Pooling and harmonizing of safety data for a robust statistical analysis

Gian Thanei, EFSPI workshop 23-24, 9.19

Safety analysis in clinical development

Traditionally safety insights have been driven through a patient to patient assessment by safety scientists (excel sheet data wrangling)

Safety analysis in clinical development

Traditionally safety insights have been driven through a patient to patient assessment by safety scientists (excel sheet data wrangling)

- Safety profile established on a single trial (for drug label)
- Statistical analysis of safety data has been mostly counts/frequencies and incidences

Safety analysis in clinical development

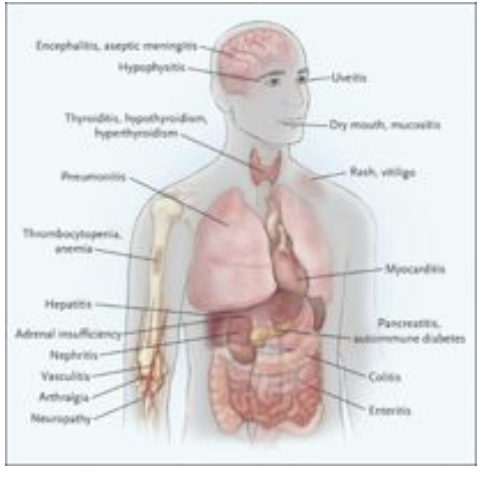
Traditionally safety insights have been driven through a patient to patient assessment by safety scientists (excel sheet data wrangling)

- Safety profile established on a single trial (for drug label)
- Statistical analysis of safety data has been mostly counts/frequencies and incidences

New interest in safety data: associations of safety occurrence with biomarkers (due to bigger data sources), treatment differentiation (understanding benefit/risk)

Context

Cancer Immunotherapy (CIT) delivers progress in treatment but comes with increased numbers/frequencies in Immune-related Adverse Events (IrAE)



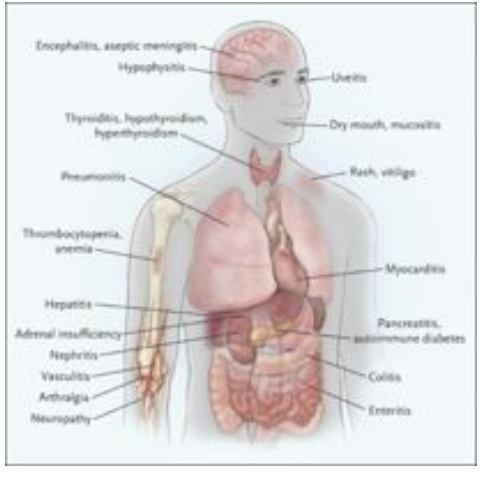
Context

Cancer Immunotherapy (CIT) delivers progress in treatment but comes with increased numbers/frequencies in Immune-related Adverse Events (IrAE)

Data: Studies across many indications in thousands of patients

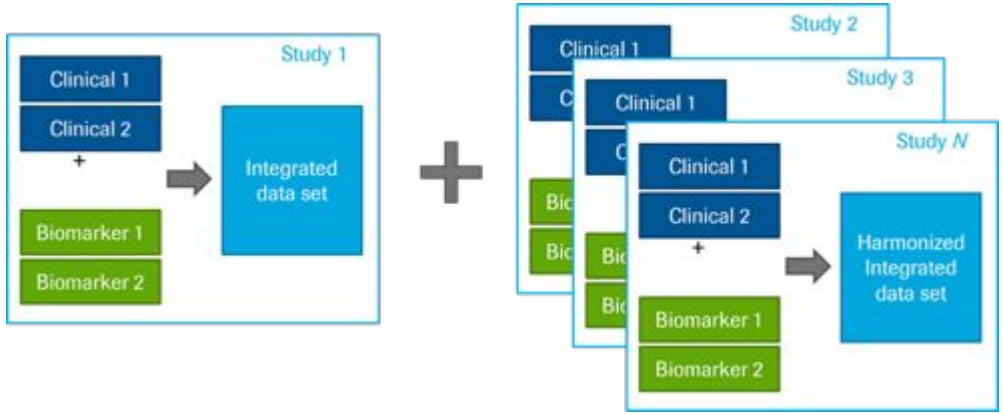
Statistics: Model IrAE's over different trials

Which infrastructure is needed?



Infrastructure: Data marts

Data marts are a collection of data sets that are **harmonized** in such a way that endpoints, biomarker measurements etc are comparable over different data sets: i.e. *"Immune-related Hepatitis has the same definition over different trials"*



Data marts at Roche

EDIS is an effort within Roche to pool and harmonize data from different trials that cover the same disease/molecule area

Data marts at Roche

EDIS is an effort within Roche to pool and harmonize data from different trials that cover the same disease/molecule area

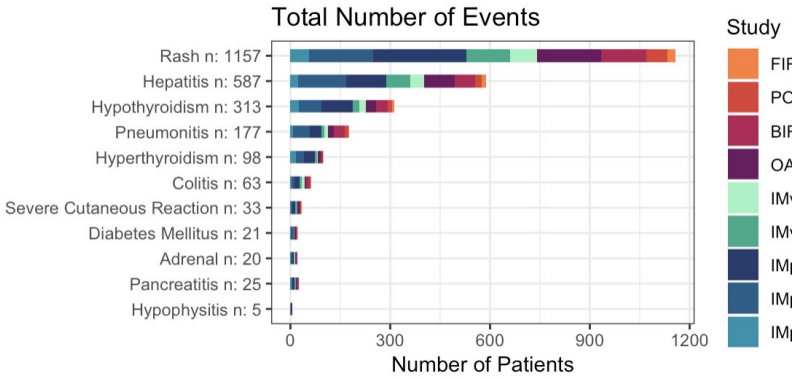
The long term objectives are:

- Robust statistical modeling of outcomes/relationships across trials
- Inform future trial design (enrichment)
- Quick querying of data in future analyses

CIT safety data mart

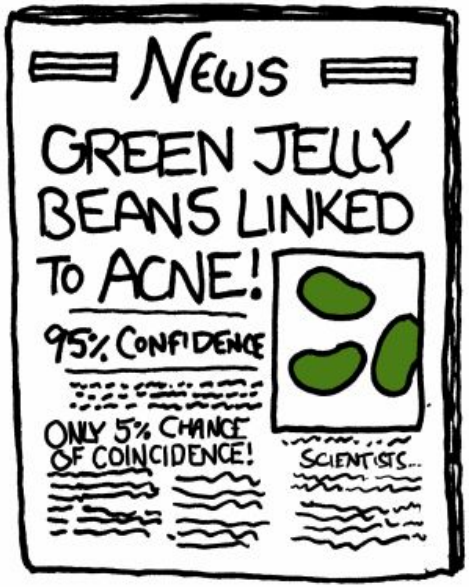
To develop a robust CIT safety profile we harmonize all CIT trials into a single safety data mart:

- 9 RCT's (more than 6000 patients in lung and bladder under CIT)
- Clinical data, biomarker, rna-seq, germline DNA, microbiome
- Rash (skin), Hepatitis (liver), Pneumonitis (lung)



Exploratory analysis of a data mart

Can you just hand the data mart to analysts (statisticians and data scientists) and hope for magic?



Exploratory analysis of a data mart

Can you just hand the data mart to analysts (statisticians and data scientists) and hope for magic?

No! You need scientists, you need to make the data accessible to the subject experts.

Shiny apps based on the teal-framework

Teal-framework is an R package to build Shiny apps (developed at Roche in the SPA-DA group)

For the CIT safety mart: The primary purpose of the CIT app is to guide clinical scientists through the data to get an overview of standard safety outputs and summaries of basic quantitative analyses

EDIS CIT Safety Explorer: Biomarker Analysis for UC22

Study Info | Data Table | Variable Browser | Demography | Distribution | Bivariate | **Safety** | Time-to-Event

Encodings

Analysis data: AST

Facetting by

STUDYID - Study ID

Color by

-- no selection --

Variable (Endpoint)

- (•) Aggregate all endpoints
- ( ) Aggregate selected endpoints

- [x] Relative Frequencies
- [ ] Stacked bars
- [x] Descending order

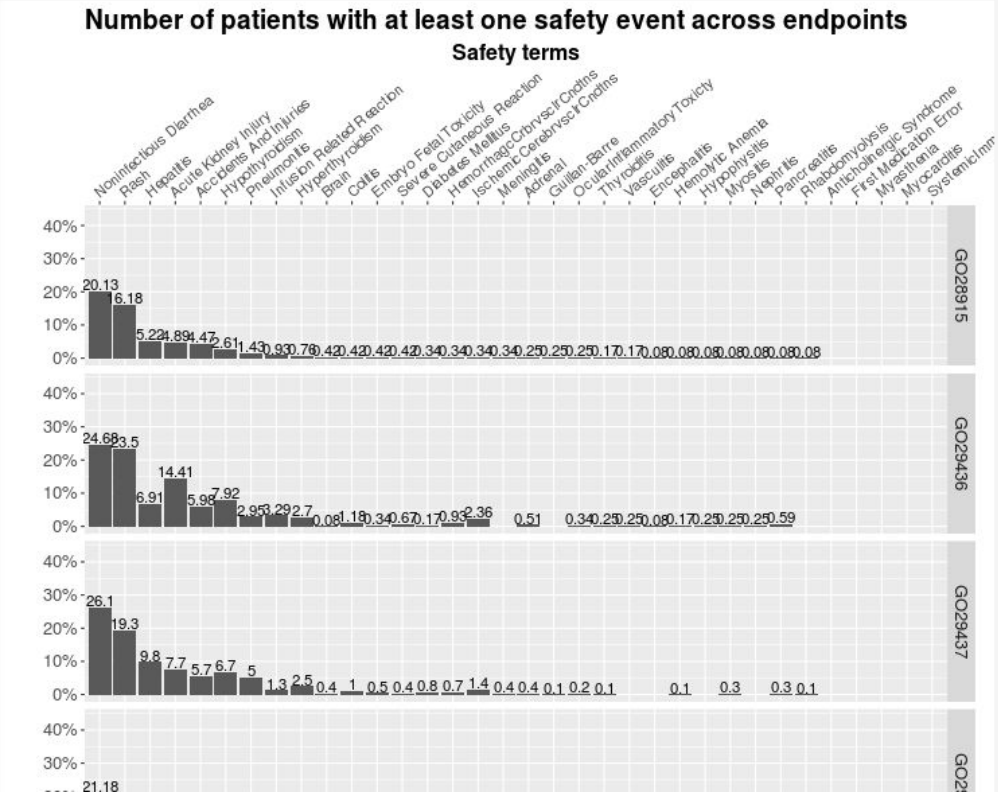
plot height

200 | 1,724 | 5,000

- [ ] Plot configuration

Show R Code

Currently filtering by AST variables not supported.



Active Filter Variables

Add Filter Variables

ASL

AST

Study Info | Data Table | Variable Browser | Demography | Distribution | Bivariate | Safety | **Time-to-Event**

Encodings

Analysis data: AST

Safety (Endpoint)

Hepatitis

Biomarker Variable

Age - Age

- (•) Dichotomization
- ( ) Trichotomization

Dichotomization for Age

0% 31% 100%

0 10 20 30 40 50 60 70 80 90 100

Range=18-92, Actual value=59

Arm Variable

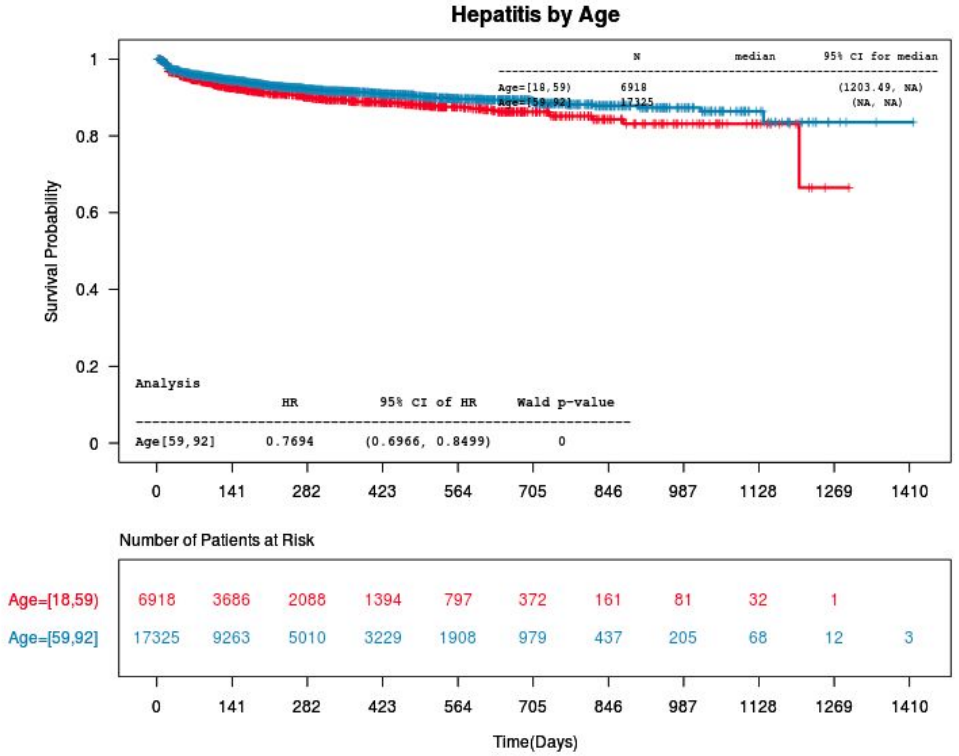
-- no selection --

plot height

600 5,000

Show R Code

Currently only filtering by AST variable time supported.



Active Filter Variables

Add Filter Variables

ASL

AST

Example: IPD-Meta Analysis

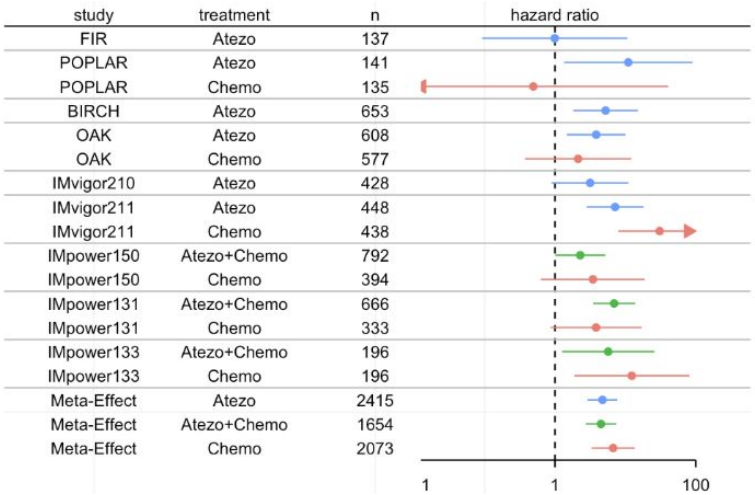
Select covariates and use individual patient data meta analysis (IPD) to assess associations:

$$\log(\lambda_{ij}) = \log(\lambda_j(t)) + \theta_j b_{ij} + \gamma_{ij}^T X_{ij} + \varepsilon_{ij}$$

time to first IrAE

biomarker of interest

other AE driving covariates



Example findings: prognostic factors to identify patients at higher risks to develop IrAE.

Conclusions: What does robust mean?

- Harmonized data marts: ensuring data is comparable

Conclusions: What does robust mean?

- Harmonized data marts: ensuring data is comparable
- Including science: humane access to large data resources

Conclusions: What does robust mean?

- Harmonized data marts: ensuring data is comparable
- Including science: humane access to large data resources
- Sample size: reproducing results over multiple trials

Acknowledgment

Works of many:

- Biostats PHC-group: Laurent Essioux, Daria Rukina
- Shiny App: Adrian Waddell, Tadeusz Lewandowski, Vincent Wolowski
- Safety Science: Scott Chambler, Flavia di Nucci, Sergio Ley