From mock shells to Output deliverables - transforming a menace to metadata

Bernadett Seregi (Fortrea, Senior Biostatistician)

Dubravka Georgijevic (Fortrea, Senior Principal Statistical Programmer)

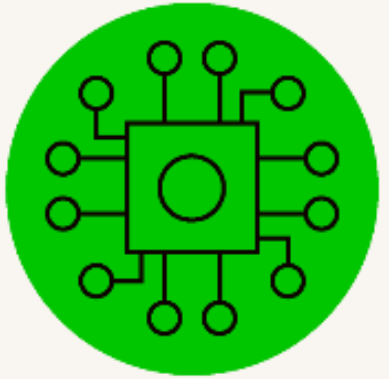
Daniel Windgassen (AZ, Statistical Programming Associate Director)

Fortrea

I will be talking about

1. Challenging aspects of clinical trial data analysis and reporting (A&R)
2. A powerful sponsor tool for creating and maintaining mock TLF shells
3. How the CRO uses the tool for downstream automations to address some of the A&R challenges, facilitating sponsor-CRO collaboration and traceability during the operational phase of a clinical trial

1. The operational challenge at hand



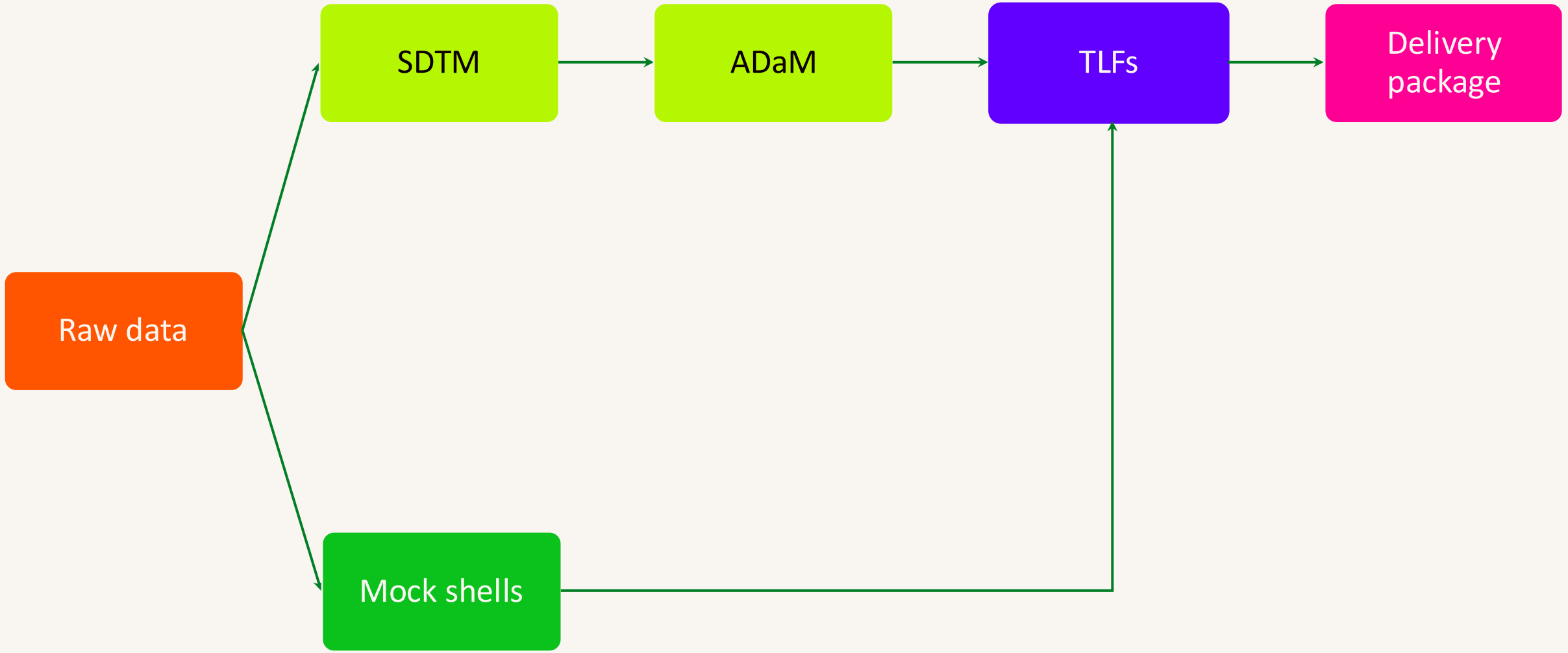
The operational incentive

1. Standardize
2. Predefine
3. Fine-tune
4. Deliver on time, with high quality

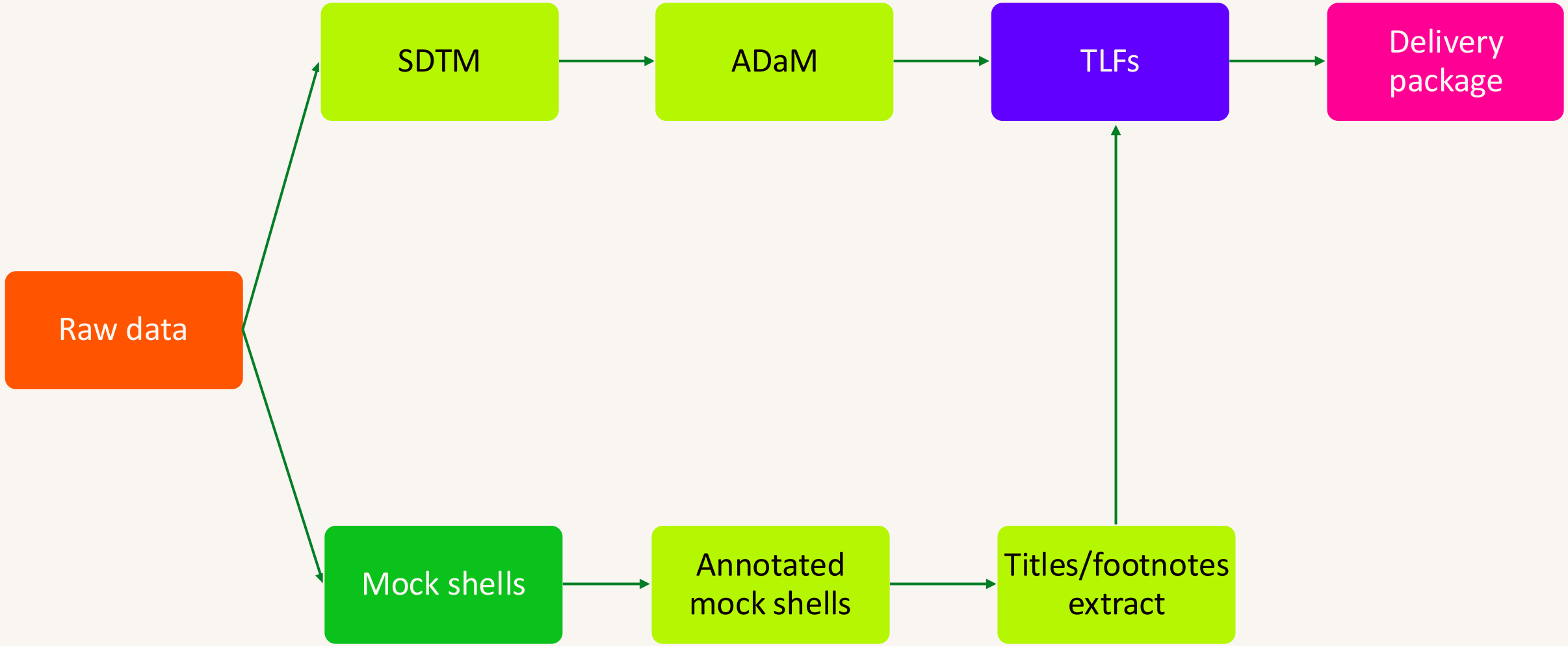
Translating clinical requirements for data professionals

- CSP, SAP
- Data set and TLF standard libraries
- Mock shells
- Data set specifications
- Programming guidelines

The A&R pipeline

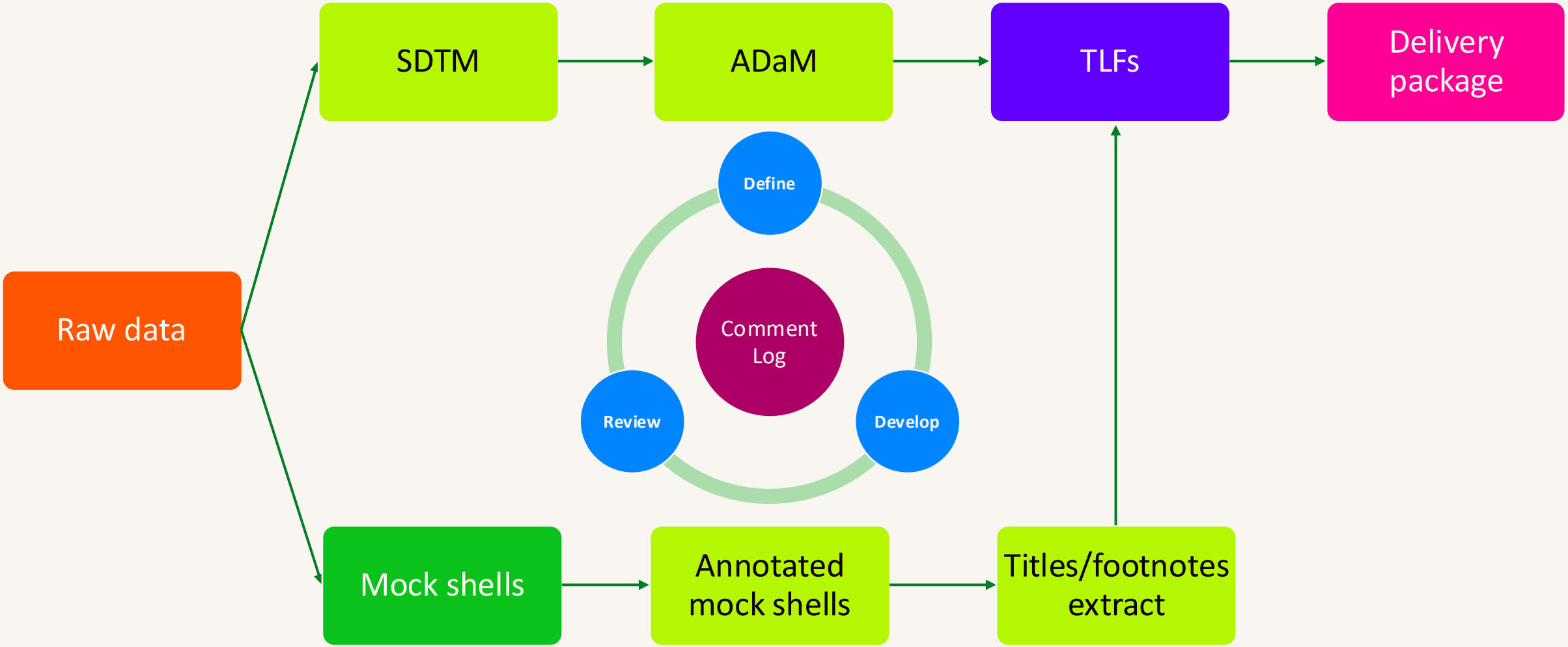


The A&R pipeline



Fortrea

The A&R pipeline



Fortrea

… but why should we care so much about the Comment Log?

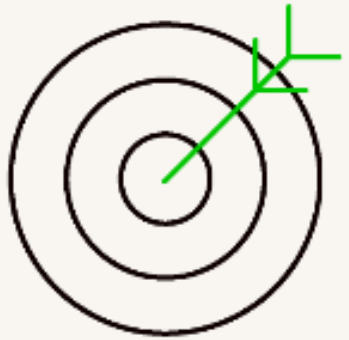
- **ICH E6 R3**.
  - ***Section 3.16.1 Data Handling***
    - (b) The sponsor should apply **quality control to the relevant stages of data handling** to ensure that the data are of sufficient quality to generate reliable results.
    - (e) The sponsor should ensure that documented processes are implemented to ensure the **data integrity for the full data life cycle**.
  - ***Section 3.16.2. Statistical Programming and Data Analysis***
    - (b) The sponsor should ensure that appropriate and documented **quality control of statistical programming and data analysis is implemented** (e.g., for sample size calculations, analysis results for IDMC review, outputs for clinical trial report, statistical or centralised monitoring).
    - (c) The sponsor should ensure the **traceability of data transformations and derivations** during data processing and analysis.
    - (f) The sponsor should retain the **statistical programming records that relate to the output** contained or used in reports of the trial results, including quality control/validation activities performed. Outputs should be traceable to the statistical software programs, dated and time stamped, protected against any changes, and have access controls implemented to avoid inappropriate viewing of information that may introduce bias.
  - ***Section 4.2.3 Review of Data and Metadata***

    Procedures for review of **trial-specific data, audit trails and other relevant metadata** should be in place. It should be a planned activity, and the extent and nature should be risk-based, adapted to the individual trial and adjusted based on experience during the trial.

… but why should we care so much about the Comment Log?

- Operations are not “one-and-done”.
- Challenges that we may encounter:
  - With the core document changes the mock shells and down-stream documents also need to be adjusted.
  - Phase in the development program: exploratory phase trials are deeply impacted
  - Complex study designs, e.g. modular studies
  - Changes in strategical goals, e.g. introduction of interim analyses
  - Changes in the study team personnel – new expectations, questions, ideas, differences in preferences when it comes to phrasing
- Traceability of comments, discussions through all the changes is challenging, especially when output titles and numbers change – The typical dilemma is: which comments are referring to the same output?

2. Well designed tool from the Sponsor, facilitated integration



The MOSAIC system

- Medical Data Review tool developed internally by AstraZeneca and presented to the wider public in 2022 on CDISC US Interchange.
- From the operational perspective
  - The tool contains the AstraZeneca output standard library
  - Modular and centralized build-up of the mock shells
  - Easy comparison between mock shells of different clinical trials
  - Delivery specific mock shell creation

Feature 1. – Titles / footnotes extraction

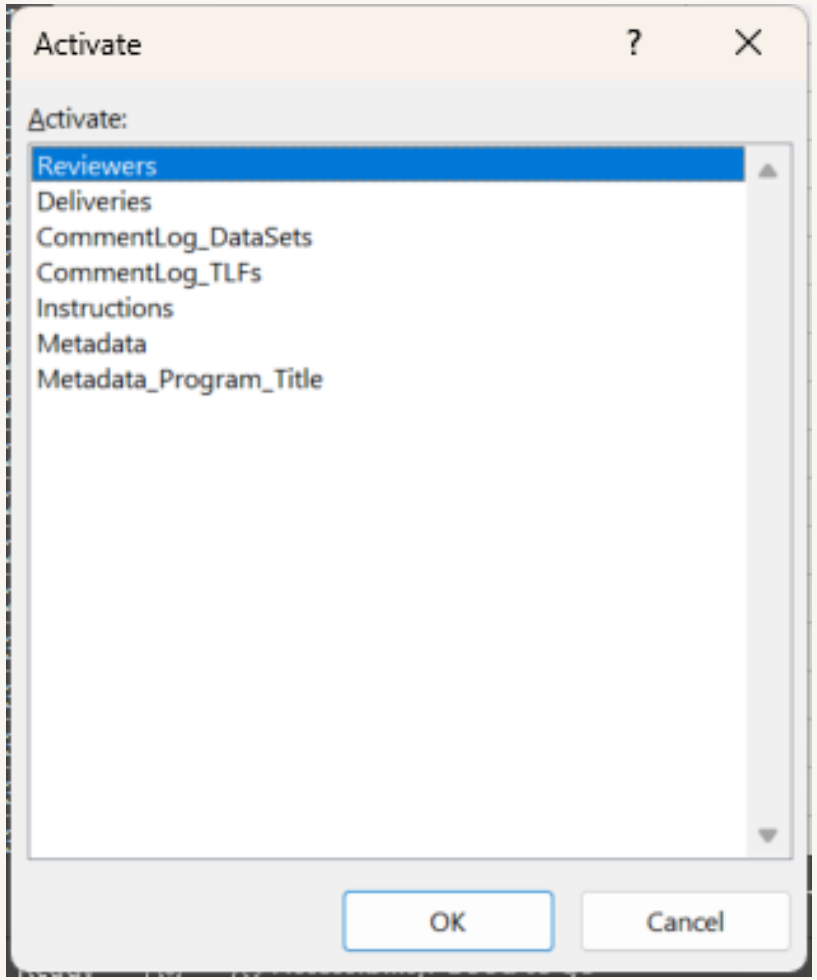
sect_num	sect_ttl	program	suffix	parm	value
14.1	Study Population	t_ds200	a	outfile	
14.1	Study Population	t_ds200	a	outtype	rtf
14.1	Study Population	t_ds200	a	azsolid	AZTSP01
14.1	Study Population	t_ds200	a	tocnumber	14.1.1
14.1	Study Population	t_ds200	a	title1	j=L 'AstraZeneca' j=R 'Page x of y'
14.1	Study Population	t_ds200	a	title2	j=L 'Study number D1234C00001 <<Study name>> - <<Deliverable if not final>>, <<Dummy Treatment>>, <<Data cut-off ddmmmyyyy>>'
14.1	Study Population	t_ds200	a	title4	j=C 'Table 14.1.1'
14.1	Study Population	t_ds200	a	title5	j=C 'Disposition (All Enrolled Set)'
14.1	Study Population	t_ds200	a	footnote1	j=L 'Screened subjects are those from whom informed consent has been received.'
14.1	Study Population	t_ds200	a	footnote2	j=L 'Percentages are based on the number of subjects started treatment.'
14.1	Study Population	t_ds200	a	footnote3	j=L '<<Period>> is defined as <<specify>>.'
14.1	Study Population	t_ds200	a	footnote4	j=L '<<Milestone>> is defined as <<xxx>>. '
14.1	Study Population	t_ds200	a	footnote5	j=L 'n Number of subjects per category.'
14.1	Study Population	t_ds200	a	footnote6	j=L '<<output program path>> <<output file name>> <<date/time>>'

Feature 2. – The program name (suffix) as unique ID

sect_num	sect_ttl	program	suffix	parm	value
14.1	Study Population	t_ds200	a	outfile	
14.1	Study Population	t_ds200	a	outtype	rtf
14.1	Study Population	t_ds200	a	azsolid	AZTSP01
14.1	Study Population	t_ds200	a	tocnumber	14.1.1
14.1	Study Population	t_ds200	a	title1	j=L 'AstraZeneca' j=R 'Page x of y'
14.1	Study Population	t_ds200	a	title2	j=L 'Study number D1234C00001 <<Study name>> - <<Deliverable if not final>>, <<Dummy Treatment>>, <<Data cut-off ddmmmyyyy>>'
14.1	Study Population	t_ds200	a	title4	j=C 'Table 14.1.1'
14.1	Study Population	t_ds200	a	title5	j=C 'Disposition (All Enrolled Set)'
14.1	Study Population	t_ds200	a	footnote1	j=L 'Screened subjects are those from whom informed consent has been received.'
14.1	Study Population	t_ds200	a	footnote2	j=L 'Percentages are based on the number of subjects started treatment.'
14.1	Study Population	t_ds200	a	footnote3	j=L '<<Period>> is defined as <<specify>>.'
14.1	Study Population	t_ds200	a	footnote4	j=L '<<Milestone>> is defined as <<xxx>>.'
14.1	Study Population	t_ds200	a	footnote5	j=L 'n Number of subjects per category.'
14.1	Study Population	t_ds200	a	footnote6	j=L '<<output program path>> <<output file name>> <<date/time>>'

3. From menace to metadata – the comment log

Phase 1. Standard excel functionalities for standardization



- Drop down menu for standard fields
  - Reviewer initials
  - Deliveries – unique delivery names can be assigned
  - Output types
  - Comment resolution status
  - Developmental level where the comment will need to be addressed

Output type	Status	Root cause
SDTM	Open	Raw Data
SDTM Specs	Ongoing (Sponsor)	SAP
ADaM	Ongoing (Data Manager)	SDTM
ADaM Specs	Ongoing (Programming)	SDTM Specs
Adhoc	Ongoing (Statistics)	ADaM
SAP	Resolved	ADaM Specs
	No change	Mock Shells
	On-hold	TLF program
		Change request
		(Duplicate)

Phase 2. Linking the comment log and the mock shells with unique IDs

- MOSAIC output numbers and titles can be transformed to be used as a drop-down list
  - Helps to avoid typos in TFL number / title – facilitated discussion between the study team and the A&R team
- Program names as unique identifiers can be associated with certain outputs for each comment (XLOOKUP function)
  - Traceable history of comments on a certain output from mock shells until final delivery, even if table number or title changes

Phase 2. Linking the comment log and the mock shells with unique IDs

Unique combination of program name and TLF number and titles as a key in the comment log (Metadata_Program_Title sheet)	
program	title_full
t_ds200	Table 14.1.1 Disposition
t_dm200	Table 14.1.2 Analysis sets
t_dc200	Table 14.1.3.1 Disease characteristics at baseline (Full analysis set)

Use of the program name and output number and title combination (drop-down menu) in the CommentLog_TFL sheet

Delivery	Item#	Program Name	Output Number and Title	Reviewer	Date	Review Comment	…
Delivery1_dry-run	1	t_ds200	Table 14.1.1 Disposition	SBdt	7-Jan		

The impact for the CRO

Table number, title and footnotes are defined in a central place

Unique program names enhance traceability in the comment log

Fortrea

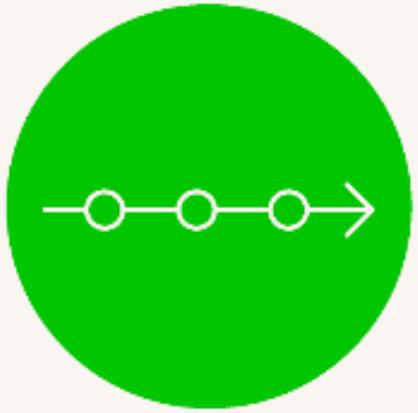
Thank you for your attention!

Fortrea

Q & A

Fortrea

Highlights



Operational journey

- Mock shells are just the beginning
- Early and precise definition facilitates effective production, and change management in later phases of the clinical trial
- Reduction in the manual text management reduces complexity and facilitates delivery (e.g. via a system like MOSAIC)

ICH E6(R3)

- Increased emphasis on the data flow and governance

Developmental metadata

- The Comment Log is a significant tool in establishing the paper trail in clinical trial development.
- Identifying outputs by unique IDs allows us to
  - conserve this trail even when table numbers, titles and footnotes had changed.
  - Increase clarity by minimizing the wrongful linking of outputs and comments, reducing typos etc.

Fortrea