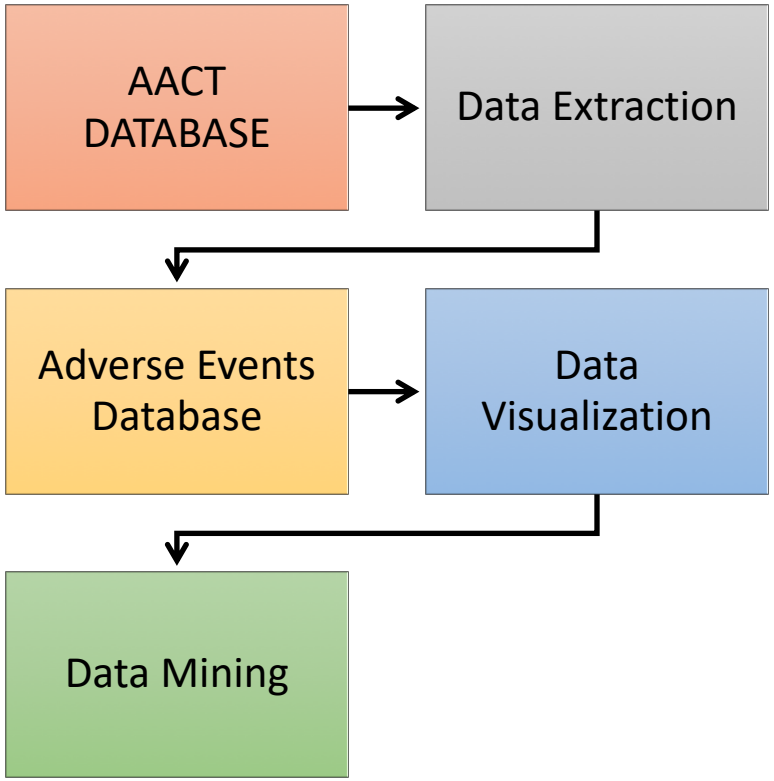
Symbiance
a relationship for success

Objective

- To convert huge volume of Adverse event data available in *clinicaltrials.gov* into valuable data insights.
- Prediction of Pattern from frequently occurring Adverse events

Introduction

- AACT is a publicly available relational database
- Contains protocol and result data elements about every study registered in *clinicaltrials.gov.*



Symbiance
a relationship for success

Application of AACT Database

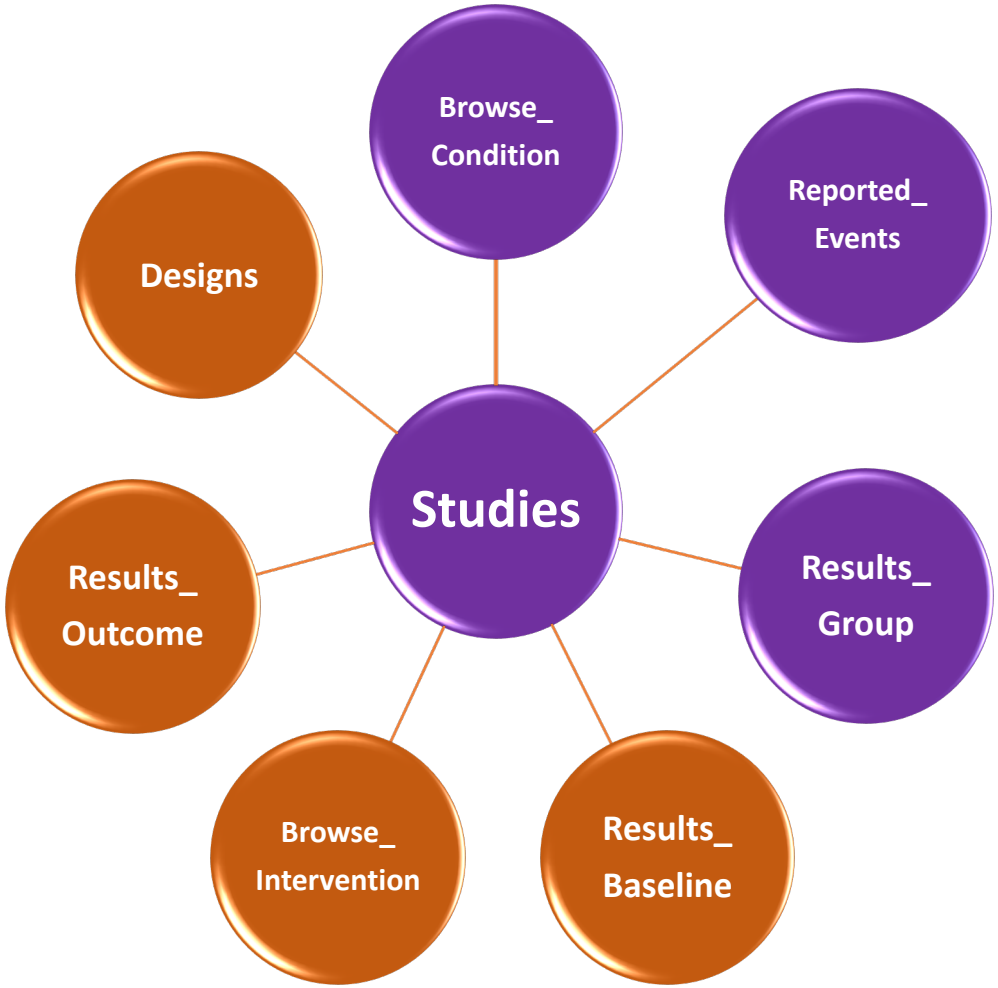
- Observe Clinical Trial Trends
- Systematic Review.
- Meta-analysis
- Compare the efficacy and safety of different types of therapies.

Symbiance
a relationship for success

Data Extraction

- As a cloud-hosted PostgreSQL database
- Static copy of the database can be downloaded.
- AACT can be directly accessed via standard query & analytic tools such as
  - R and R studio
  - SAS

Database Schema



Symbiance
a relationship for success

Datasets

Studies

nct_id	enrollment	phase	number_of_arms	overall_status	completion_dat
NCT00003869	186	Phase 3	2	Completed	2008-05-31
NCT00001575	87	Phase 1/Phase 2	1	Completed	2013-11-30

Condition_Browse

id	nct_id	mesh_term	downcase_mesh_term
2797321	NCT00003869	Lung Neoplasms	lung neoplasms
2809649	NCT00001575	Lymphoma	lymphoma

Reported_Events

nct_id	ctgov_gro	default_vocab	organ_system	adverse_event_term	event_type	subjects_affected	subjects_at_risk
NCT00001575	E1	CTCAE v3	Blood and lymphatic system disorders	BLOOD/BONE MARROW:: Leukocyte	other	32	64
NCT00001575	E1	CTCAE v3	Blood and lymphatic system disorders	BLOOD/BONE MARROW:: Platelets	other	37	64
NCT00003869	E1	MedDRA 5	Gastrointestinal disorders	Nausea	other	55	90
NCT00003869	E2	MedDRA 5	Blood and lymphatic system disorders	Anemia	other	51	92

AACT Results Domain

AACT provides four general categories of result information:

- Participant Flow (Milestones & Drop/Withdrawals)
- Baselines
- Outcomes
- Reported Events

Result_Groups table

NCT_ID	RESULT_TYPE	CTGOV_GROUP _CODE	GROUP TITLE	EXPLANATION
NCT001	Baseline	B1	Experimental Group	All Baseline_Measures associated with this study's experimental group link to this row.
NCT001	Baseline	B2	Control Group	All Baseline_Measures associated with this study's control group link to this row.
NCT001	Outcome	O2	Experimental Group	All Outcome_Measures associated with this study's experimental group link to this row.
NCT001	Outcome	O1	Control Group	All Outcome_Measures associated with this study's control group link to this row.
NCT001	Reported Event	E1	Experimental Group	All Reported_Events associated with this study's experimental group link to this row.
NCT001	Reported Event	E2	Control Group	All Reported_Events associated with this study's control group link to this row.

Symbiance
a relationship for success

Data Dictionary

AACT is composed of 41 tables

Name	Row Count	DB section	Rows per Study
Studies	284,374	Protocol & Results	one
Browse_Conditions	505,111	Protocol	many
Browse_Interventions	295,994	Protocol	many
Reported_Events	4,057,877	Results	many

Symbiance
a relationship for success

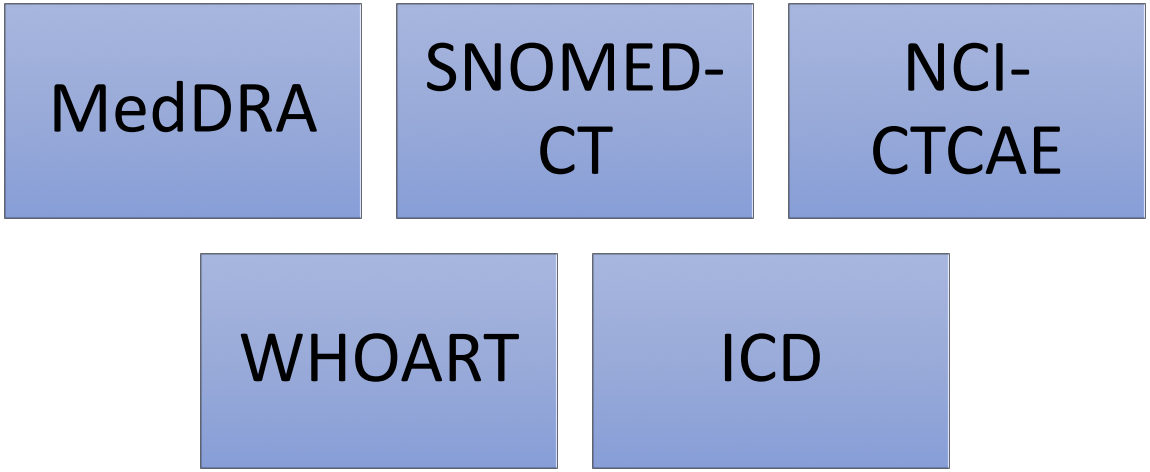
Oncology Trials

MeSH_Term variable in ***condition_browse*** could be used to filter clinical trials in several therapeutic area.

Data Table	Data Examples	Total Data Volume
Clinical Trials	NCT00004092, NCT00090844,etc	76998 (Trials)
Adverse Events	Tumour pain, Dehydration,etc	345601(Events)
Disease Conditions	Breast Neoplasm, Lymphoma,etc	227 (Condition)

Reported Events

- It contains Event name, Subject Affected, Subject at Risk and Event Categorization.
- More than 6 terminology standards were used



Challenges

- Huge Data volume
- Dynamic Data(refreshed daily)
- Standardization of adverse event (coding)
- Computationally difficult to exhaust all possible association patterns of adverse events.

Standardization

Unified Medical Language System (UMLS)

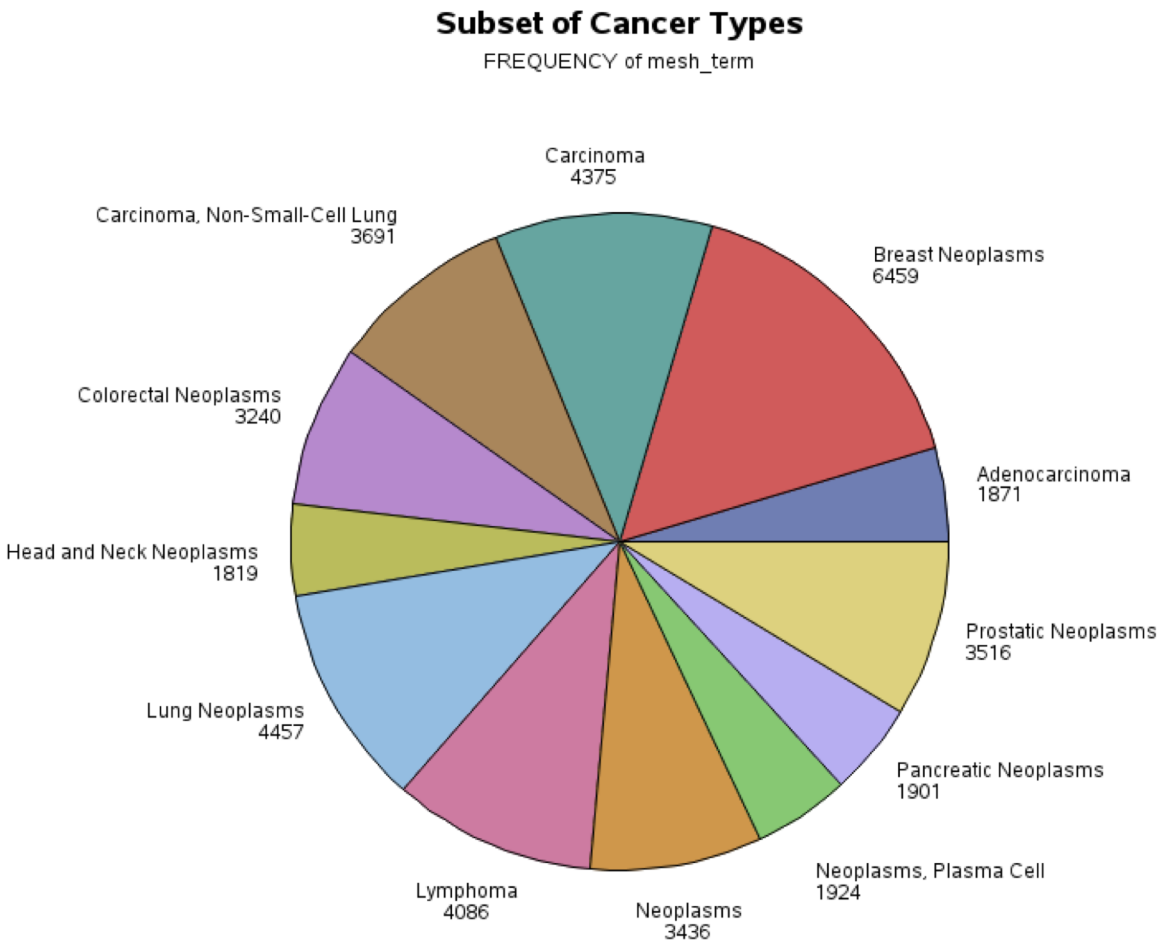
Example:

Adverse events: *Dysentery, Infectious Diarrhea and Diarrhoea Infection*

Mapped to : *concept "C0013369"*

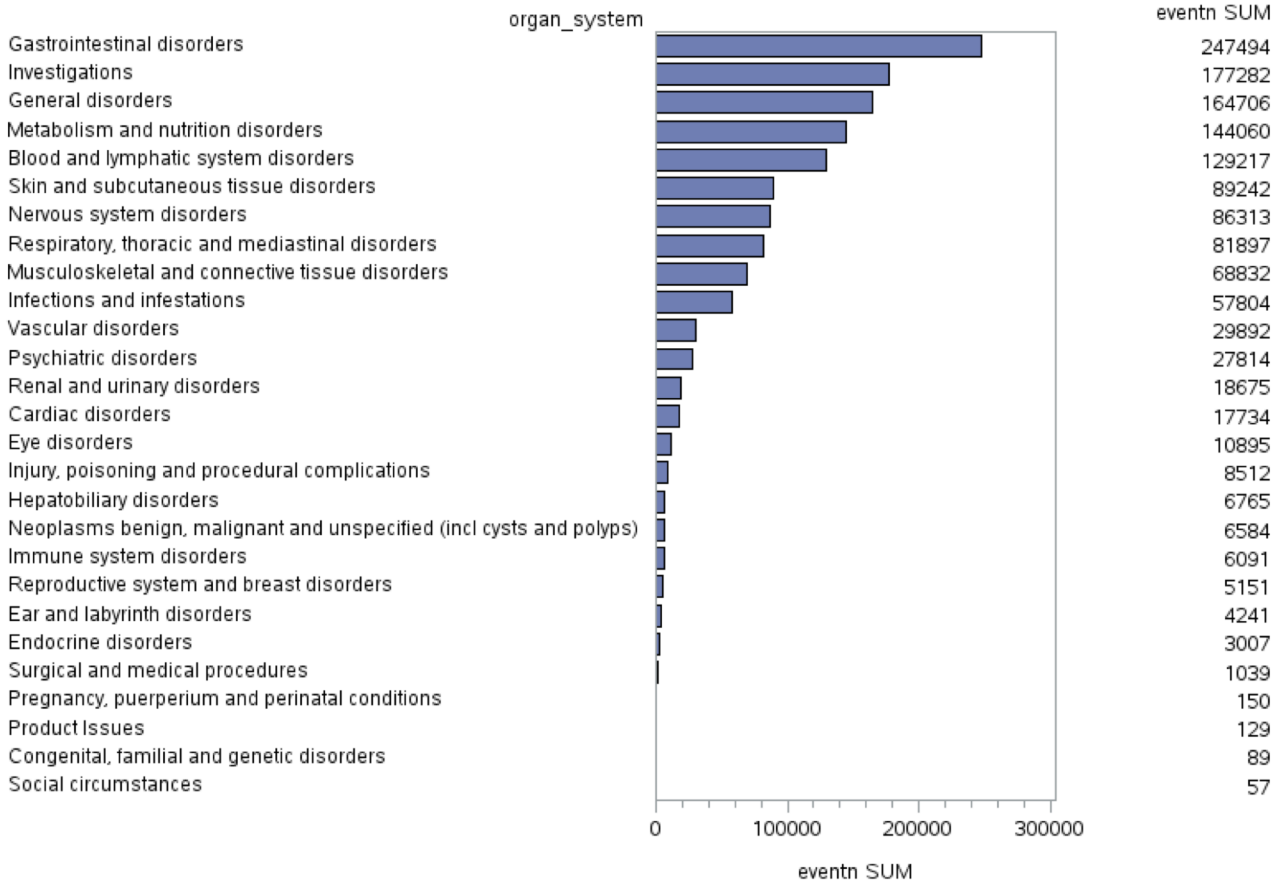
Symbiance
a relationship for success

Data insights



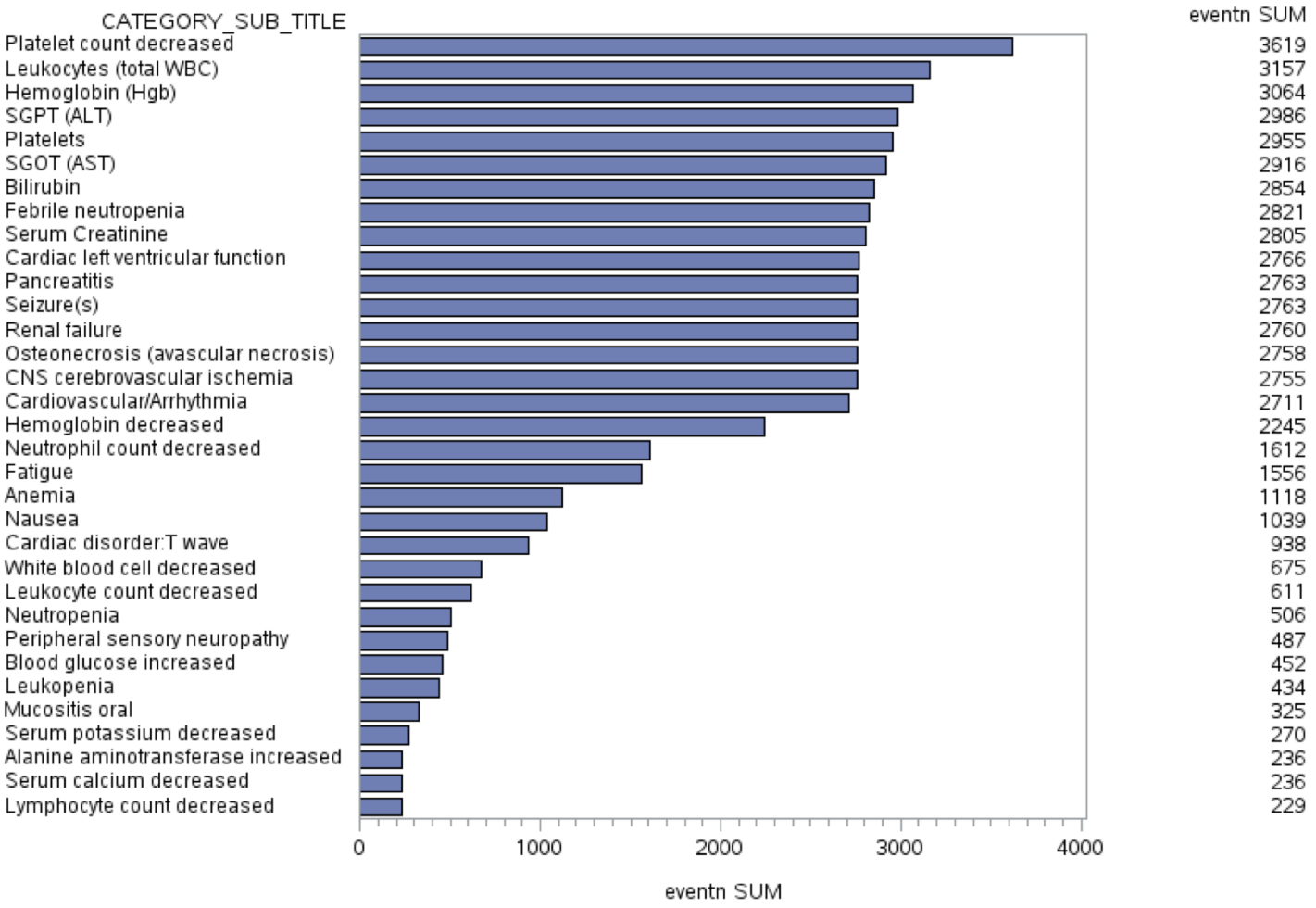
Symbiance
a relationship for success

Frequently Occuring AE SOC in Oncology Clinical Trials



Symbiance
a relationship for success

Frequently Occuring AE in Lymphoma



Symbiance
a relationship for success

Pattern mining

- Pattern mining is used to explore association patterns among adverse events.

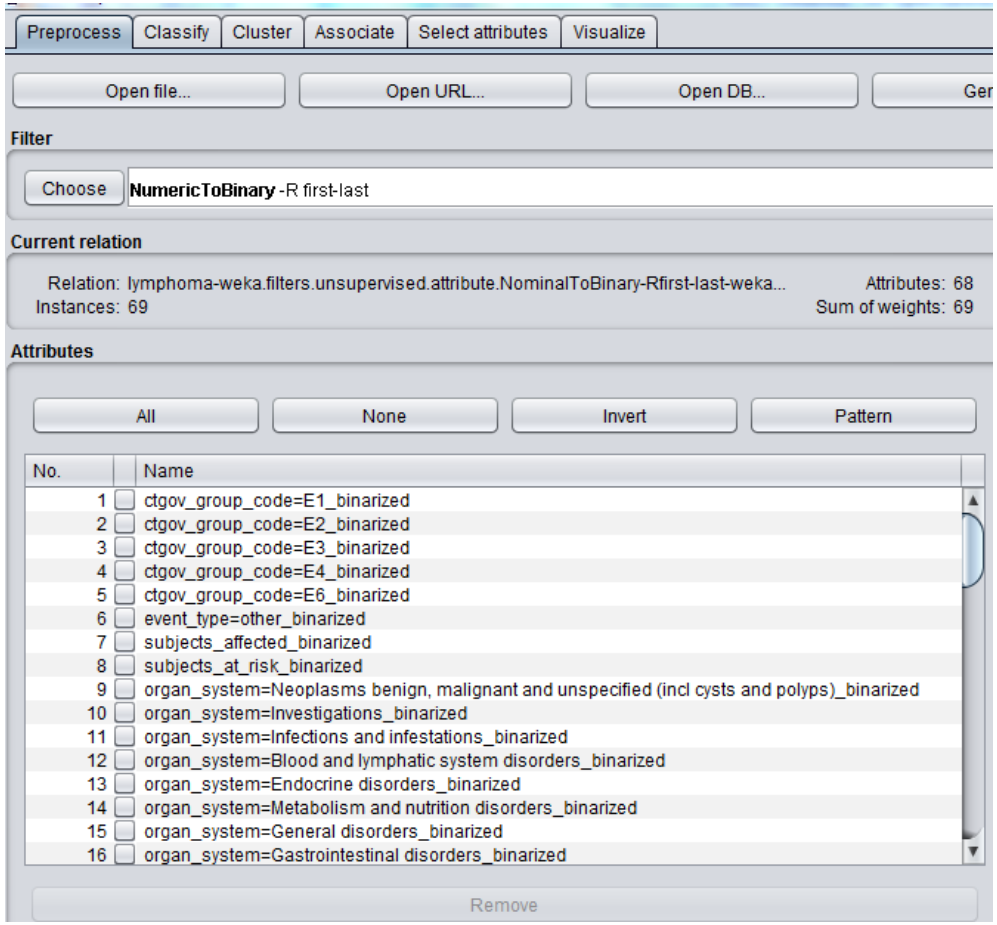
- FP- Tree algorithm: Frequent Pattern Tree

A set of patterns that provides association of Adverse Events with each other

Pattern Between two sets of adverse event is defined by X => Y.

Y -> Pattern Head ; X -> Pattern Body

WEKA software



Symbiance
a relationship for success

Input Dataset for Pattern Mining

- Adverse Events with standardised terms
- Event type (Eg. Serious AE flag etc.)
- Number of subjects affected
- Number of subjects at risk

Symbiance
a relationship for success

Pattern Mining Results

- Lift scores is a measure of strength of an association pattern
- The lift score measures how often pattern body *X* and pattern head *Y* occur together more frequently than expected if they were statically independent.
- Lift scores larger than 1 indicate positive association between *X* and *Y*.

Pattern Mining Results

Rank	Pattern Body	Pattern Head	Lift
1	Injection Site pain, Injection Site Swelling	Injection Site Erythema	13.28
2	Alanine Aminotransferase Increased, Blood Alkaline Phosphatase Increased	Aspartate Aminotransferase Increased	11.74
3	Decreased Appetite, Pain	Weight Decreased	11.10
4	Febrile Neutropenia, Neutropenia	Pneumonia	9.71
5	Dizziness, Alopecia	Hypertension	7.44
6	Anxiety, Depression	Dysgeusia	7.09
7	Pyrexia, Sepsis	Deep Vein Thrombosis	6.37
8	Myocardial Infarction, Atrial Fibrillation	Congestive Cardiac failure	5.74

Rank	Pattern Body	Pattern Head	Lift
1	Cardiac Failure, Convulsion	Death	6.76
2	Renal Failure, Osteoarthritis	Death	5.81
3	Respiratory Failure, Hypokalaemia	Death	5.75
4	Dyspnoea, Bronchopneumonia	Death	5.61
5	Angina Pectoris, Pleural Effusion	Death	5.25
6	Cerebral Haemorrhage, Pyrexia	Death	4.77
7	Cerebrovascular Accident, Hypertension	Death	4.69
8	Gastrointestinal Haemorrhage, Bronchitis	Death	4.68
9	Syncope, Infection	Death	4.43
10	Sepsis, Hypokalaemia	Death	4.26

CI Methods

- Direct Method:

  Asymmetric CI on the difference in AE rates of two groups to directly compare the two groups.

- Indirect Method:

  By comparing the lower confidence limit (LL) of the CI on the AE rate in one group to the upper confidence limit (UL) of the CI on the AE event rate in the other group.

Limitation

- Pattern mining only reveal general correlation relation between adverse events.
- Implied Clinical meaning of mined patterns need to be interpreted by domain expert.
- Tricky to handle large volume of data.

Conclusion

- Huge Volume of data, results of clinical trial could be used as a support for safety and efficacy analysis
- Mined Patterns can be used to help predict adverse events and to improve the safety of patients.

Symbiance
a relationship for success

Reference

- Harpaz R, DuMouchel W, Shah N, Madigan D, Ryan P, Friedman C. Novel Data-Mining Methodologies for Adverse Drug Event Discovery and Analysis. Clinical Pharmacology & Therapeutics. 2012;91(6):1010–1021.
- Luo Z, Johnson SB, Weng C. Semi-Automatic Induction of Semantic Classes from Free-Text Clinical Research Eligibility Criteria Using UMLS; American Medical Informatics Association Annual Symposium; Washington, DC. 2010. pp. 487–491.

Symbiance
a relationship for success

Acknowledgement

- Clinical Trials Transformation Initiative
- UMLS
- WEKA software

Symbiance
a relationship for success

Symbiance
a relationship for success