2024 CDISC + TMF
EUROPE INTERCHANGE BERLIN

24-25 APRIL: CONFERENCE & EXPO | 22, 23, 26 APRIL: TRAININGS

ARD Generator:
A SAS tool for metadata-driven analysis using the
Analysis Result Standard

Karl Wallendszus
Senior Analyst/Programmer, Oxford Population Health

Meet the Speaker

Karl Wallendszus

Title: Senior Analyst/Programmer

Organization: Clinical Trial Service Unit, Oxford Population Health

Karl Wallendszus is a statistical programmer in the Oxford Population Health, a department of the University of Oxford. Over the last 30 years he has worked on a number of large clinical trials in cardiology and kidney disease as well as the RECOVERY trial of treatments for COVID-19. He has been using CDISC standards since 2010, which at the time was unusual for an academic institution. Karl was a volunteer in the CDISC team which developed the new Analysis Results Standard (ARS) and participated in the CDISC ARS Hackathon the summer of 2023.

Disclaimer and Disclosures

- *The views and opinions expressed in this presentation are those of the author(s) and do not necessarily reflect the official policy or position of CDISC.*

- *The author(s) have no real or apparent conflicts of interest to report.*
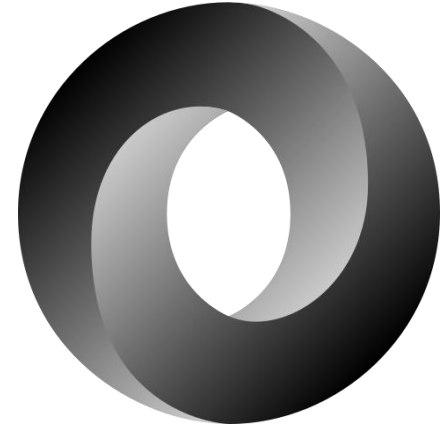
cdisc

Agenda

The Analysis Result Standard

ARS and JSON

We use JSON as a representation of the ARS.

But JSON is not part of that standard.

And other representations of ARS are possible.

CDISC Foundational Standards

Data Collection
CDASH

Data Aggregation
SDTM

Analysis
ADaM

Results
???

ARM for Define.XML

ARS: Separating results from displays

Analysis Results that Save Trees - #KillTFLs

Chris Decker, d-wise, Morrisville, NC, USA

ABSTRACT

In today's clinical programming environment, we spend large amounts of time (and money) producing Tables, Figures, and Listings – single use, throw-away artifacts that I would challenge focus more on the cosmetics then the content. Artifacts that, by their own admission, are rarely looked at by regulatory reviewers. We build these reports for one narrowly focused outcome instead of making that result, and many more, available for many different uses. What if we could capture and store analysis results that gave the results context and allowed a clinician or reviewer to dynamically search the results within and across studies without sifting through 300 pages of paper? We can do this! This presentation will provide an overview of a project where a graph-based analysis results data store was designed and populated to capture and make analysis results available to many different stakeholders reducing their workload by 50%. We can leverage new technology and/or methodologies to optimize our archaic processes and #killTFLs.

https://phuse.s3.eu-central-1.amazonaws.com/Archive/2019/Connect/US/Baltimore/PAP_DH12.pdf

ARS Model Supported Workflow and Entry Points

ARS model
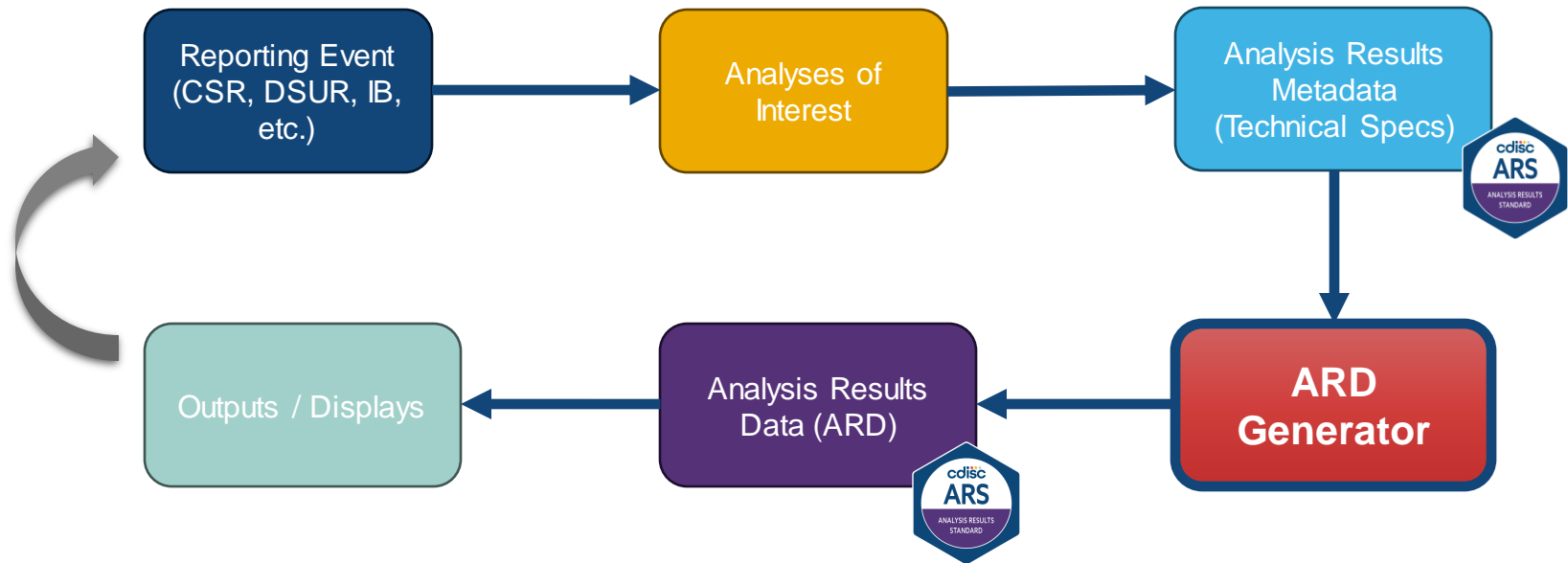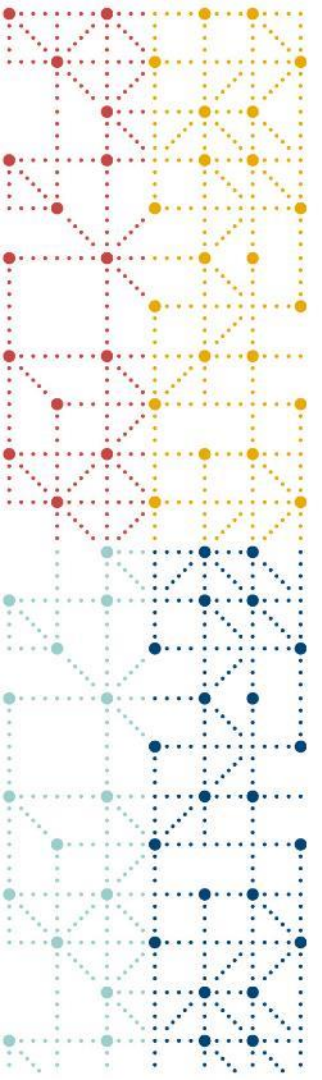
The ARD Generator tool

ARD Generator

- Written in SAS

- Ingests ARS metadata from JSON → SAS datasets
- Processes analyses specified in that metadata
- Uses a macro library to automate analysis
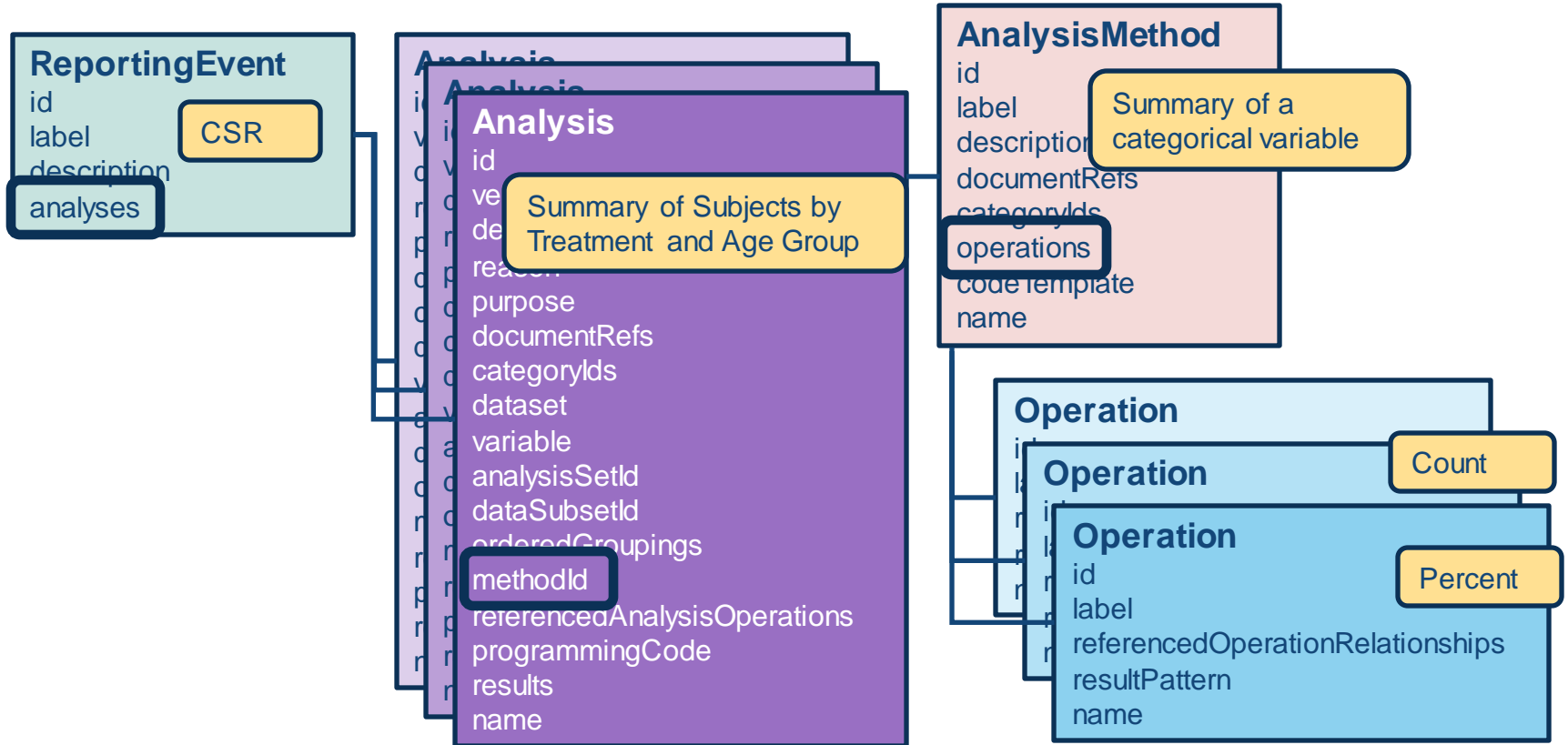- Output to Analysis Result Dataset (ARD)

Ingesting JSON

```json
"analyses": [
  {
    "id": "An03_02_AgeGrp_Comp_ByTrt",
    "reason": {
      "controlledTerm": "SPECIFIED IN SAP"
    },
    "purpose": {
      "controlledTerm": "PRIMARY OUTCOME MEASURE"
    },
    "methodId": "Mth03_CatVar_Comp_PChiSq",
    "version": 1,
    "categoryIds": [
      "Catn_01_Grp_1_Pop",
      "Catn_02_Dclass_1_Sbj",
      "Catn_03_SbjDType_1_Dm"
    ],
    "documentRefs": [
      {
        "referenceDocumentId": "CDISCPILOT01_SAP",
        "pageRefs": [
          {
            "refType": "PhysicalRef",
            "label": "9.2. Demographic and Baseline Characteristics",
            "pageNumbers": [
              13
            ]
          }
        ]
      }
    ],
    "analysisSetId": "AnalysisSet_02_SAF",
```

VIEWTABLE: Jsonmd.Analyses

	id	version	name
1	An01_05_SAF_Summ_ByTrt	1	Summary of Subjects by Treatment
2	An03_01_Age_Summ_ByTrt	1	Summary of Age by Treatment
3	An03_01_Age_Comp_ByTrt	1	Comparison of Age by Treatment
4	An03_02_AgeGrp_Summ_ByTrt	1	Summary of Subjects by Treatment and Age Group
5	An03_02_AgeGrp_Comp_ByTrt	1	Comparison of Age Group by Treatment
6	An03_03_Sex_Summ_ByTrt	1	Summary of Subjects by Treatment and Sex
7	An03_03_Sex_Comp_ByTrt	1	Comparison of Sex by Treatment
8	An03_04_Ethnic_Summ_ByTrt	1	Summary of Subjects by Treatment and Ethnicity
9	An03_04_Ethnic_Comp_ByTrt	1	Comparison of Ethnicity by Treatment
10	An03_05_Race_Summ_ByTrt	1	Summary of Subjects by Treatment and Race
11	An03_05_Race_Comp_ByTrt	1	Comparison of Race by Treatment
12	An03_06_Height_Summ_ByTrt	1	Summary of Height by Treatment
13	An03_06_Height_Comp_ByTrt	1	Comparison of Height by Treatment
14	An07_01_TEAE_Summ_ByTrt	1	Summary of Subjects with At Least One TEAE, by Treatment
15	An07_01_TEAE_Comp_ByTrt_PlacLow	1	Comparison of Subjects with TEAEs by Treatment - Placebo vs Low Dose
16	An07_01_TEAE_Comp_ByTrt_PlacHigh	1	Comparison of Subjects with TEAEs by Treatment - Placebo vs High Dose
17	An07_02_RelTEAE_Summ_ByTrt	1	Summary of Subjects with At Least One Related TEAE, by Treatment
18	An07_03_SerTEAE_Summ_ByTrt	1	Summary of Subjects with At Least One Serious TEAE, by Treatment
19	An07_04_RelSerTEAE_Summ_ByTrt	1	Summary of Subjects with At Least One Related Serious TEAE, by Treatment
20	An07_05_TEAELd2Dth_Summ_ByTrt	1	Summary of Subjects with At Least One TEAE Leading to Death, by Treatment
21	An07_06_RelTEAELd2Dth_Summ_ByTrt	1	Summary of Subjects with At Least One Related TEAE Leading to Death, by Treatment
22	An07_07_TEAELd2DoseMod_Summ_By	1	Summary of Subjects with At Least One TEAE Leading to Dose Modification, by Treatment

How analyses are modelled

ReportingEvent
id
label
description
analyses

CSR

Analysis
Analysis

Analysis
id
version
description
reason
purpose
documentRefs
categoryIds
dataset
variable
analysisSetId
dataSubsetId
orderedGroupings
methodId
referencedAnalysisOperations
programmingCode
results
name

Summary of Subjects by Treatment and Age Group

AnalysisMethod
id
label
description
documentRefs
categoryIds
operations
codeTemplate
name

Summary of a categorical variable

Operation
Operation

Count

Operation
id
label
referencedOperationRelationships
resultPattern
name

Percent

A macro library allows automation

ard-generator / macros / source /

KarlWallendszus Added operations for chis

Name
📁 ..
📄 append_addcols.sas
📄 build_expression.sas
📄 build_work_dataset.sas
📄 define_analset.sas
📄 op_catvar_comp_pchisq.sas
📄 op_catvar_count_bygrp_n.sas
📄 op_catvar_summ_bygrp_pct.sas
📄 op_contvar_summ_bygrp.sas
📄 outline_ard.sas
📄 run_analysis.sas
📄 run_method.sas
📄 run_operation.sas
📄 standardize_grouping.sas
📄 standardize_groupings.sas

```sas
* Loop through analyses;
%local ianal analid;
%let ianal = 1;
%do %while(%scan(&analids., &ianal., '|') ne );
    %let analid = %scan(&analids., &ianal., '|');

    * Run this analysis;
    %run_analysis(mdlib=&mdlib., datalib=&datalib., ardlib=&ardlib.,
        analid=&analid., debugfl=&debugfl.);
```

```sas
* Run the analysis method;
%run_method(mdlib=&mdlib., datalib=&datalib., ardlib=&ardlib.,
```

```sas
%* Loop through analysis operations;
%do iop = 1 %to &noperations.;

    %* Execute this operation;
    %run_operation(mdlib=&mdlib., datalib=&datalib.,
        opid=&&opid&iop., opseq=&iop., nop=&noperations., methid=&methid.,
        analid=&analid., analsetid=&analsetid., datasubsetid=&datasubsetid.,
        groupingids=&groupingids., analds=workds, analvar=&analvar.,
        ard=&ardlib..ard, debugfl=&debugfl.);

%end;
```

cdisc

Operation macros

KarlWallendszus Added operations for chisq a

Name

..

append_addcols.sas

build_expression.sas

build_work_dataset.sas

define_analset.sas

op_catvar_comp_pchisq.sas

op_catvar_count_bygrp_n.sas

op_catvar_summ_bygrp_pct.sas

op_contvar_summ_bygrp.sas

outline_ard.sas

run_analysis.sas

run_method.sas

run_operation.sas

standardize_grouping.sas

standardize_groupings.sas

```sas
proc freq data = &dsin.;
    tables &tabreq. / chisq;
    output out = rawres chisq pchi;
run;
```

VIEWTABLE: Work.Rawres

	N	_PCHI_	DF_PCHI	P_PCHI
1	254	1.7166152	2	0.4238788

```sas
%* Set the result column to use;
%local rescol;
%if %bquote(&oplabel.) = chisq %then %let rescol = _PCHI_;
%else %if %bquote(&oplabel.) = df %then %let rescol = DF_PCHI;
%else %if %bquote(&oplabel.) = %str(p-value) %then %let rescol = P_PCHI;
```

cdisc
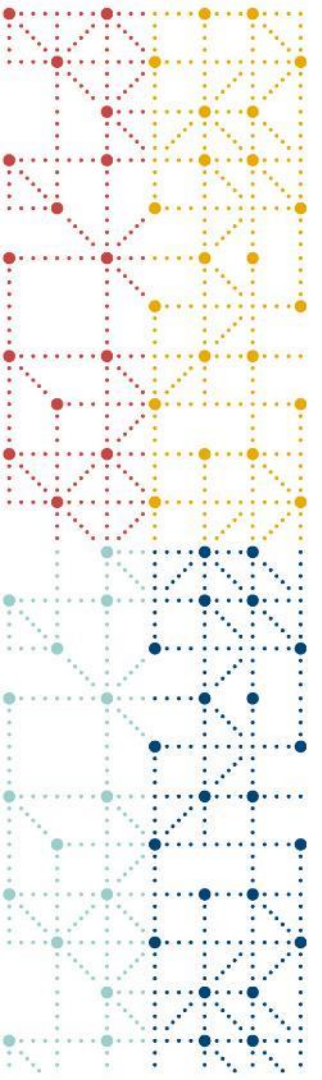
Usage example

Comparison of Age Group by Treatment

Comparison of Age Group by Treatment

Variable	Value
Id	An03_02_AgeGrp_Comp_ByTrt
Name	Comparison of Age Group by Treatment
Reason	SPECIFIED IN SAP
Purpose	PRIMARY OUTCOME MEASURE
analysisSetId	AnalysisSet_02_SAF
groupingId1	AnlsGrouping_01_Trt
resultsByGroup1	FALSE
groupingId2	AnlsGrouping_03_AgeGp
resultsByGroup2	FALSE
Dataset	ADSL
Variable	USUBJID
Method_id	Mth03_CatVar_Comp_PChiSq

Analysis sets

VIEWTABLE: Jsonmd.Analysissets

	id	label	compndExpression_logi	condition_dataset	condition_variable	condition_comparator	condition_value	level	order
1	AnalysisSet_01_ITT	Intent-to-Treat Population		ADSL	ITTFL	EQ	Y	1	1
2	AnalysisSet_02_SAF	Safety Population		ADSL	SAFFL	EQ	Y	1	1

```
* Derive expressions in SAS syntax for conditions;
%build_expression_all(dsin=jsonmd.analysissets, datalib=adam,
    dsout=jsonmd.expressions);
%build_expression_all(dsin=jsonmd.datasubsets, datalib=adam,
    dsout=jsonmd.expressions);
%build_expression_all(dsin=jsonmd.analysisgroupings, idvar=group_id,
    labelvar=group_label, levelvar=group_level,
    ordervar=order, logopvar=group_logicalOperator,
    dsvar=group_condition_dataset, varvar=group_condition_variable,
    compvar=group_condition_comparator, valvar=condition_value,
```

VIEWTABLE: Jsonmd.Expressions

	id	label	dsconds	expression
1	AnalysisSet_01_ITT	Intent-to-Treat Population	ADSL	ITTFL EQ "Y"
2	AnalysisSet_02_SAF	Safety Population	ADSL	SAFFL EQ "Y"
3	D_01_TEAE	Treatment Emergent Adverse	ADAE	TRTEMFL EQ "Y"

cdisc

Analysis groupings

VIEWTABLE: Jsonmd.Analysisgroupings

|	grouping_type	id	label	groupingVaria	dataDriven	group_id	group_label	group_logic	group_conditio	group_condi	group_cond	condition_value	gr
1	Subject	AnlsGrouping_01_Trt	Treatment	TRT01A	FALSE	AnlsGrouping_01_Trt_1	Placebo		ADSL	TRT01A	EQ	Placebo	1
2	Subject	AnlsGrouping_01_Trt	Treatment	TRT01A	FALSE	AnlsGrouping_01_Trt_2	Xanomeline Low Dose		ADSL	TRT01A	EQ	Xanomeline Low Dose	1
3	Subject	AnlsGrouping_01_Trt	Treatment	TRT01A	FALSE	AnlsGrouping_01_Trt_3	Xanomeline High Dose		ADSL	TRT01A	EQ	Xanomeline High Dose	1
4	Subject	AnlsGrouping_02_Sex	Gender	SEX	FALSE	AnlsGrouping_02_Sex_1	Male		ADSL	SEX	EQ	M	1
5	Subject	AnlsGrouping_02_Sex	Gender	SEX	FALSE	AnlsGrouping_02_Sex_2	Female		ADSL	SEX	EQ	F	1
6	Subject	AnlsGrouping_03_AgeGp	Age Group	AGEGR1	FALSE	AnlsGrouping_03_AgeGp_1	< 65 years		ADSL	AGEGR1	EQ	<65	1
7	Subject	AnlsGrouping_03_AgeGp	Age Group	AGEGR1	FALSE	AnlsGrouping_03_AgeGp_2	≥ 65 years		ADSL	AGEGR1	IN	65-80	>80	1
8	Subject	AnlsGrouping_04_Race	Primary Race	RACE	FALSE	AnlsGrouping_04_Race_1	American Indian or Alaska Native		ADSL	RACE	EQ	AMERICAN INDIAN OR ALASKA NATIVE	1
9	Subject	AnlsGrouping_04_Race	Primary Race	RACE	FALSE	AnlsGrouping_04_Race_2	Asian		ADSL	RACE	EQ	ASIAN	1
10	Subject	AnlsGrouping_04_Race	Primary Race	RACE	FALSE	AnlsGrouping_04_Race_3	Black or African American		ADSL	RACE	EQ	BLACK OR AFRICAN AMERICAN	1
11	Subject	AnlsGrouping_04_Race	Primary Race	RACE	FALSE	AnlsGrouping_04_Race_4	Native Hawaiian or Other Pacific Islander		ADSL	RACE	EQ	NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER	1
12	Subject	AnlsGrouping_04_Race	Primary Race	RACE	FALSE	AnlsGrouping_04_Race_5	White		ADSL	RACE	EQ	WHITE	1
13	Subject	AnlsGrouping_04_Race	Primary Race	RACE	FALSE	AnlsGrouping_04_Race_6	Multiple		ADSL	RACE	EQ	MULTIPLE	1
14	Subject	AnlsGrouping_04_Race	Primary Race	RACE	FALSE	AnlsGrouping_04_Race_7	Not Reported		ADSL	RACE	EQ	NOT REPORTED	1
15	Subject	AnlsGrouping_04_Race	Primary Race	RACE	FALSE	AnlsGrouping_04_Race_8	Unknown		ADSL	RACE	EQ	UNKNOWN	1
16	Subject	AnlsGrouping_04_Race	Primary Race	RACE	FALSE	AnlsGrouping_04_Race_9	Other		ADSL	RACE	EQ	OTHER	1
17	Subject	AnlsGrouping_05_Ethnic	Ethnicity	ETHNIC	FALSE	AnlsGrouping_05_Ethnic_1	Hispanic or Latino		ADSL	ETHNIC	EQ	HISPANIC OR LATINO	1
18	Subject	AnlsGrouping_05_Ethnic	Ethnicity	ETHNIC	FALSE	AnlsGrouping_05_Ethnic_2	Not Hispanic or Latino		ADSL	ETHNIC	EQ	NOT HISPANIC OR LATINO	1
19	Data	AnlsGrouping_06_Soc	System Organ Class	AESOC	TRUE	AnlsGrouping_06_Soc	System Organ Class						
20	Data	AnlsGrouping_07_Pt	Preferred Term	AEDECOD	TRUE	AnlsGrouping_07_Pt	Preferred Term						
21	Data	AnlsGrouping_08_Param	Parameter	PARAMCD	FALSE	AnlsGrouping_08_Param_1	Systolic Blood Pressure (mmHg)		ADVS	PARAMCD	EQ	SYSBP	1
22	Data	AnlsGrouping_08_Param	Parameter	PARAMCD	FALSE	AnlsGrouping_08_Param_2	Diastolic Blood Pressure (mmHg)		ADVS	PARAMCD	EQ	DIABP	1
23	Data	AnlsGrouping_08_Param	Parameter	PARAMCD	FALSE	AnlsGrouping_08_Param_3	Pulse Rate (beats/min)		ADVS	PARAMCD	EQ	PULSE	1
24	Data	AnlsGrouping_08_Param	Parameter	PARAMCD	FALSE	AnlsGrouping_08_Param_4	Temperature (C)		ADVS	PARAMCD	EQ	TEMP	1

Output in ARD

VIEWTABLE: Ard.Ard

	analysisId	analysisName	analysisSetId	analysisSetLabel	dataSubsetId	dataSubsetLabel	analysisD	analysisVariable	
29	An03_02_AgeGrp_Comp_ByTrt	Comparison of Age Group by Treatment	AnalysisSet_02_SAF	Safety Population			ADSL	USUBJID	
30	An03_02_AgeGrp_Comp_ByTrt	Comparison of Age Group by Treatment	AnalysisSet_02_SAF	Safety Population			ADSL	USUBJID	
31	An03_02_AgeGrp_Comp_ByTrt	Comparison of Age Group by Treatment	AnalysisSet_02_SAF	Safety Population			ADSL	USUBJID	
32	An03_02_AgeGrp_Summ_ByTrt	Summary of Subjects by Treatment and Age...	AnalysisSet_02_SAF	Safety Population			ADSL	USUBJID	

VIEWTABLE: Ard.Ard

	methodId	methodName	operationId	operationName	
29	Mth03_CatVar_Comp_PChiSq	Pearson's chi-square test group comparison for a categorical variable	Mth03_CatVar_Comp_PChiSq_1_chisq	Chi-squared	A
30	Mth03_CatVar_Comp_PChiSq	Pearson's chi-square test group comparison for a categorical variable	Mth03_CatVar_Comp_PChiSq_2_df	Degrees of freedom	A
31	Mth03_CatVar_Comp_PChiSq	Pearson's chi-square test group comparison for a categorical variable	Mth03_CatVar_Comp_PChiSq_3_pval	P-value	A

VIEWTABLE: Ard.Ard

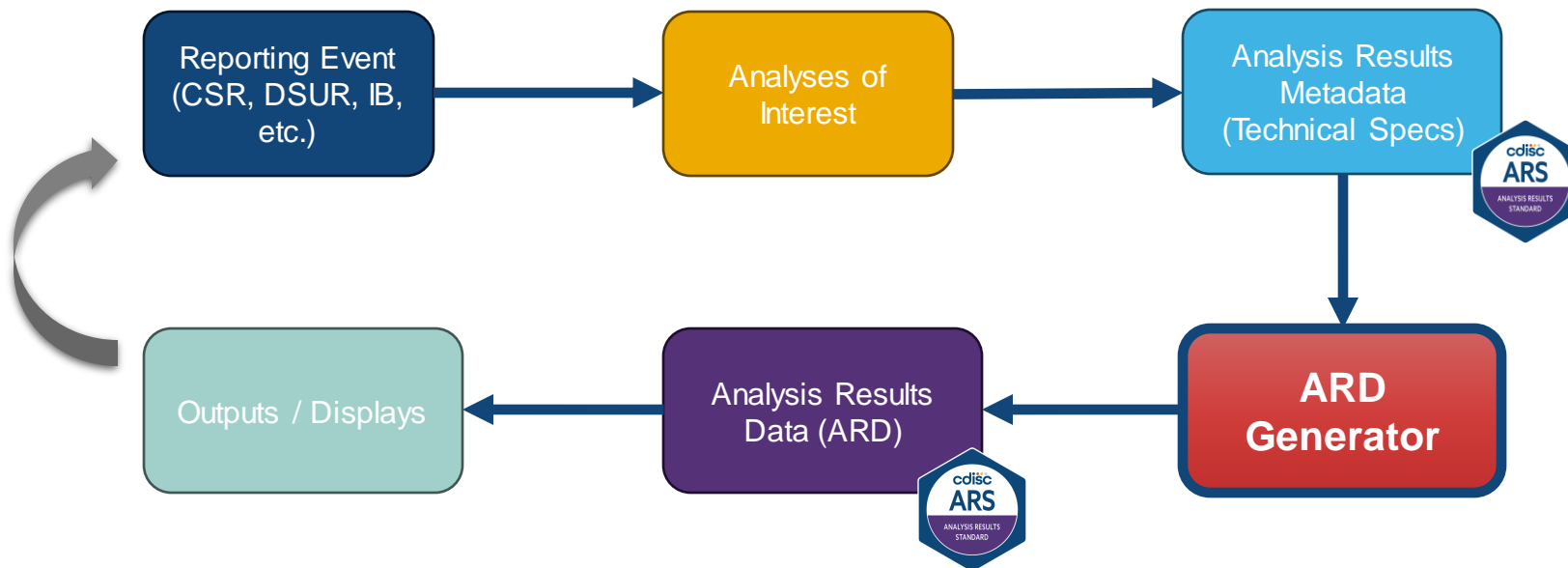|	resultGroup1_groupingId	resultGroup2_groupingId	rawValue
29	AnlsGrouping_01_Trt	AnlsGrouping_03_AgeGp	1.7166151981
30	AnlsGrouping_01_Trt	AnlsGrouping_03_AgeGp	2
31	AnlsGrouping_01_Trt	AnlsGrouping_03_AgeGp	0.4238788486
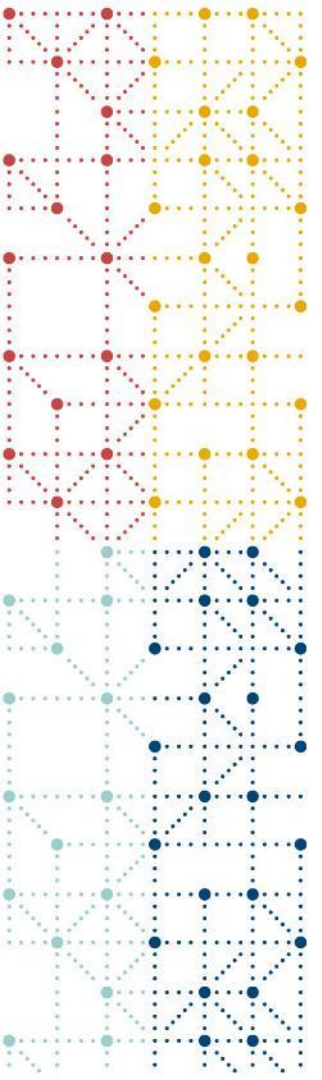
# Conclusions

Existing workflow
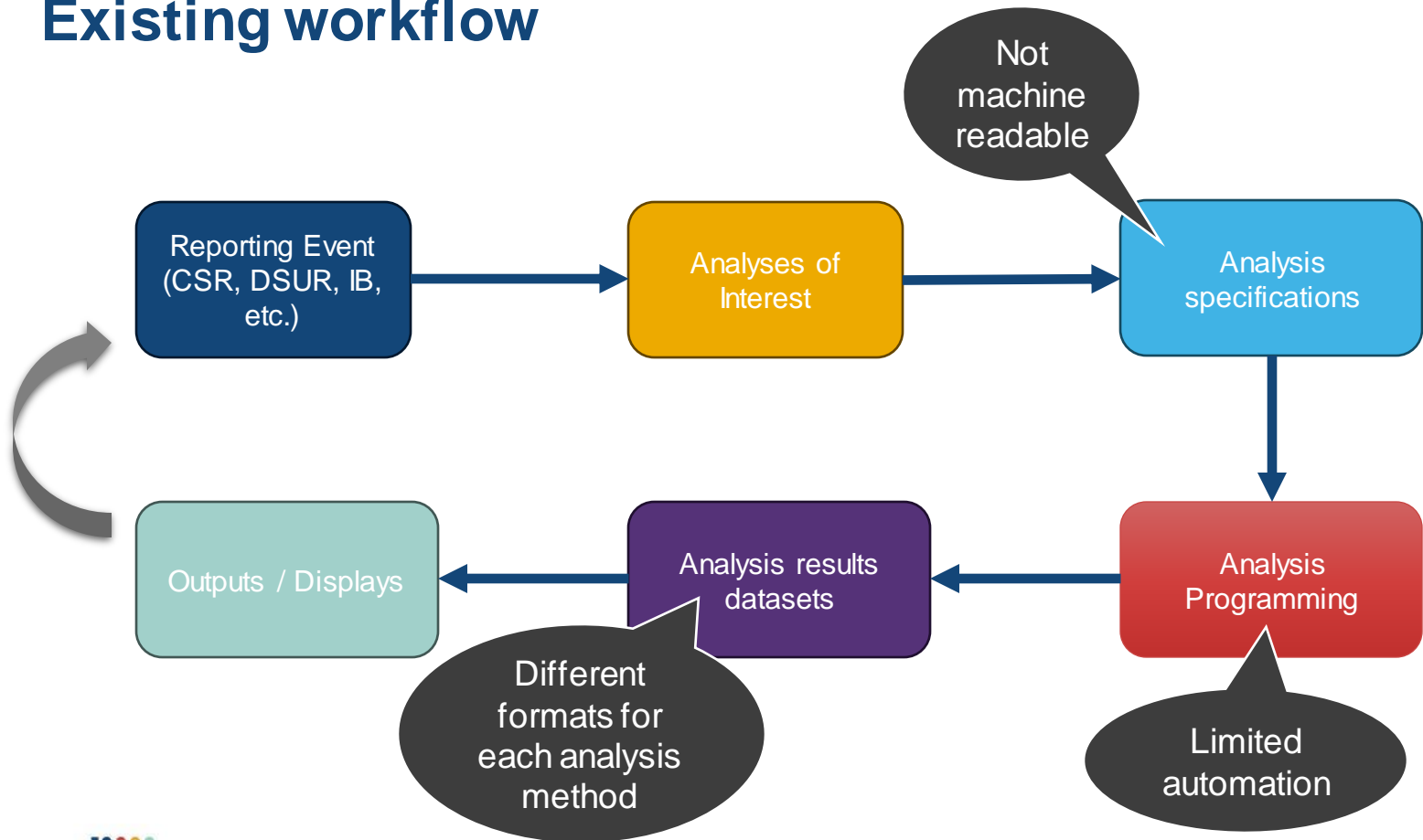
Future workflow

Future development

- Handle more analysis operations and methods.
- Implement display patterns.
- Output ARD as JSON
- Complement ARD Generator with other tools

Summary

- The Analysis Results Standard (ARS) extends the coverage of CDISC standards to the representation of analysis results in their full context.

- Our SAS tool, ARD Generator, makes use of the ARS to automate the generation of results.

- The tool ingests ARM in JSON format and maps its deep hierarchical structure to SAS datasets. It then uses a macro library to perform the specified analyses and writes results to an ARS-compliant Analysis Result Dataset (ARD).

- The tool is highly modular with much re-use of code. It largely automates the analysis process, as once the relevant macros exist, all the information needed to apply them can be taken from the ARM.

- ARD Generator came out of work for a CDISC Hackathon.
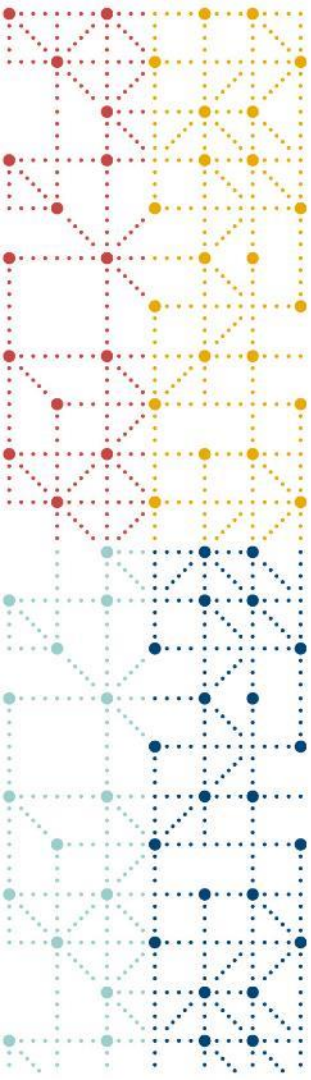
Thank You!

Karl Wallendszus
Clinical Trial Service Unit, Oxford Population Health
University of Oxford, UK
karl.wallendszus@ndph.ox.ac.uk

https://github.com/KarlWallendszus/ard-generator