

## 1. Background

Standards play a pivotal role in ensuring **quality and consistency** in clinical data, with Define-XML for ADaM providing a standardized format for describing the structure and content of the analysis datasets.

The Origin element specifies the source for the variable, aiding transparency and traceability in analysis and regulatory submissions. Having observed inconsistencies within our own analysis standards in how Origin had been assigned, we conducted a survey revealing varied understanding and implementation practices across our Biostatistics department.

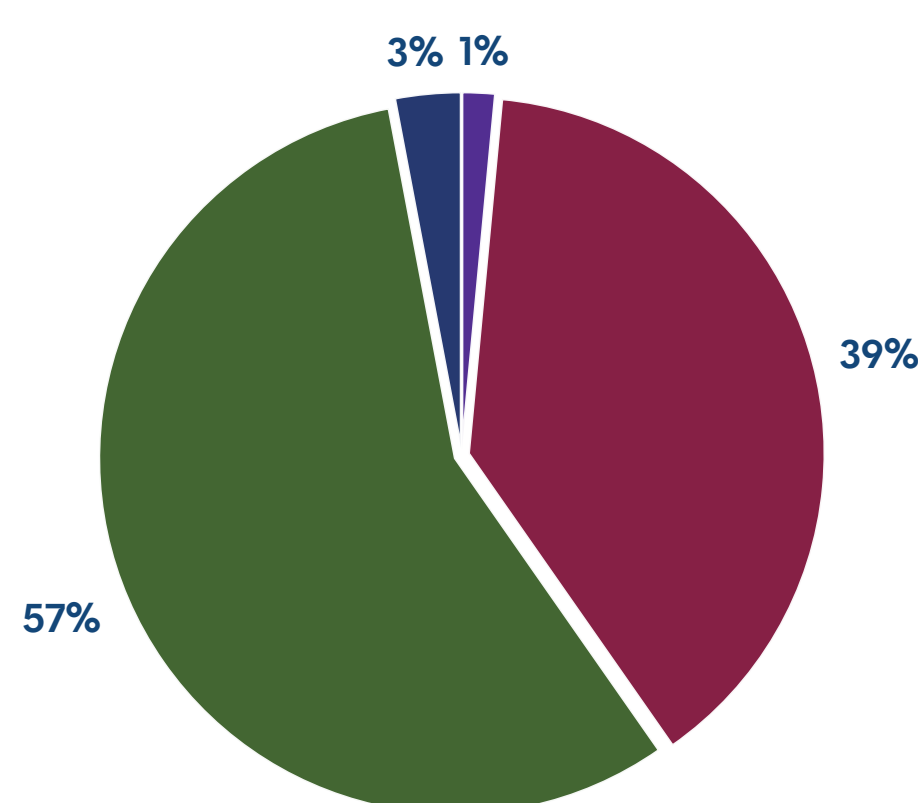
This poster presents a **solution to streamline Origin assignment**, with the aim of improving ADaM specification consistency. To facilitate the preferred application of Origin within our organization, we have created a flowchart to support both standards creators and users.

## 2. Survey

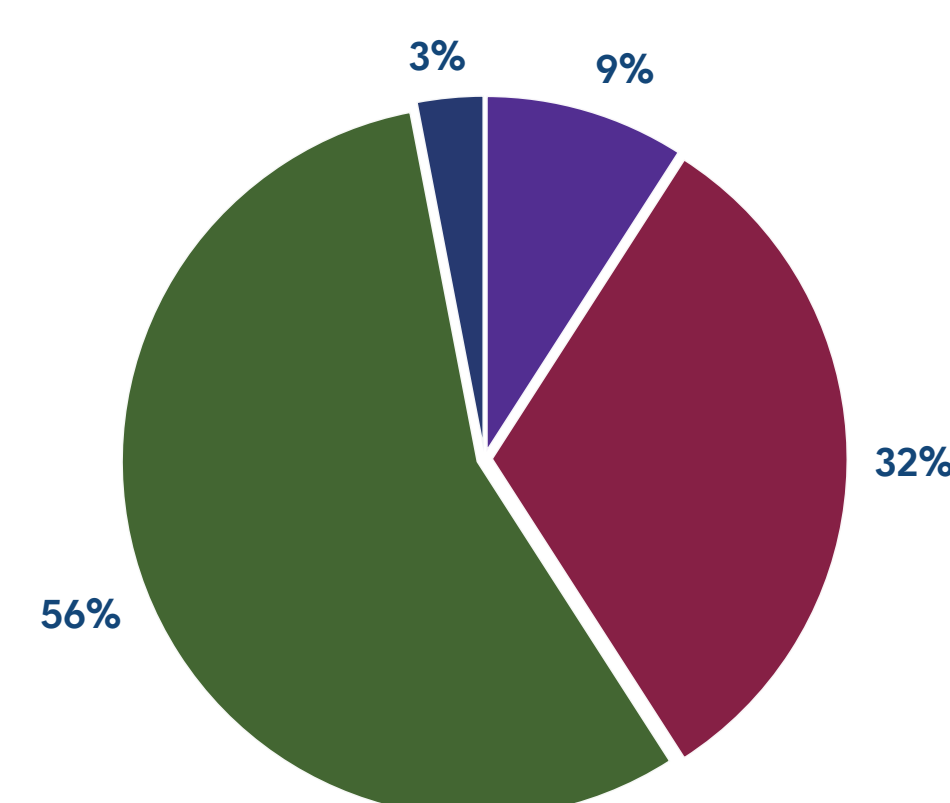
We released the survey within our Biostatistics team, covering a selection of scenarios a typical user might encounter. There were a total of 68 responders across several Therapeutic Areas, incorporating a range of experience from new starters to study leads.

A selection of the results are presented here, clearly showing a variation in understanding:

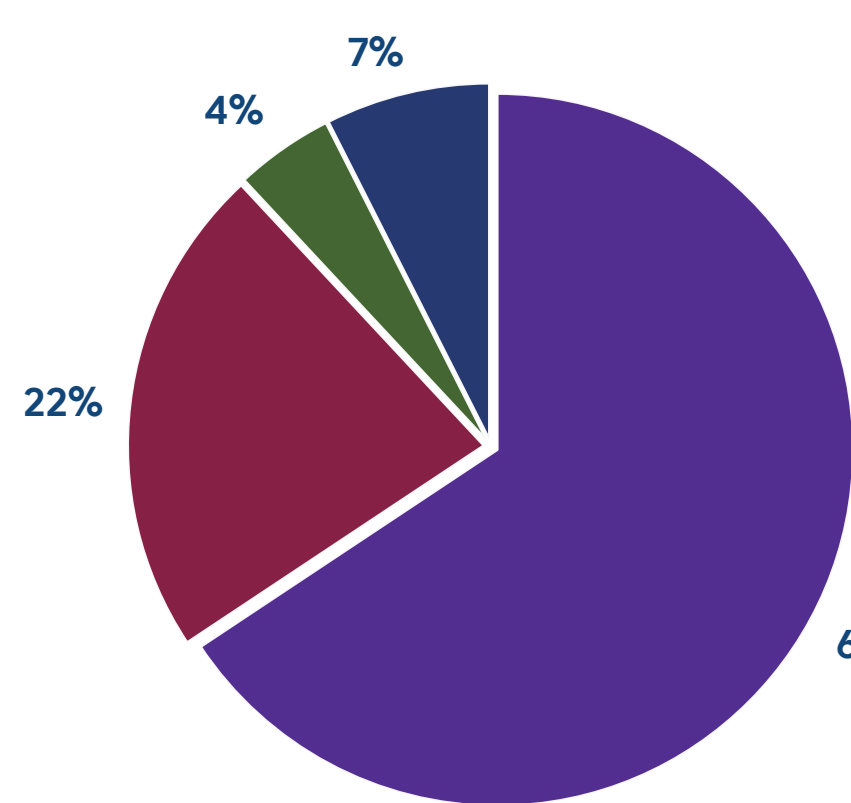
Q1. ADSL.BRTHDTF:  
BRTHDTF = Y if the year is imputed. BRTHDTF = M if year is present and month is imputed. BRTHDTF = D if only day



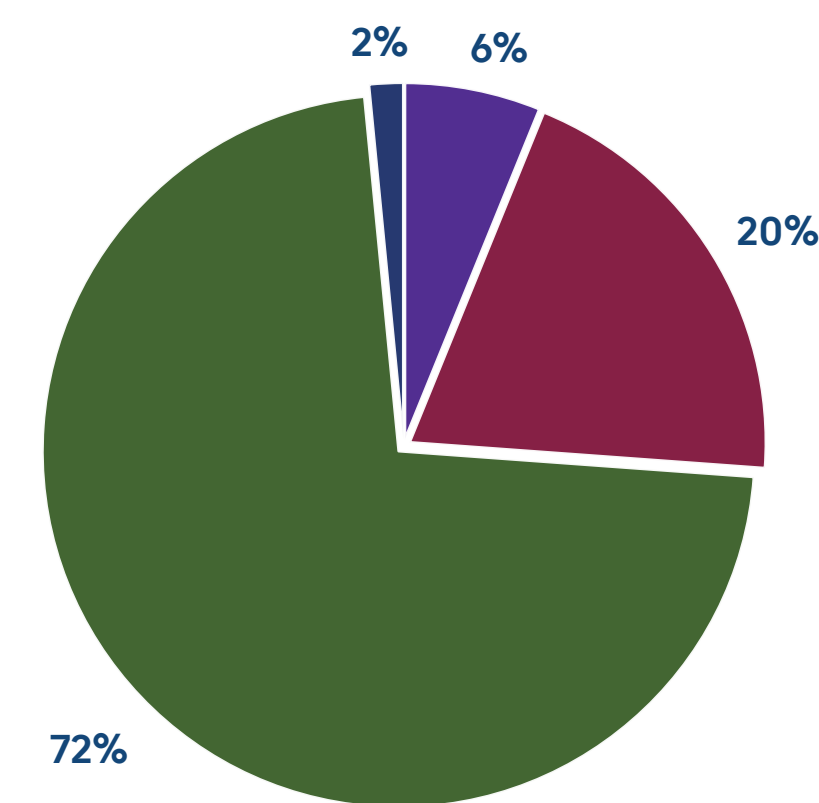
Q2. ADDS.AVALC (PARAMCD EQ SJSTAT):  
DS.DSDECOD where DSSCAT = "SUBJECT CONTINUATION" and DSDECOD = "CONTINUED". Otherwise "WITHDRAWN".



Q3. ADAE.AESPROT  
SUPPAE.QVAL where QNAM = "AESPROT"



Q4. ADVS.BTOXGR  
ATOXGR when ABLFL="Y". One value per Subject and Parameter. Populate as a constant across all rows within that Subject and Parameter.



■ PREDECESSOR ■ ASSIGNED ■ DERIVED ■ NOT SURE

## 3. Take Part!

We **invite you** to contribute to the ongoing discussion on improving the quality and reliability of analysis standards. Please take part in this industry wide survey!



## 4. Flowchart

The flowchart serves as a comprehensive guide for both the Analysis Standards Team and study programmers, simplifying the process of assigning Origin and promoting consistency and clarity.

### Step 1

- Is the variable copied from another SDTM or ADaM dataset?  
E.g. ADSL.SEX = DM.SEX, ADAE.SAFFL = ADSL.SAFFL
  - Is the variable copied from another SDTM or ADaM dataset with a different name?  
E.g. ADVS.AVISIT = VS.VISIT
  - Is the variable a subset of records copied from another SDTM or ADaM dataset?  
E.g. ADAE.AETRTEM = SUPPAE.QVAL where QNAM = "AETRTEM"  
E.g. ADSL.HEIGHTBL = VS.VSSTRESN where VS.VSTESTCD = "HEIGHT" and VS.VISITNUM = xxx
- Note, copied means the values remain unchanged

If YES to any

PREDECESSOR

If NO

### Step 2

- Is the variable populated from an external dictionary?  
E.g. ADAE.SMQzzNAM
- Is the variable a numeric or character counterpart of a primary variable or assigned using a codelist?  
E.g. ADSL.SEXN, ADSL.APERIODC, ADSL.ACOUNTRY
- Are the variable values set independently of any subject-related data values?  
E.g. CRITy, ADURU, SRCDOM, DTYPE
- Is the variable a BDS parameter related variable?  
E.g. PARAM, PARAMCD, PARAMN, PARAMLBL, PARCATY

If YES to any

ASSIGNED

If NO

### Step 3

- Is the variable derived\* by an algorithm or reproducible rule that is dependent upon other data values?  
\* Derivations include transformation, calculation, categorization, aggregation, modeling, imputation etc  
E.g. ADY = ADT - ADSL.TRTSDT + 1  
E.g. CHG = AVAL - BASE  
E.g. ANL0IFL = Flag as Y for worst case post baseline records  
E.g. BRTHDT = Numeric date part of DM.BRTHDTC.  
See SAP for imputation rules.

If YES to any

DERIVED

If NO

### Action

Query with Analysis Standards Team\*\*

\*\*Reaching this step indicates a potential misunderstanding or inadequacy in meeting Origin assignment criteria. Consultation with the Analysis Standards Team may prompt updates to the standards or further user education.

## 5. Next Steps

### We plan

- to roll this out internally to support our study teams resulting in greater efficiency and improved consistency
- to share on other platforms to get further insight from industry peers
- to collate the feedback and prepare a formal paper, with a view to working with CDISC/PHUSE to create guidance and training material

## References

PHUSE Define-XML Version 2.0 Completion Guidelines

## Acknowledgements

We would like to express thanks to the Clinical Analysis and Reporting Standards Team and to all survey responders.

## Contact

We welcome any comments or suggestions. Please reach us at:  
[lucy.x.watson@gsk.com](mailto:lucy.x.watson@gsk.com), [saqwindar.x.singh@gsk.com](mailto:saqwindar.x.singh@gsk.com)