verisian

Traceability and AI for Better Understanding, Communication, and QC of Trials

Richardus Vonk
Strategic Advisor, Verisian

Disclaimer

- The views expressed in this presentation are the personal views of the presenter and do not necessarily represent the view of his past or current employer(s) and/or collaboration partners

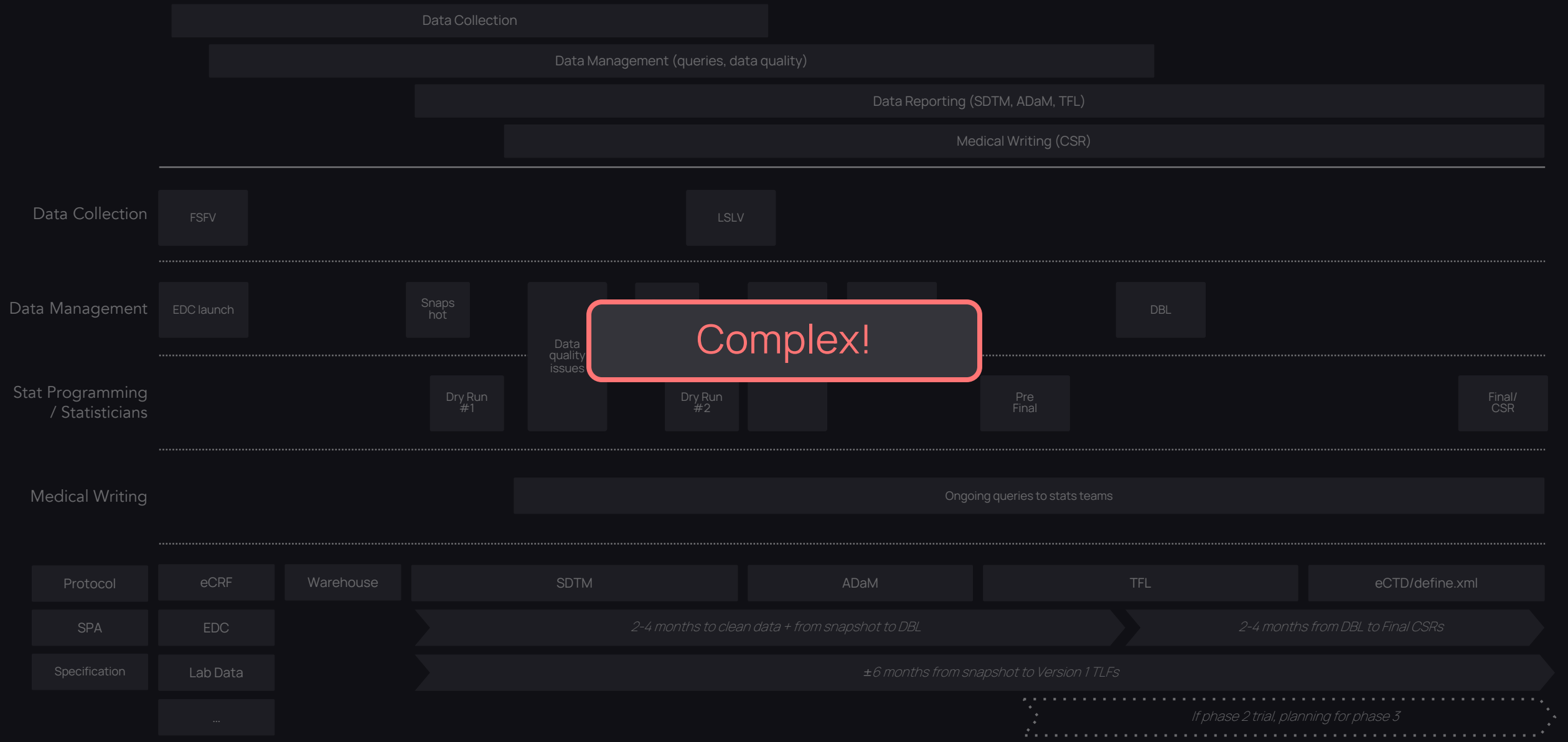
High attrition rates, high costs, and long development times
call for more agile processing of data,
to quantify the risks and facilitate timely decisions

Leading to more agile drug development

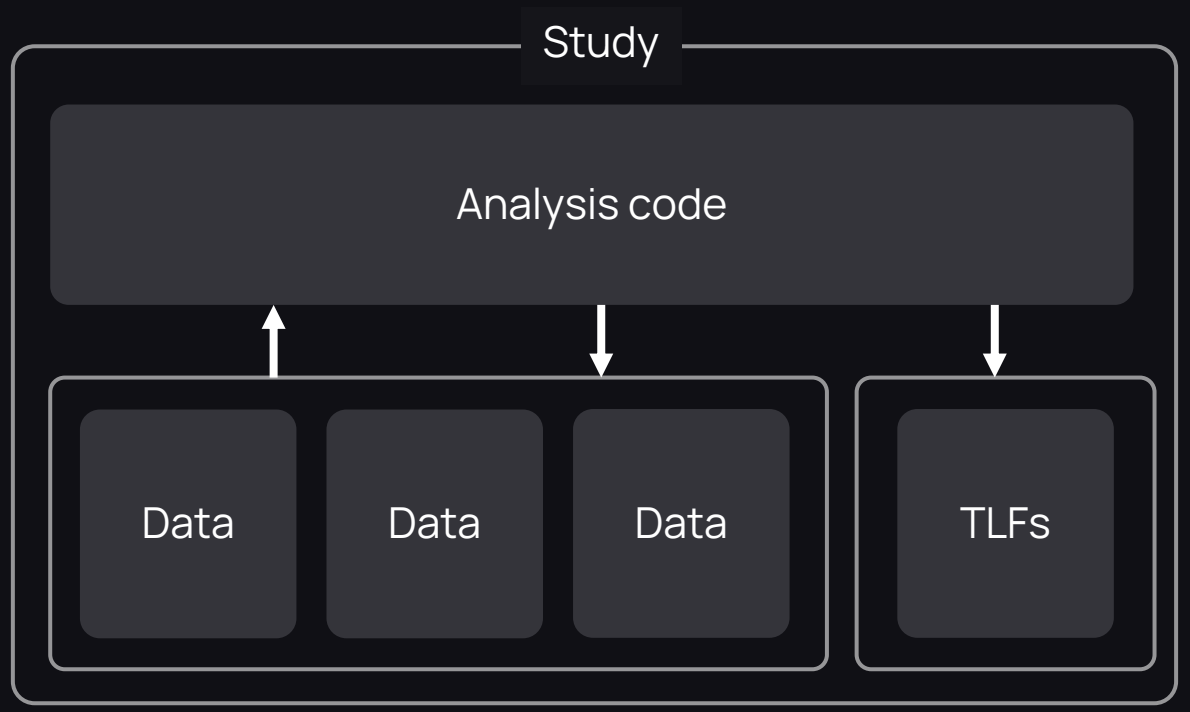
The availability of data
and the ability to derive action from data
will determine the success of drug development programs

Why can’t we release a drug immediately after the last pivotal trial?

Drug submission data workflow

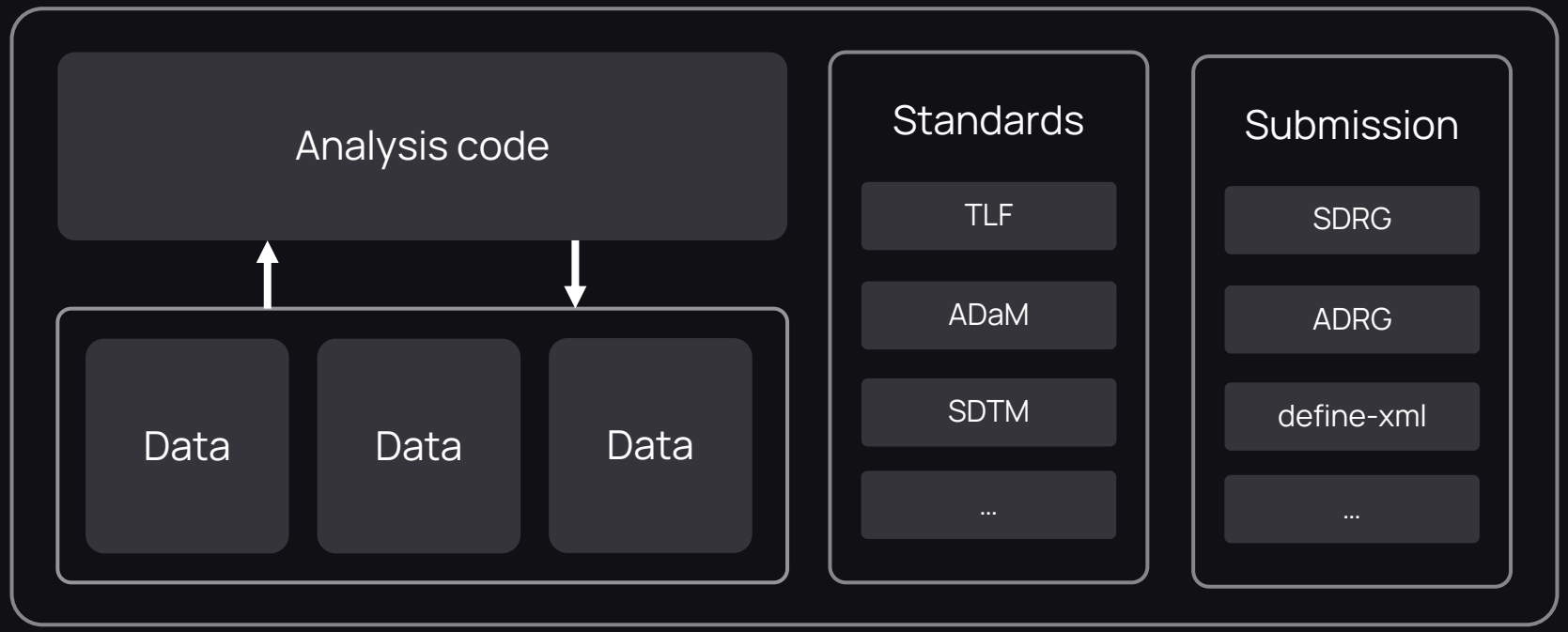


The challenge is to understand data and code, and how they interact

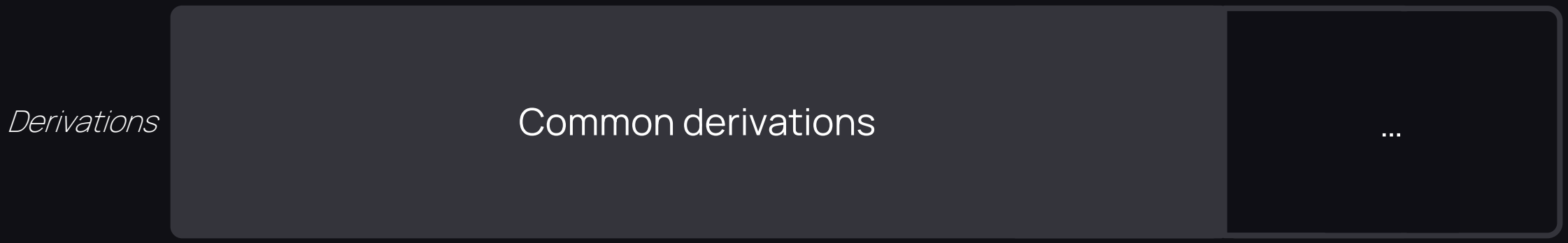


Building Trust:
Validity, Traceability, Reproducibility

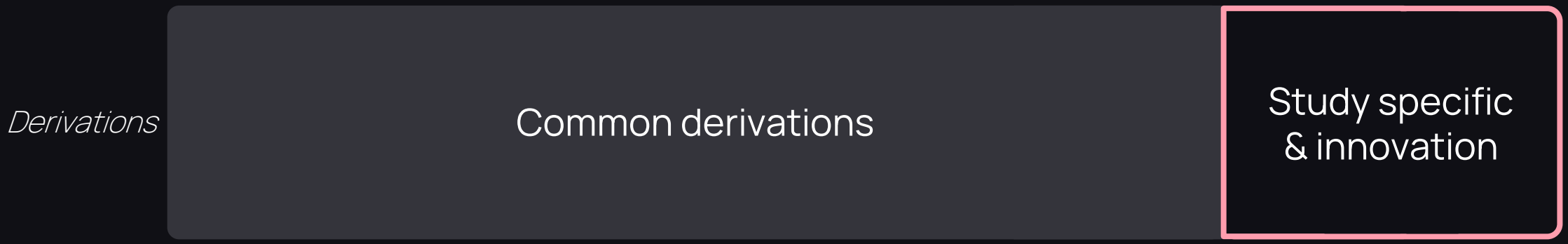
Once data are collected: what is a clinical trial?



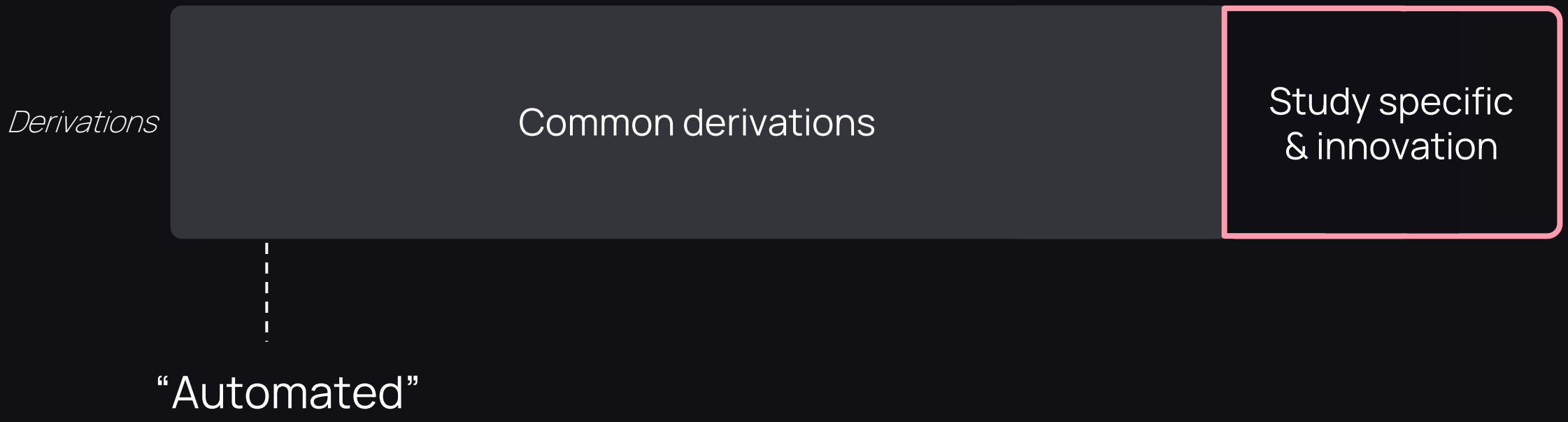
Most derivations in a study are standardized



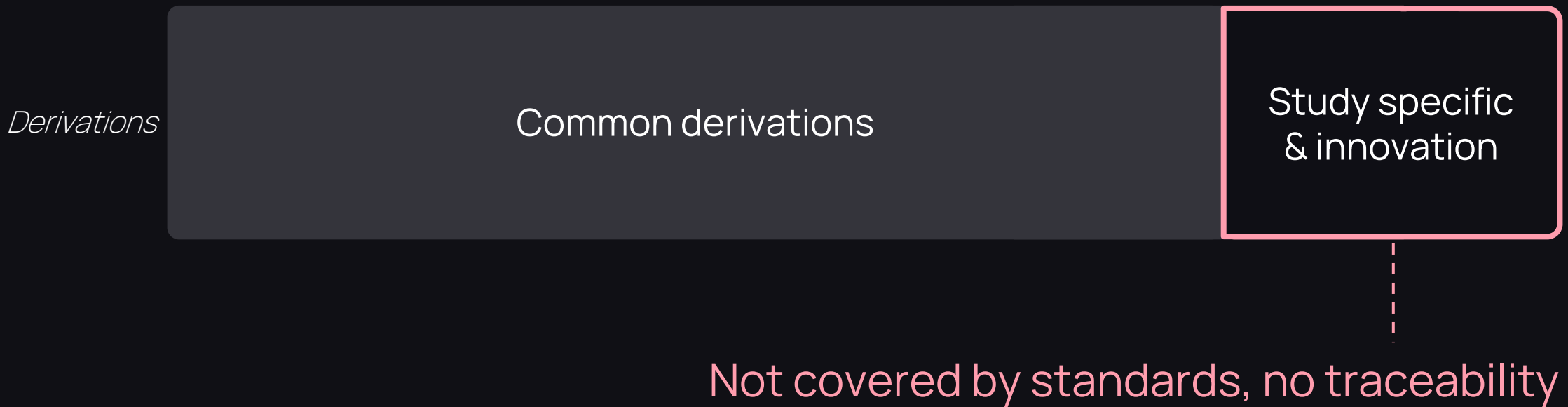
Most derivations in a study are standardized



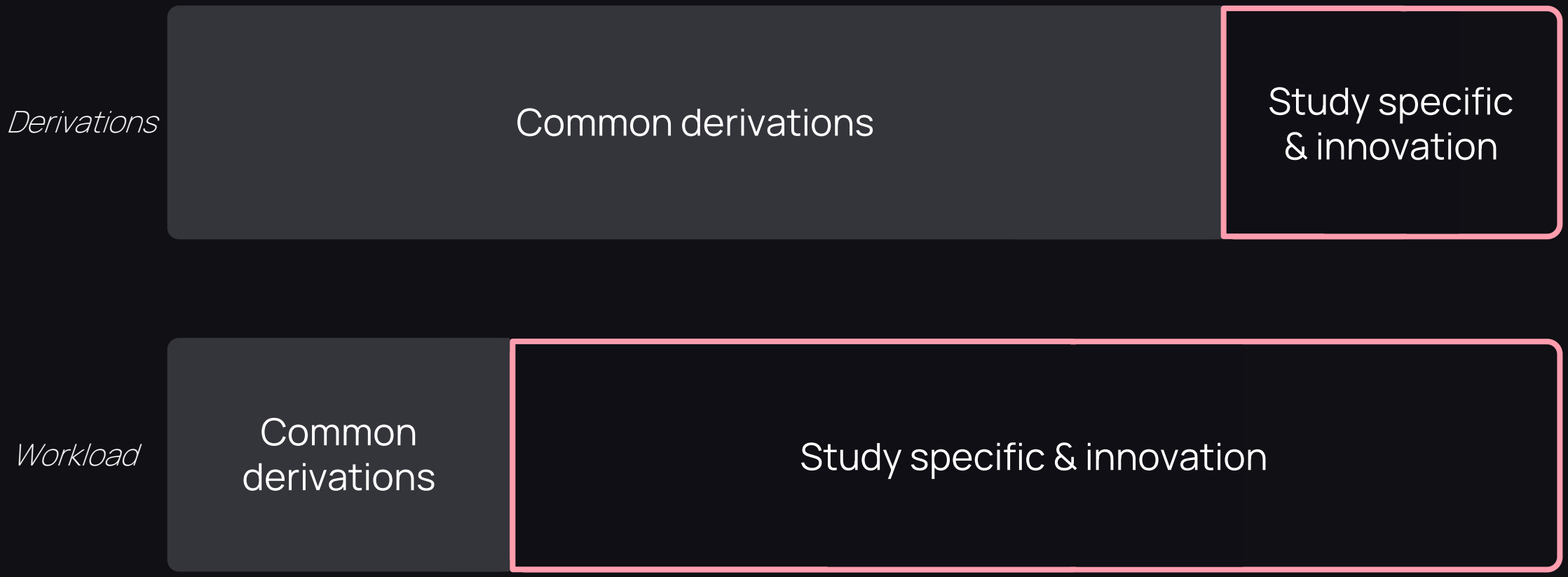
Most derivations in a study are standardized



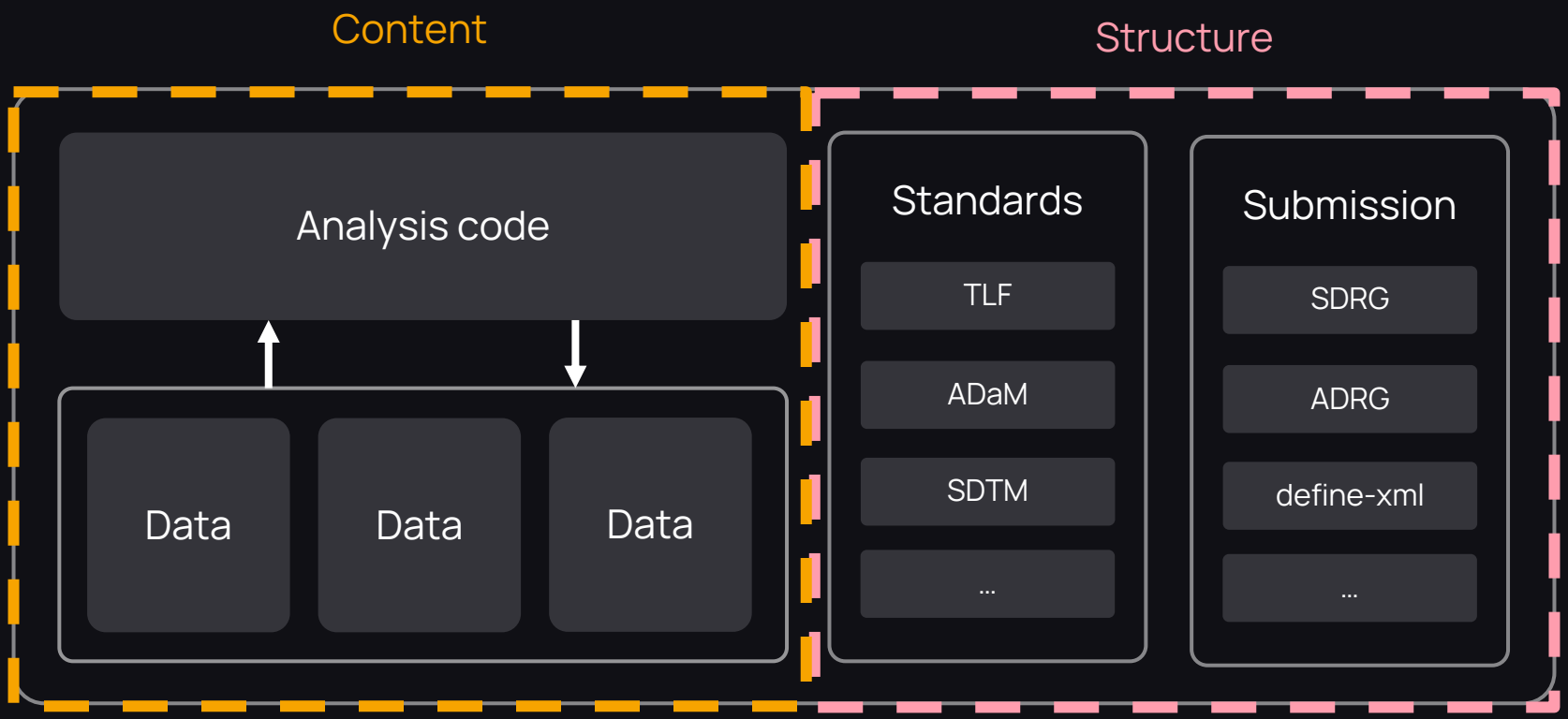
Most derivations in a study are standardized



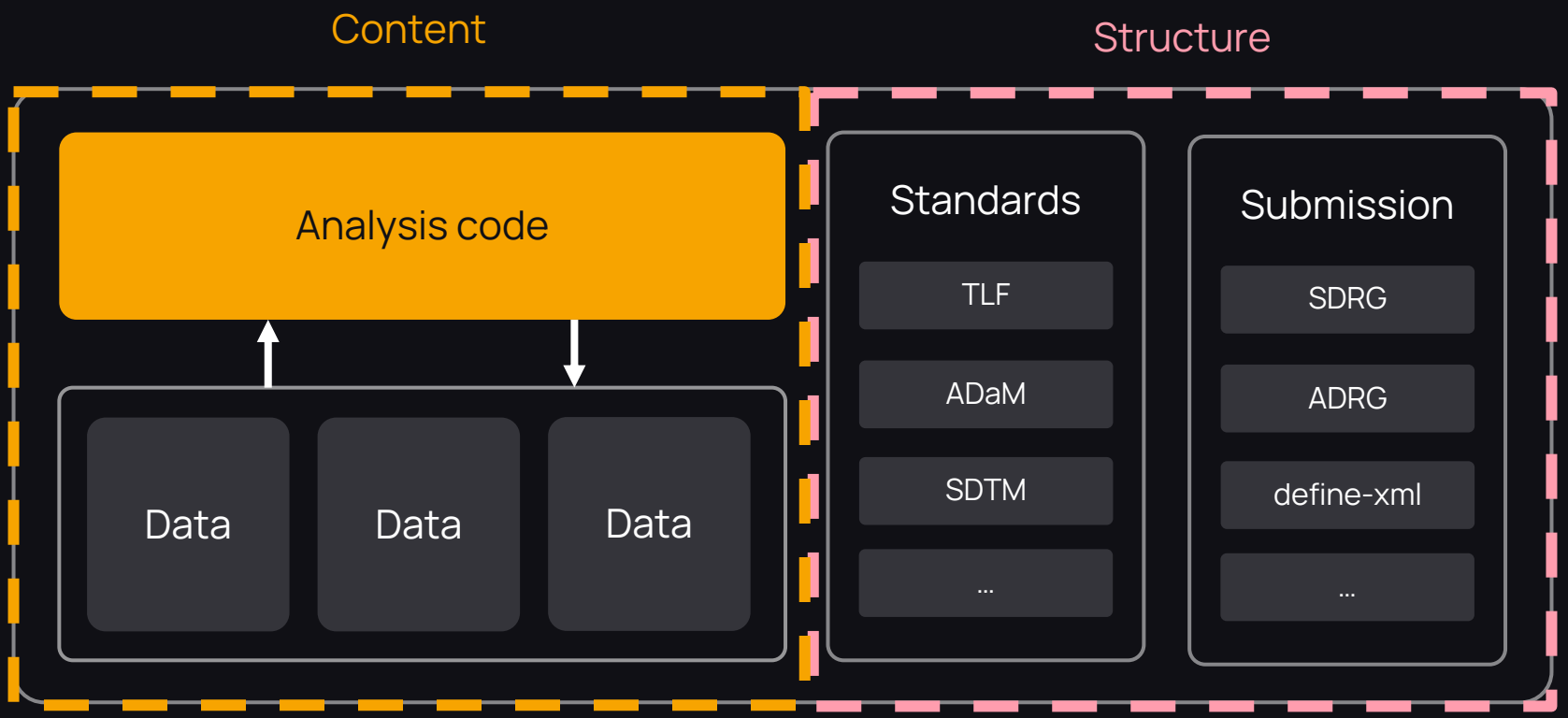
Most derivations in a study are standardized



What is a clinical trial?



The source of truth of content of a trial analysis is its code



What matters is the content of the study, not the structure

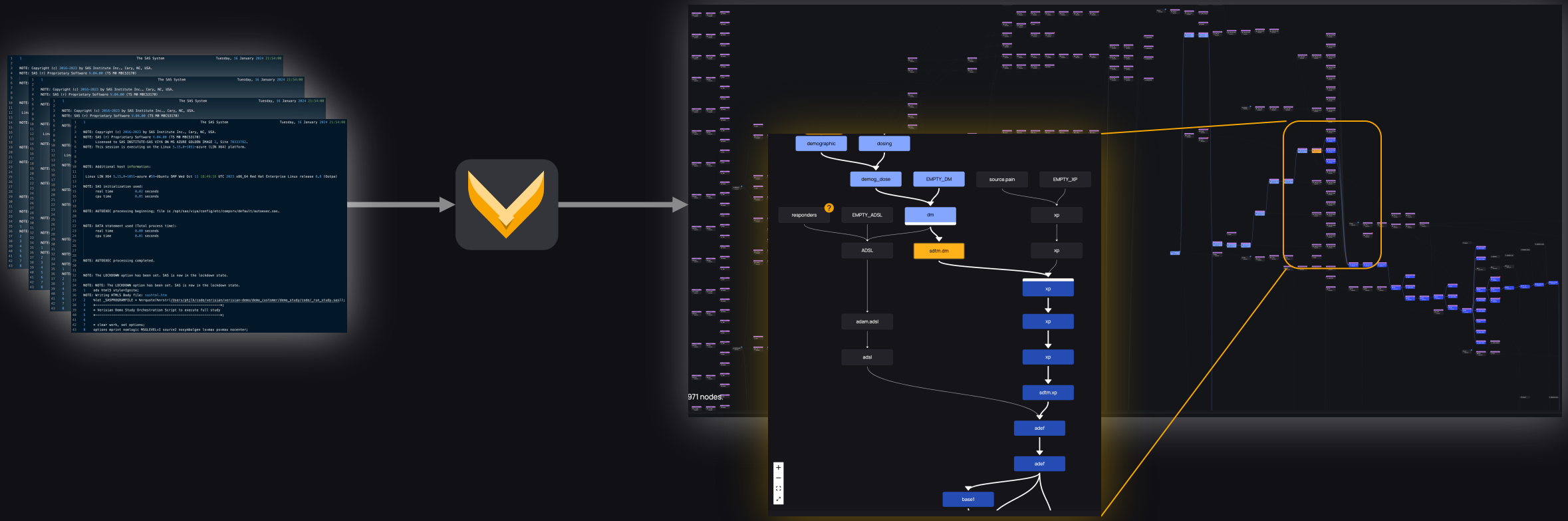
Non-standard, study-specific, scientific innovation

Code traceability

Tracing the content interactions between data, code and results

Independent of standards and languages

All you need are logs...



Search variables...

ADaM

- adsl.arm
- adsl.trt01p
- adsl.trt01pn
- adef.trtpn

SDTM

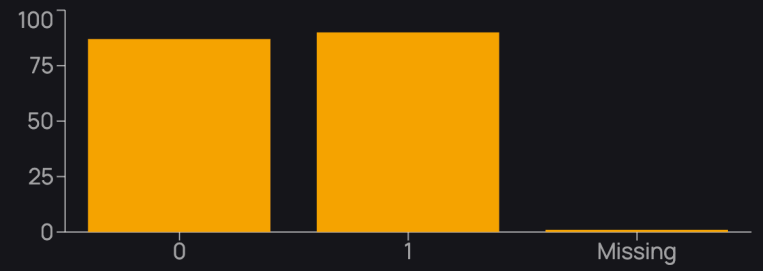
- dm.arm

SOURCE

- demographic.trt

Data | Variable | trtpn

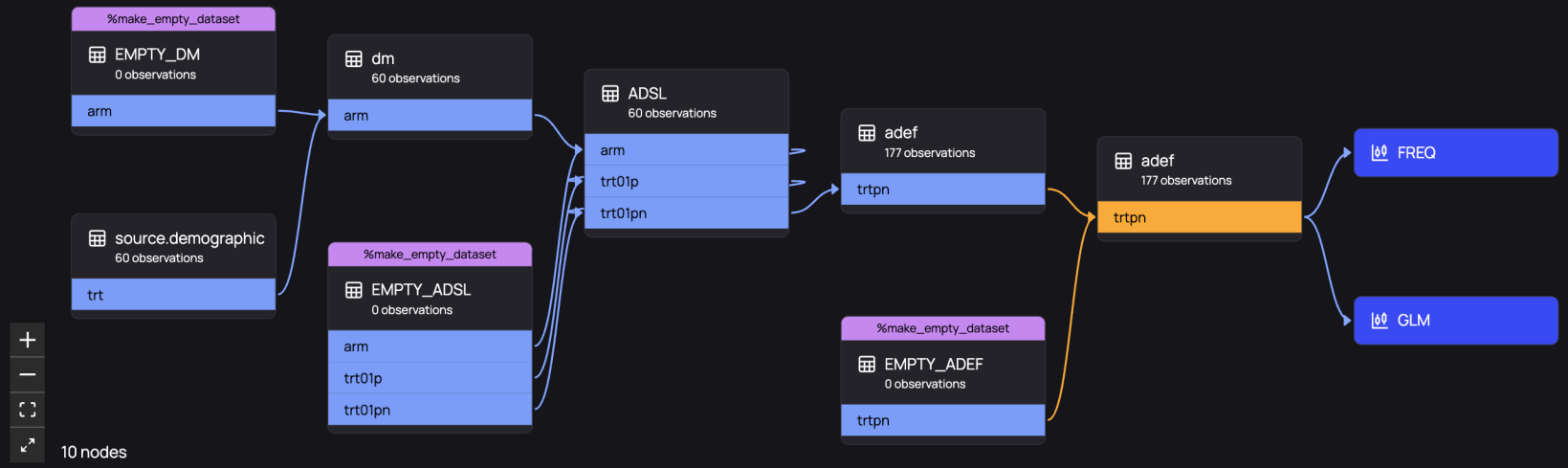
ADaM.adef > trtpn



Variable | Full Derivation

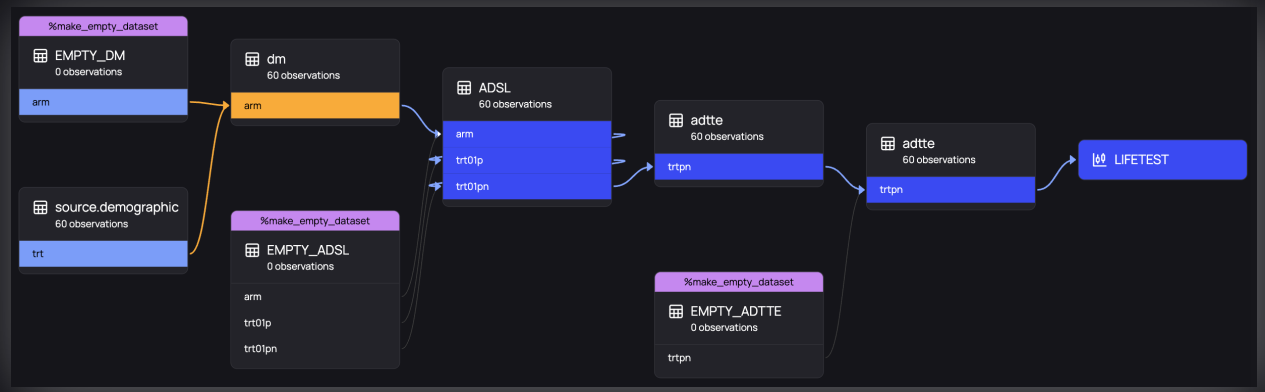
(Macro-Resolved)

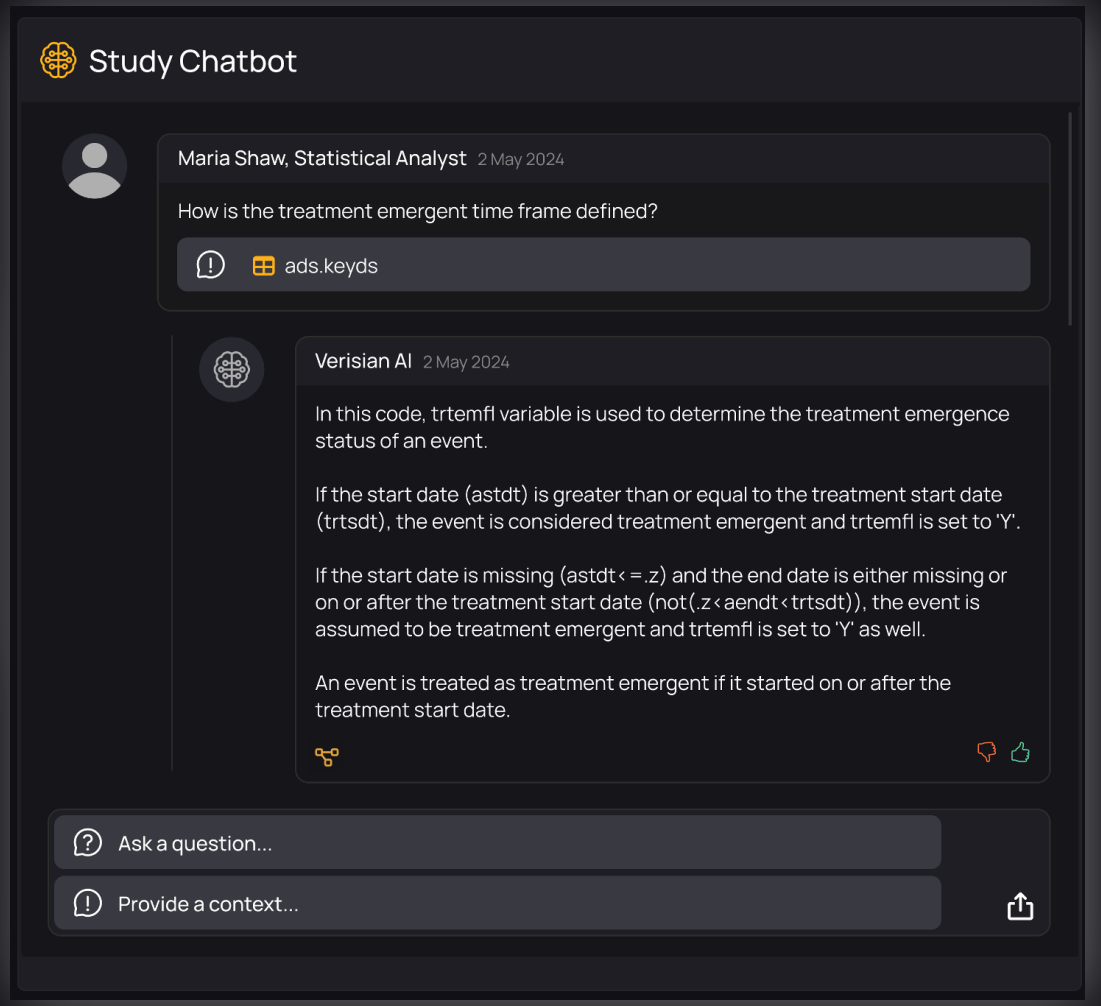
```
63   *...;
64 run;
65
66 * adef.sas;
67 data adef;
68   retain studyid usubjid siteid country randdt age agegr1n a
69   set EMPTY_ADEF adef;
70 run;
71
72 * tlf-responder-rate.sas;
73 proc freq data=adef;
74   *...;
75   tables trtpn * crit1fl / chisq;
```



Traceability-enabled AI

Creating an inclusive and deterministic environment





Given any natural language prompt...

User Question

```
How is time to event
determined in this study?
```

...and translate into natural language output

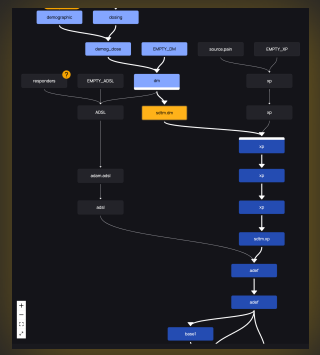
User Question

LLM 1

Clinical concepts

...
...
...
...
...
...
...

Fetch traceability



LLM 2

User Question

Clinical concepts

Traceability info

Output

Natural language

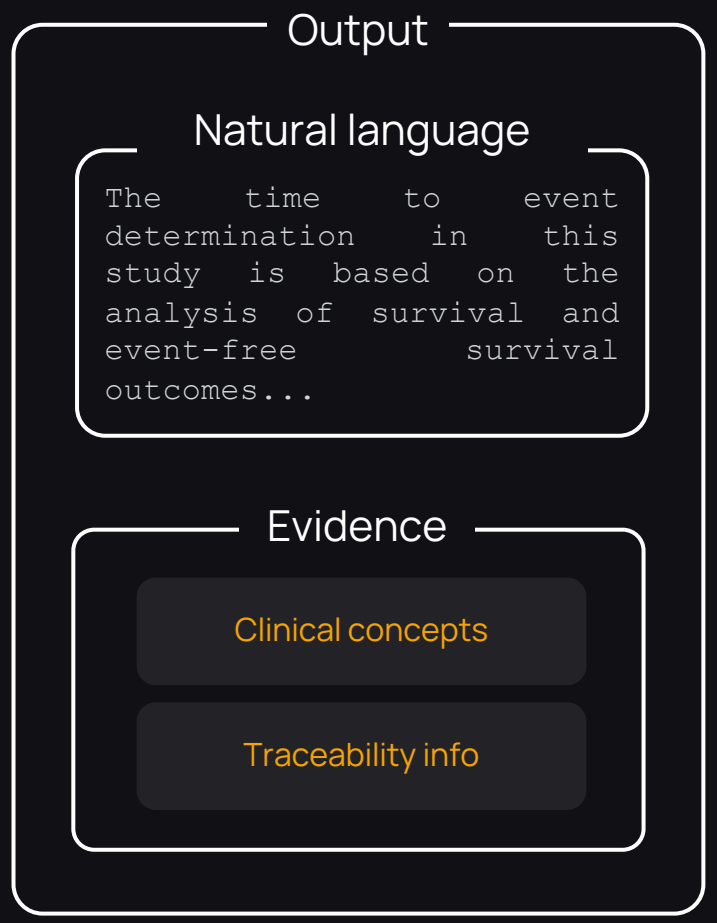
The time to event determination in this study is based on the analysis of survival and event-free survival outcomes...

Evidence

Clinical concepts

Traceability info

...always provide evidence!

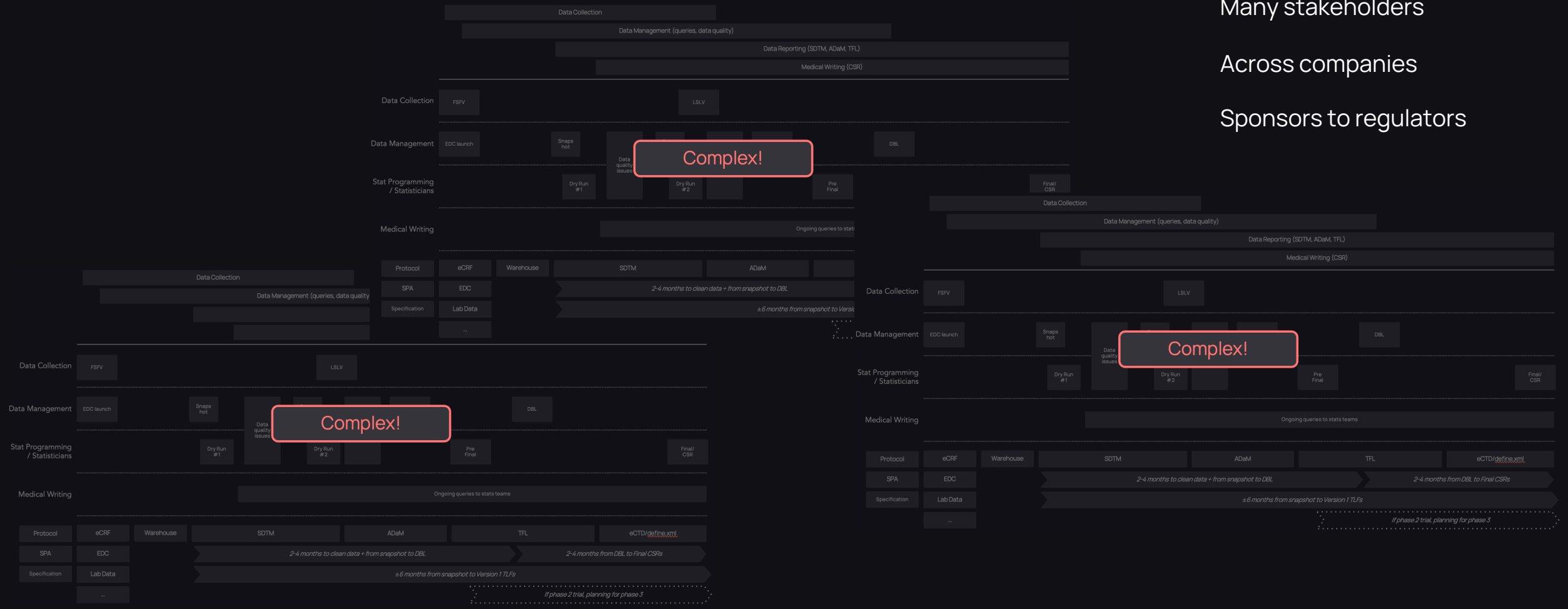


Deterministic	Evidence	Study specific
Translation of Code Traceability	Always provided to review AI outputs	Retrieving directly from a study

Traceability-enabled AI

Reproducible, evidence based

Knowledge transfer in an integrated industry



- Many stakeholders
- Across companies
- Sponsors to regulators

Tracer

Filters (x) sdtm.dm.age

Selection (x) dosing.firstdose

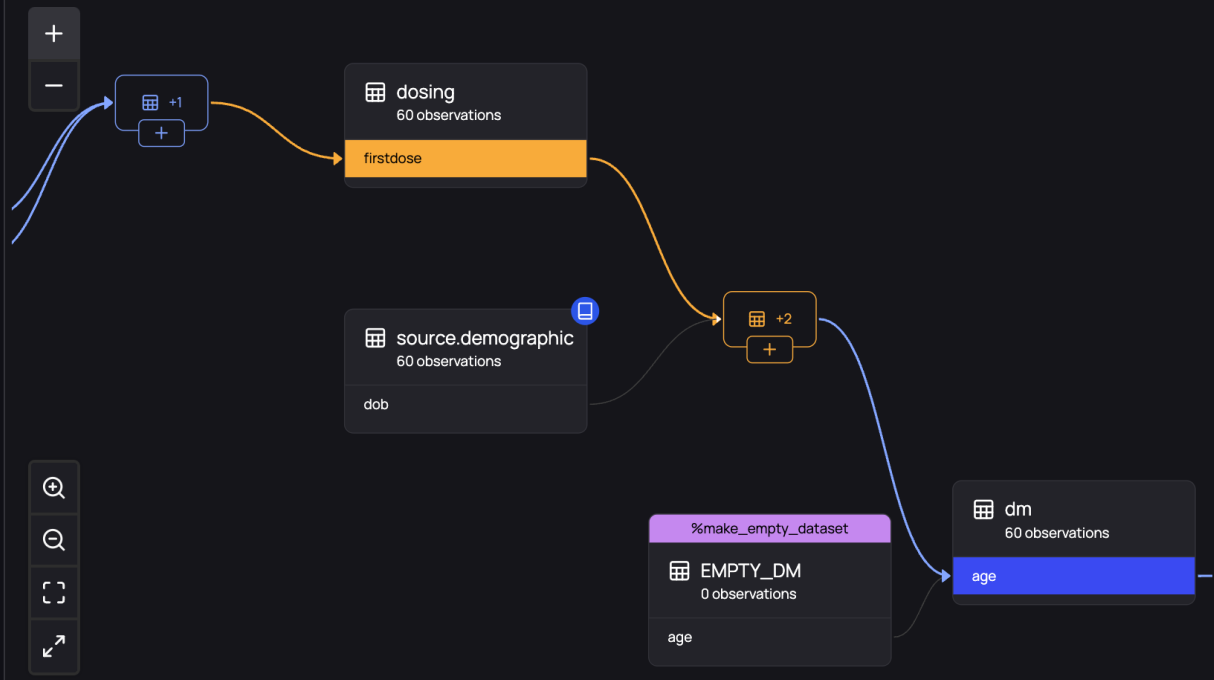
Graph | Resolved | Original | Log

(Macro-Resolved)

```
27        format startdt enddt mmddyy10.;
28        *...;
29  run;
30
31  * DM.sas;
32  data EMPTY_DM;
33      *...;
34    attrib AGE label="Age" length=8. ;
35    retain AGE . ;
36    *...;
37  run;
38
39  * DM.sas;
40  data dosing;
41      *...;
42    by subject;
43    *...;
44    retain firstdose lastdose;
45    *...;
46    if first.subject then do;
47      firstdose=.;
48      *...;
49    firstdose=min(firstdose,startdt,enddt);
50    *...;
51  run;
52
53  * DM.sas;
54  data dm;
55      *...;
56    age=floor ((intck('month',dob,firstdose) - (day(firstdose) < day(dob)
57    if age ne . then ageu='YEARS';
```

Graph | Overview Graph

17 nodes

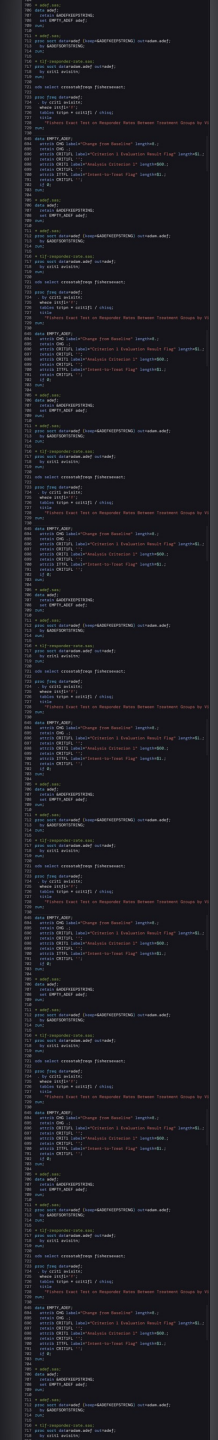


Variable | Datasets

Name
firstdose

Type
unknown

Inferred Values
‹unknown›



From 700+ lines to 37 lines

```sas
* demographic.sas;
data source.demographic;
 . label trt="Treatment";
  input trt 5;
  *...;
run;

* DM.sas;
data dm;
  *...;
  arm=put(trt, arm_demographic_trt.);
 . *...;
run;

* adsl.sas;
data ADSL;
  *..;
  trt01p=ARM;
  *...;
  trt01pn=input(put(trt01p, $trt01pn.), best.);
  *...;
run;

* adef.sas;
data adef;
  * ...;
  rename trt01pn=trtpn;
  *...;
run;

proc freq data=adef;
  by crit1 avisitn;
  where ittfl='Y';
  tables trtpn * crit1fl / chisq;
  title
    "Fishers Exact Test on Responder Rates Between Treatment Groups by Vi
run;

```

Using these techniques can
fundamentally transform
clinical statistical programming for submissions

Reproducible, evidence based

Build, Validate, and Submit Clinical Trial Analyses

Based on Traceability. Empowered by AI.