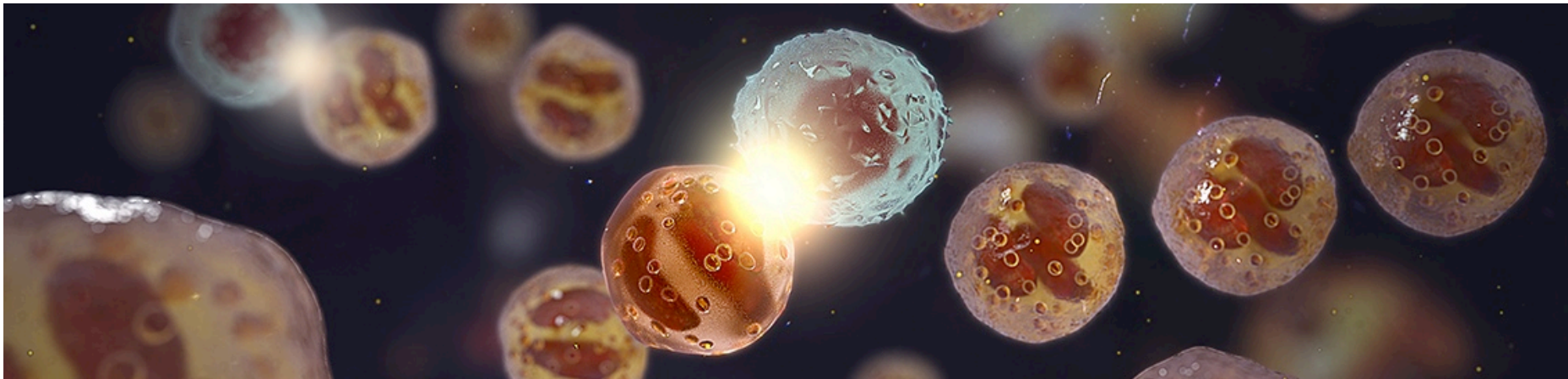


The MetaDataHub - Integrating, searching and presenting study descriptive information

Elisabeth Nyman, Kerstin Forsberg & Michail Doulis

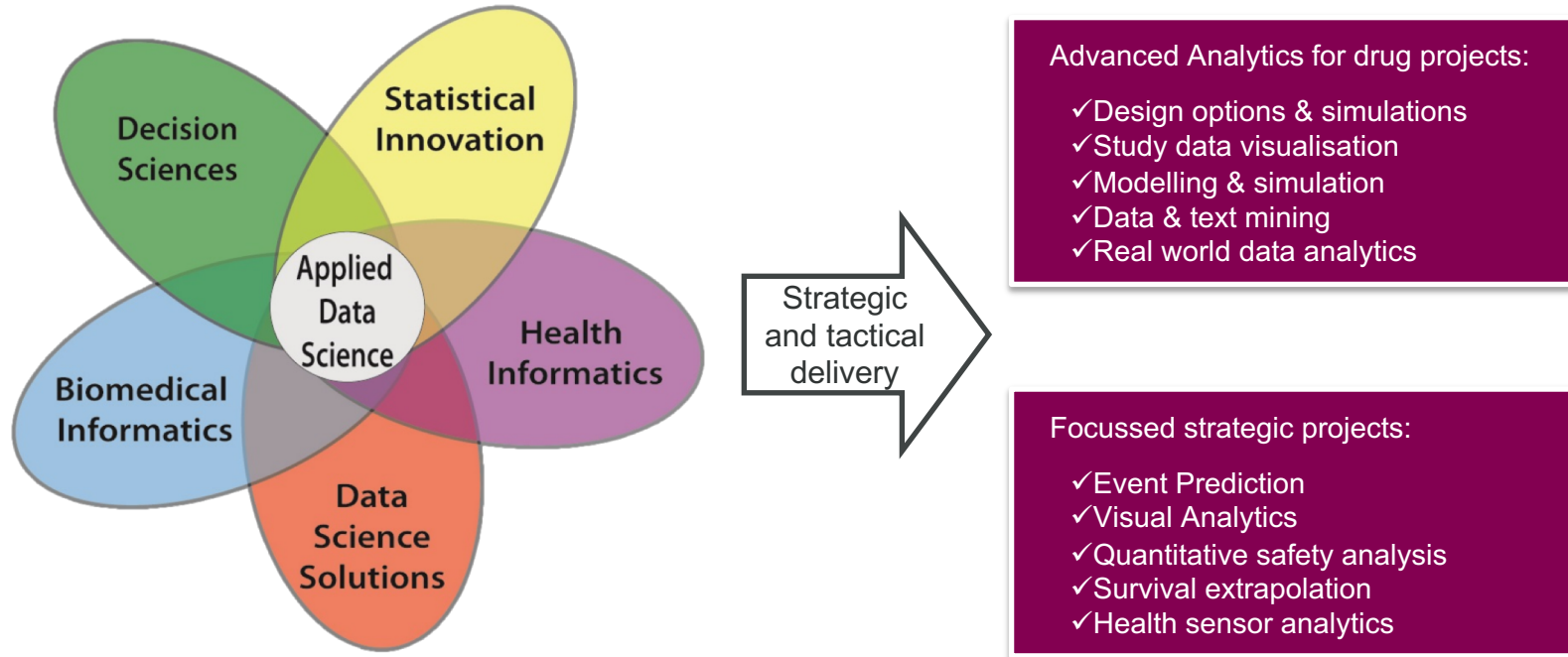
PhUSE EU Connect 2018

6 November 2018



The AZ Advanced Analytics Centre (AAC)

Transforming drug development decision making through applied data science



>40 clinical data scientists

In UK (Cambridge & Cheshire); Sweden (Gothenburg); US (Gaithersburg, MD)



The MetaDataHub - Integrating, searching and presenting study descriptive information

Presentation outline

- The purpose of the MetaDataHub
- Functionality and components
- Ongoing development & future directions
- Conclusion
- Q&A



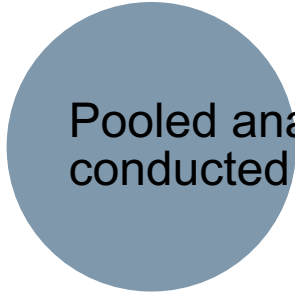
Legacy clinical data is a significant asset for pharma companies

- Used in regulatory interactions
- Re-use of data in exploring new medical opportunities
- Part of a divestment

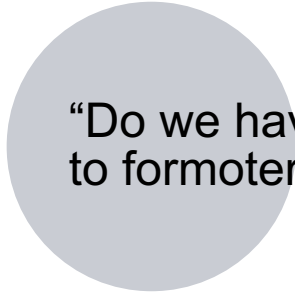
Critical that we understand our clinical studies & retain that knowledge over time



Example use cases emphasizing the need to easily identify and understand our clinical studies



Pooled analysis for a number of studies including legacy studies conducted more than 30 years ago up to recent studies



“Do we have any Turbohaler studies that compare budesonide/formoterol to formoterol with pre-dose and post-dose FEV1?”

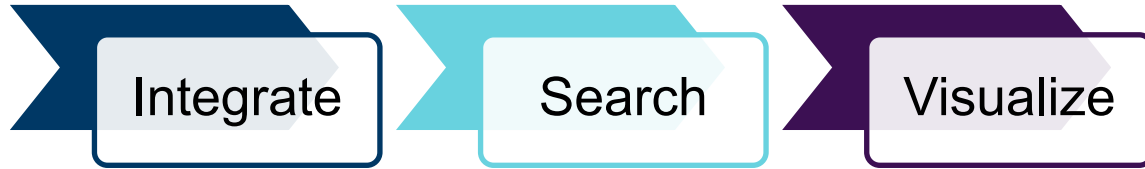


Extensive work is required to manually search multiple sources and generate comprehensive overviews

1. Find the information in different systems and documents
2. Extract the information
3. Integrate the information
4. Filter, analyze and display the information



The MetaDataHub - Simplifying identification and basic understanding of clinical studies

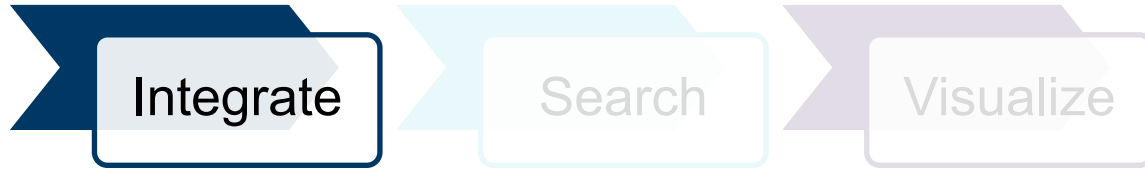


- Information accessed from available sources – no need to update
- Reusable – not a one-off solution

The MetaDataHub concept was developed as an internal crowdfunded innovation project



The MetaDataHub - Simplifying identification and basic understanding of clinical studies

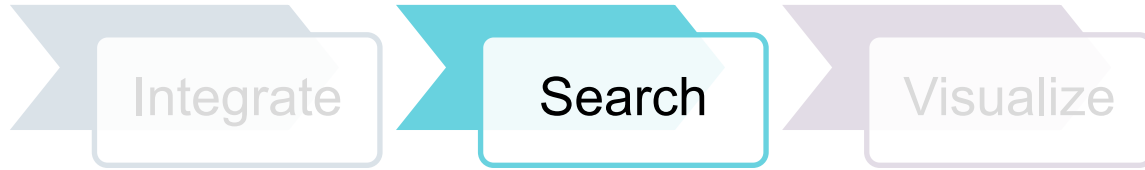


Integration of different information types is the fundament

Challenge: Common identifier across systems is needed!



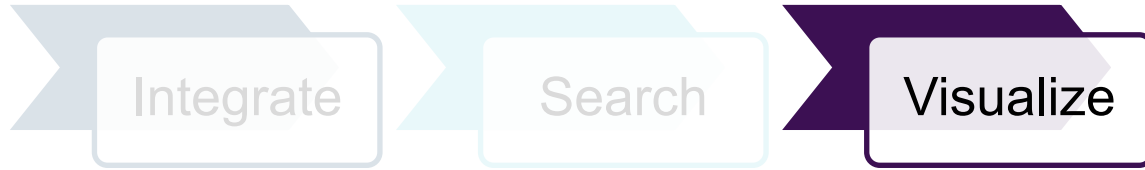
The MetaDataHub - Simplifying identification and basic understanding of clinical studies



Possibility to search across different information types, e.g. find studies with a certain length, treatment arm, and with data on a specific variable



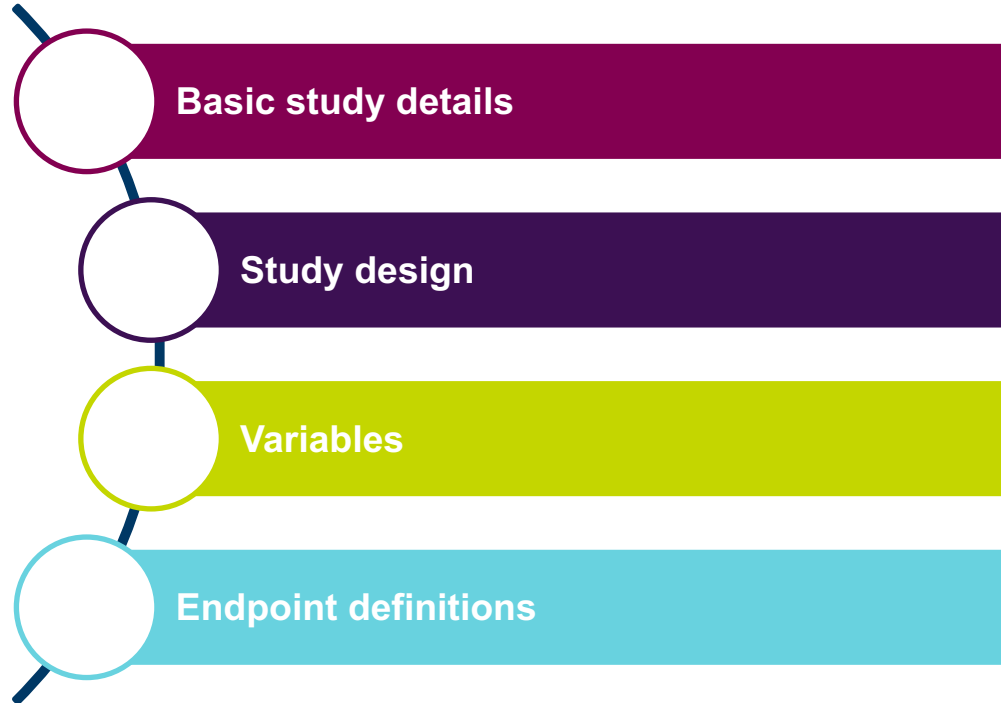
The MetaDataHub - Simplifying identification and basic understanding of clinical studies



Quickly get an overview of the studies, understand similarities and differences and find links to key documents.

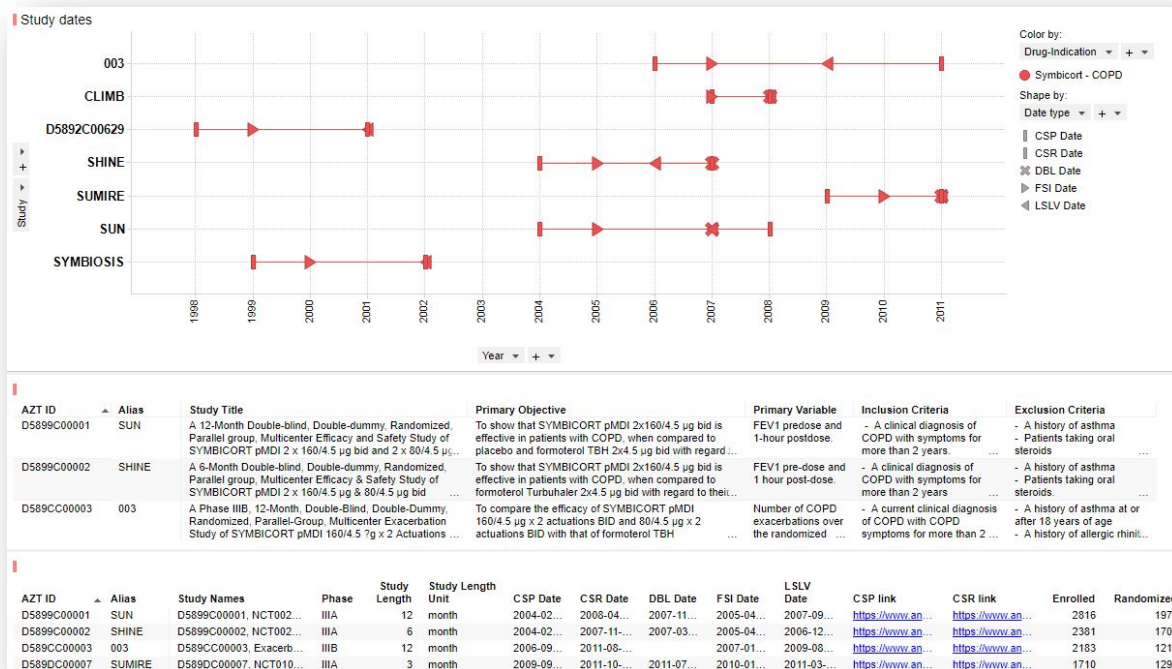


Integration of different information types is needed to get a comprehensive search and overview



Basic study details

Understanding study objectives, population and timelines



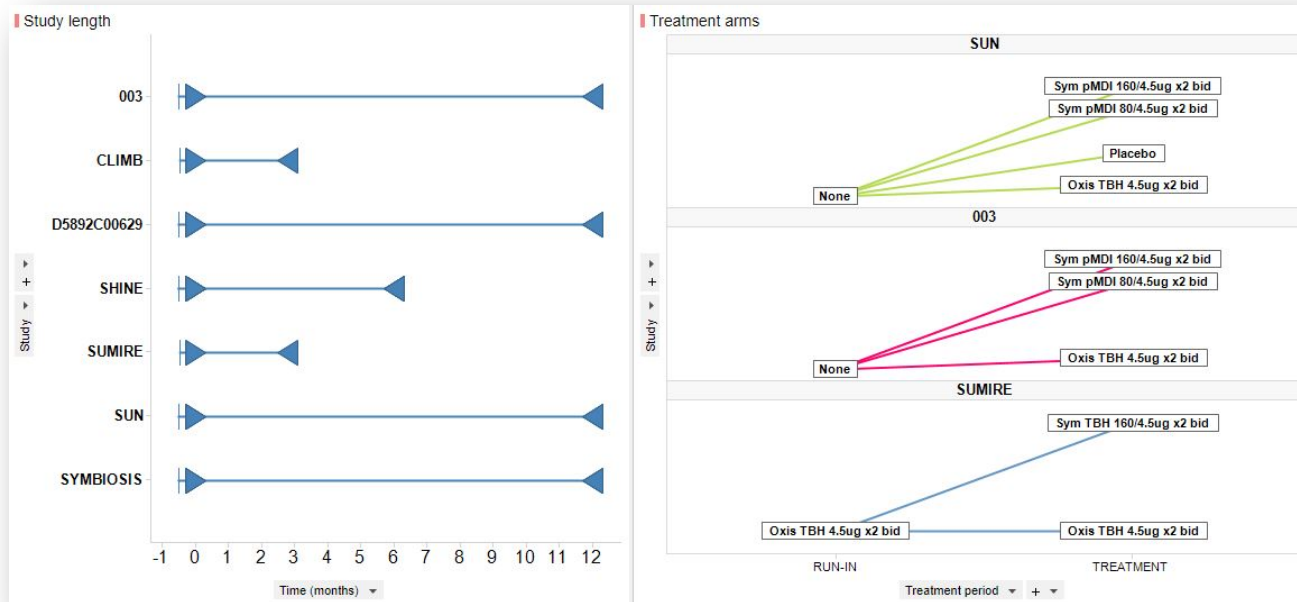
- Study synonyms
- Title
- Primary objective & variable
- Inclusion/exclusion criteria
- Phase
- Length
- Number of patients
- Dates
- Links to key documents

Sources: Internal clinical trial management systems and ClinicalTrials.gov accessed via the Aggregate Analysis of ClinicalTrials.gov (AACT)



Study design

Overview of treatment arms and study lengths



- Drug
- Dose
- Device
- Frequency
- Route
- Period

Challenge: No source with structured information on single elements

Sources: Study documents



Variables

Knowing what data is collected and where to find it



- Variable names
- Variable labels
- Dataset
- Data location

Source: Study SAS data sets, both the variables given directly in the data sets and those given as coded values



Definition of exacerbation

Hosp/ER treatment or additional corticosteroid treatment				
003	CLIMB	SHINE	SUMIRE	SUN
Hosp/ER treatment or additional corticosteroid treatment or antibiotics treatment				
D5892C00629		SYMBIOSIS		

Definition of exacerbation

AZT ID	Alias	Exacerbation definition	Original exacerbation definition	Link to source	Source document type	Source section	Created by	Date
D5899C00001	SUN	Hosp/ER treatment or additional corticosteroid treatment	Exacerbation was defined as worsening of COPD that required a course of oral steroids for treatment and/or hospitalization. Episodes of COPD worsening that ...	https://www.an...	Clinical Study Report	7.2.2.3	Elisabeth Nyman	2015-12-01
D5899C00002	SHINE	Hosp/ER treatment or additional corticosteroid treatment	A COPD exacerbation was predefined as worsening of COPD that required a course of oral steroids and/or hospitalization for treatment. Episodes of ...	https://www.an...	Clinical Study Report	7.2.2.3	Elisabeth Nyman	2015-12-01
D5899C00003	003	Hosp/ER treatment or additional corticosteroid treatment	Exacerbations were defined as worsening of COPD that required a course of oral steroids for treatment and/or hospitalization. The use of oral steroids and ...	https://www.an...	Clinical Study Report	5.5.1.1	Elisabeth Nyman	2015-12-01

- Original definition
- “Mapped definition”
- Mapping details
- Source

Capturing endpoint definition relevant in many measurements, e.g. definition of exacerbation (time to first exacerbation and number of exacerbations)

Source: Study documents



Identify studies fulfilling specific criteria by searching across information types

7 studies identified

Basic study details

AZT ID	Alias	Study Names	Phase	Study Length	Study Length Unit	CSP Date	CSR Date	DBL Date	FSI Date	LSLV Date	CSP link	CSR link	Enrolled	Randomized	Min Age	Max Age	Hospital Status
D5892C00629	D5892C006...	D5892C00629, SD-039-0...	IIIA	12	month	1998-11...	2001-12...		1999-03...	2001-07...	https://www.an...	https://www.an...	980	812	40	150	Outpatient
D5892C00670	SYMBIOSIS	D5892C00670, SD-039-0...	IIIA	12	month	1999-11...	2002-09...		2000-04...	2002-02...	https://www.an...	https://www.an...	1144	1022	40	150	Outpatient
D5899C00001	SUN	D5899C00001, NCT0020...	IIIA	12	month	2004-02...	2008-04...	2007-11...	2005-04...	2007-09...	https://www.an...	https://www.an...	2816	1978	40	150	Outpatient
D5899C00002	SHINE	D5899C00002, NCT0020...	IIIA	6	month	2004-02...	2007-11...	2007-03...	2005-04...	2006-12...	https://www.an...	https://www.an...	2381	1708	40	150	Outpatient

Studies

Indication

☒ Asthma

☒ COPD

☒ Healthy

☐ (Empty)

Primary study drug

☒ Oxis

☒ Pulmicort

☒ Symbicort

☒ Tralokinumab

Phase

Type to search in list

(All) 3 values

IIIA

IIIB

IV

Study length (month)

1

26

☒ Include empty values

Study

Type to search in list

Study details

Minimum age

0

40

☒ Include empty values

Maximum age

2

150

☒ Include empty values

Randomized patients

0

18124

DBL date

2006-03-23

2016-06-03

☒ Include empty values

Primary variable

Type to search in list

(All) 7 values

FEV1 pre-dose and 1 hour post-dose.

FEV1 predose and 1-hour postdose.

Number of COPD exacerbations ove...

Pre-dose FEV1 measured with the s...

Pre-dose forced expiratory volume i...

The number of severe exacerbation...

Design - Treatment arm components

Drug (any arm and period)

Type to search in list

(All) 6 values

Oxis

Placebo

Prednisolone

Pulmicort

Spiriva

Symbicort

Compound (any arm and period)

Type to search in list

(All) 6 values

Budesonide

Budesonide+Formoterol

Formoterol

Placebo

Prednisolone

Tiotropium

Strength (any arm and period)

Type to search in list

(All) 7 values

160

160/4.5

18

30

320/9

Design - Treatment arms

Treatment arm

Type to search in list

(All) 11 values

Oxis TBH 4.5ug x2 bid

Oxis TBHorig 4.5ug x2 bid

Placebo

Pulm pMDI 160ug x2 bid

Pulm pMDI 160ug x2 bid + Oxis TBH 4.5ug x2 bid

Pulm TBHorig 160ug x2 bid

Spi HH 18ug x1 qd (+ Bri TBHorigrig 0.4mg as r...

Sym pMDI 160/4.5ug x2 bid

Sym pMDI 80/4.5ug x2 bid

Sym TBH 160/4.5ug x2 bid

Sym TBH 320/9ug x1 bid + Spi 18ug x1 qd (+ Br...

Run-in treatment

Type to search in list

(All) 4 values

None

Oxis TBH 4.5ug x2 bid

Oxis TBHorig 4.5ug x2 bid + PRED 30mg qd

Spi HH 18ug x1 qd (+ Bri TBHorigrig 0.4mg as r...

Exacerbation definitions

Exacerbation category

☒ Definition of exacerbation

☐ Definition of mild exacerbation

☒ No exacerbation endpoint

☐ (Empty)

Definition of exacerbation

Type to search in list

(All) 2 values

Hosp/ER treatment or additional corticoster...

Hosp/ER treatment or additional corticoster...

Variables

SAS Data Set

Type to search in list

(All) 6 values

_PFT00

ADOCCT

ADPULM

ADSL

R_PULM

RD_SUBJ

Variable label

reversibility

(All) 3390 values

% of Reversibility at V1 Category

% of Reversibility at Visit 1

Baseline Reversibility (%)

Change From Baseline Reversibility (%)

FEV1 reversibility (% of P.N.) at scr

FEV1 reversibility at screening

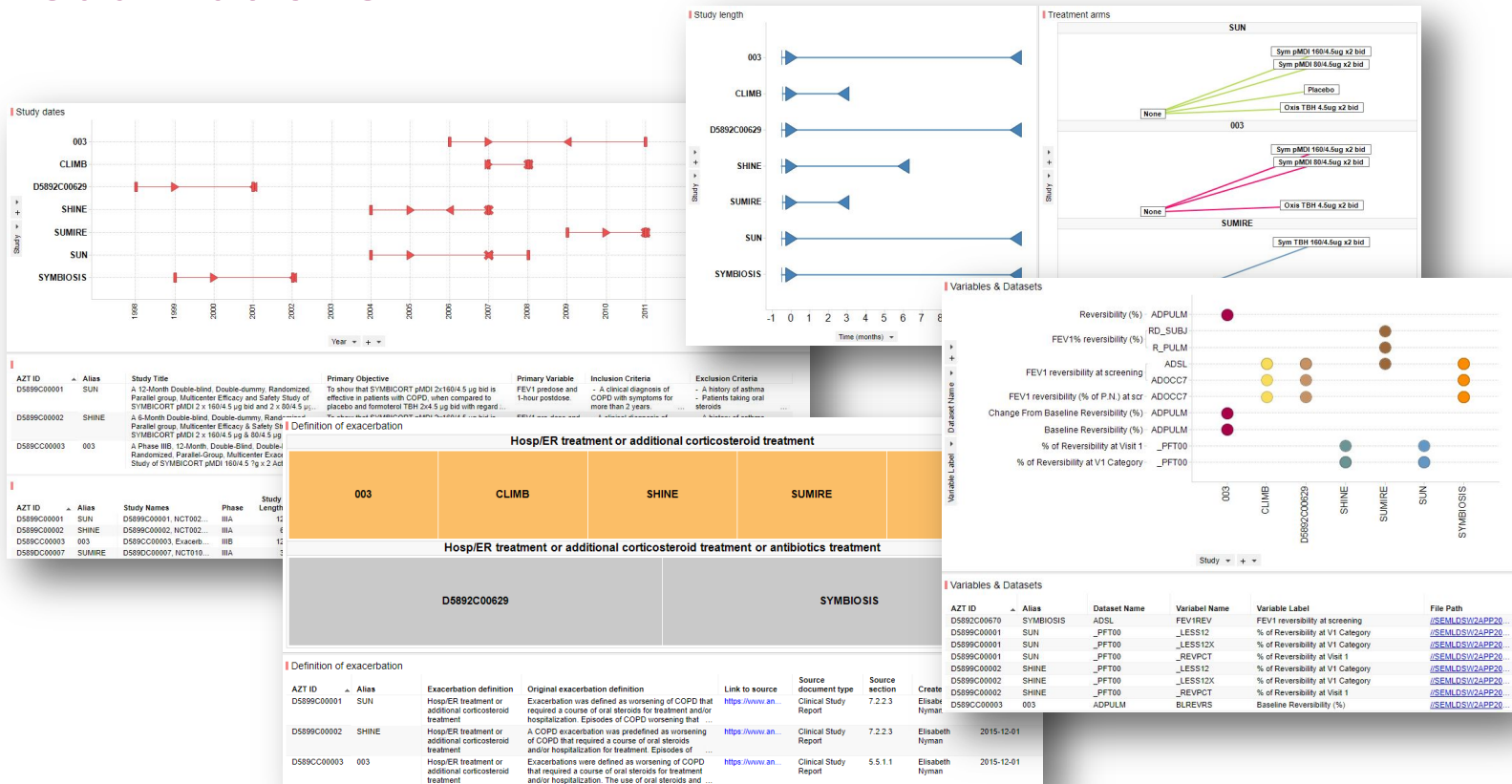
FEV1% reversibility (%)

Reversibility (%)

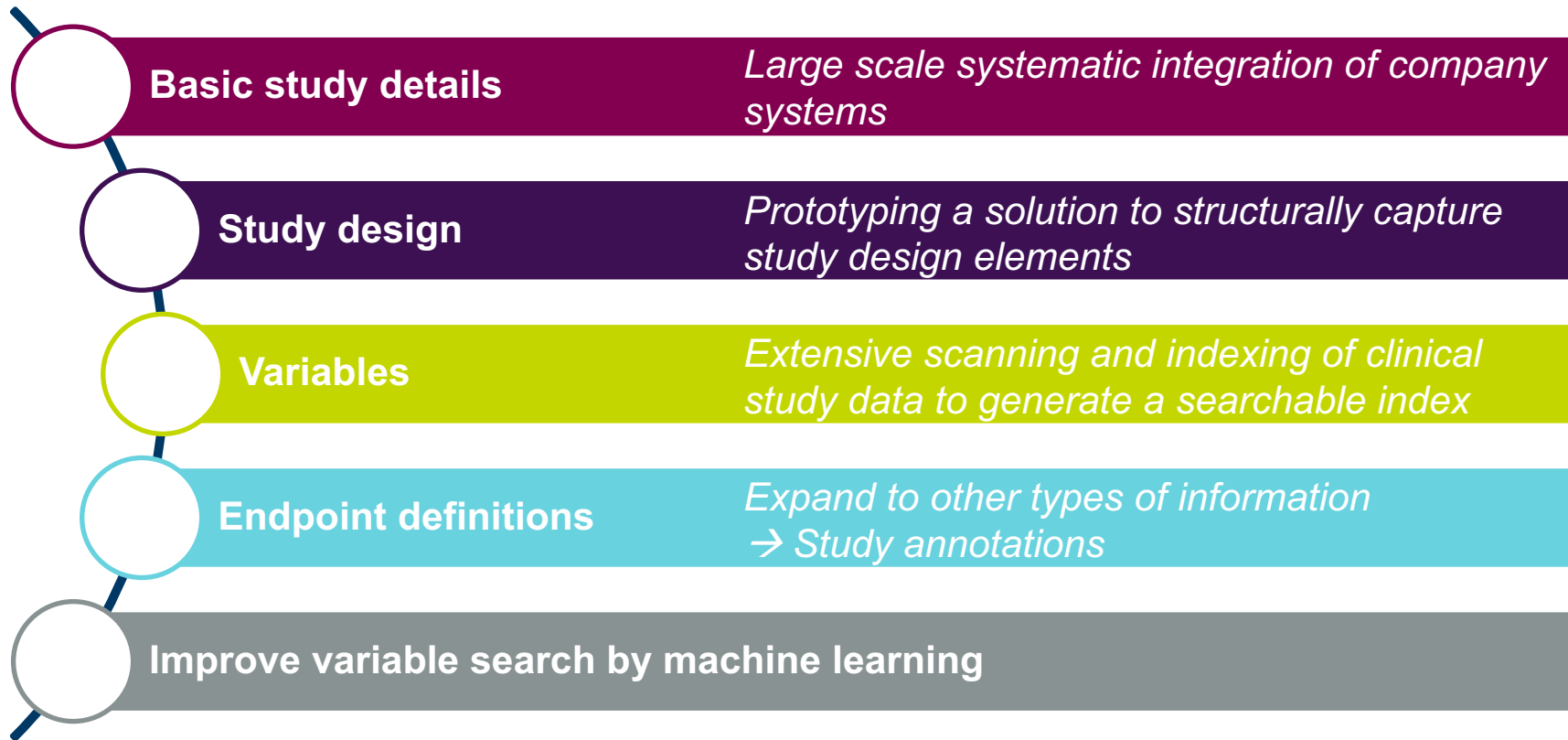
(Empty)



Building a basic understanding by exploring the different visualizations



Ongoing development & future directions



Improve variable search by machine learning

Issue

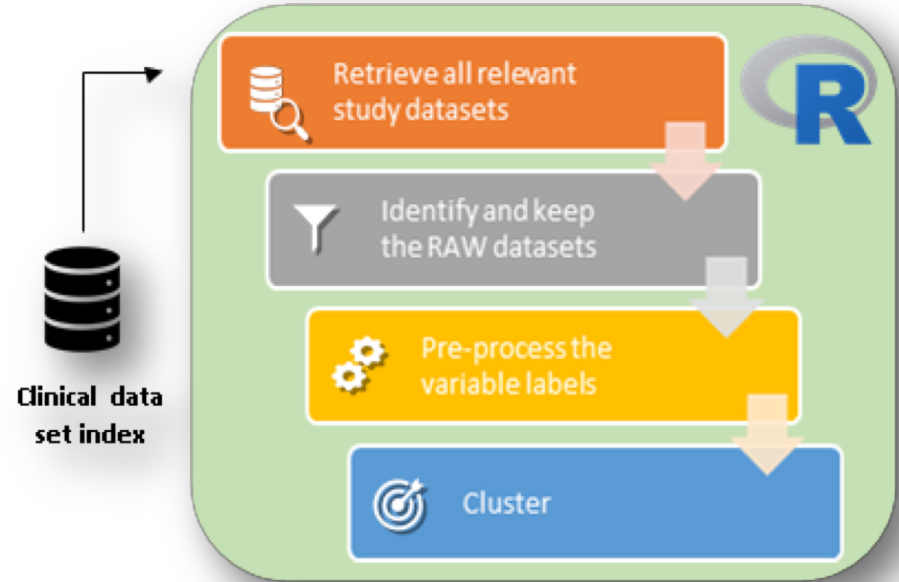
Name of the same variable may differ across studies

Objective

Explore if search of similar variables could be improved by applying machine learning algorithms

Approach

- Variable labels were used as a text corpus and a hierarchical document clustering method was applied.
- Pre-processing focused on emphasizing medical terms



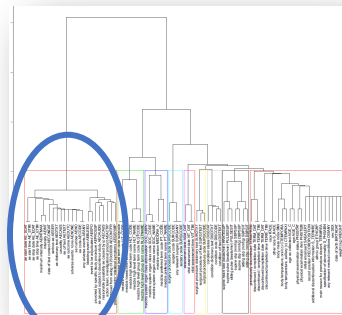
Improve variable search by machine learning

Results

Variables with similar medical content are clustered together or in adjacent clusters

Conclusion

This method can be applied e.g. when searching for similar variable across studies



ae start date AE_SDAT
ae start time AE_STIM
ae stop date AE_EDAT
ae stop time AE_ETIM
maximum intensity AEINTMAX
intensity AEINT
date of last contact AECONDAT
serious ae AESER
ae dictionary text AEDTXT
causality AECAUS
ae crf text AETEXT
ae outcome AEOUTC
adverse events: none AENONE
ae number AENO
ae code AECODE
any adverse events AEANY
bleeding event AEBLEED
event to be adjudicated AEEPO
provoked by specific study procedure AEPROSSP
ae caused subject to discontinue study AECAUDIS
reasonable possibility ae caused by ip AECAUS
action taken, investigational product AEACTIP
continued from aelog d1250c00008 AECONTFR
predefined gastro-intestinal symptoms AEGISYMP



We must be able to easily identify and understand our clinical studies



Example use cases emphasizing the need

Pooled analysis for a number of studies including legacy studies conducted more than 30 years ago up to recent studies

“Do we have any Turbohaler studies that compare budesonide/formoterol to formoterol with pre-dose and post-dose FEV1?”



The MetaDataHub is simplifying identification and basic understanding of clinical studies



The search functionality simplifies identification of studies

Comprehensive overviews helps to understand the studies

Reusable assembly of information where the information is accessed directly from the sources





Confidentiality Notice

This file is private and may contain confidential and proprietary information. If you have received this file in error, please notify us and remove it from your system and note that you must not copy, distribute or take any action in reliance on it. Any unauthorized use or disclosure of the contents of this file is not permitted and may be unlawful. AstraZeneca PLC, 1 Francis Crick Avenue, Cambridge Biomedical Campus, Cambridge, CB2 0AA, UK, T: +44(0)203 749 5000, www.astrazeneca.com

