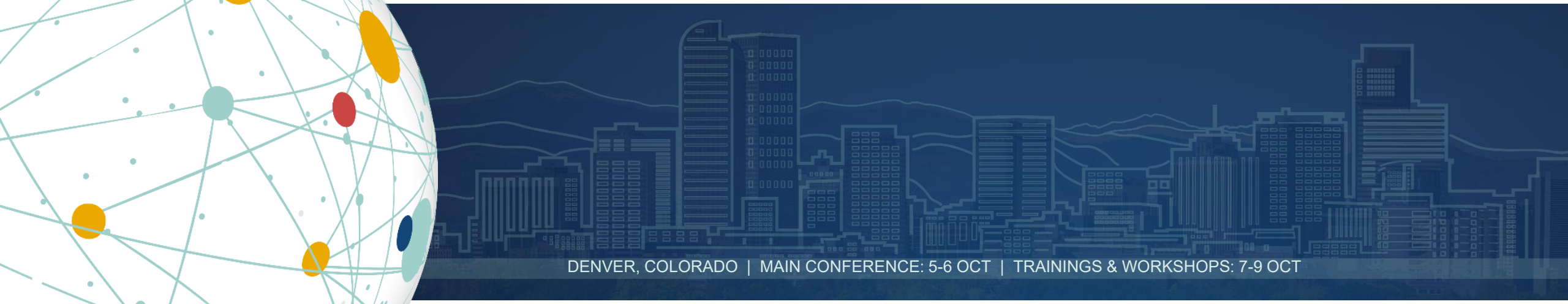


cdisc®

2026 US Interchange

THE FUTURE IS CONNECTED: STANDARDS AND AI POWERING DIGITAL TRANSFORMATION



From Manual Writing to Intelligent Drafting:
Leveraging GenAI to Create Data Reviewer's Guides for Regulatory Submissions

CDISC Conference 2026 | GenAI-Assisted Authoring

Speakers



Sirisha Davuluri is a Senior Statistical Data Sciences Lead in Vaccines at Pfizer with over 11 years of experience across multiple therapeutic areas spanning both drug side and vaccines. She has a proven track record of leading statistical programming operations across multi-phase clinical trials and global regulatory submissions. Dedicated to continuous professional growth, Sirisha is actively expanding her skills in Generative AI and intelligent automation, with a strong interest towards emerging applications in statistical data sciences and reporting.

Sirisha Davuluri
Senior Manager, Senior Statistical Data Sciences Lead
Pfizer Inc.

Speakers



Valerie Cadorett is a Manager in Statistical Data Sciences and Analytics at Pfizer Inc., where she leads initiatives focused on improving efficiency, standardization, and innovation across statistical programming processes. She is actively involved in advancing AI-enabled solutions, including agentic AI, and process governance. Valerie is passionate about leveraging technology to transform how teams work, collaborate, and generate insights, while ensuring regulatory compliance and operational excellence.

Valerie Cadorett
Manager, Statistical Programming Process Lead
Pfizer Inc.

Agenda

01 Authoring Data Reviewer’s Guides

02 GenAI: Automation Engine

03 Automating the cSDRG and ADRG

04 Outcomes

05 Trust, Citations, and Human-in-the-Loop

06 Challenges, Impact and Takeaways

Authoring Data Reviewer's Guides (DRG's)

Solving the Manual DRG Challenge: The Automated Future

Authoring DRG's



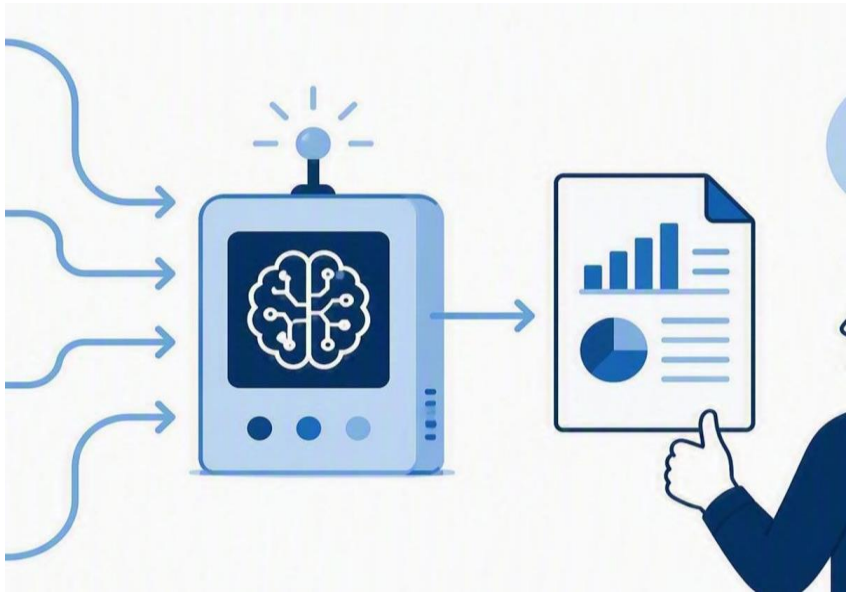
THE MANUAL STANDARD

Manual reading and re-keying

TEMPLATE GUARDRAILS
PHUSE + Pfizer DRG completion guidelines built in

FIVE SOURCE INPUTS

- **Protocol**
- **aCRF**
- **I/E Template**
- **SAP**
- **P21E DRG**



THE AUTOMATED FUTURE

Automated composition, author reviews

GenAI: Automation Engine

A Microsoft Word-integrated, generative AI platform for authoring DRGs

GenAI: Automation Engine

Word-Native Platform

A generative AI platform for authoring DRGs, built directly into Microsoft Word

Section-by-Section Drafting

Content drafted from user prompts and the study's own source files

Fact-Grounded AI Foundation

Large Language Models (LLM) paired with Retrieval-Augmented Generation (RAG)

Authors Stay in Control

An interactive AI assistant that supports the author, never replaces them



Every generated block is grounded in real study data, not invented.

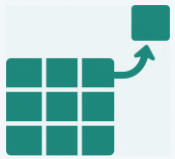
What the Engine Delivers



Section-by-Section Drafting via Predefined Templates



First-Draft Document Generation



Dynamic metadata Extraction from aCRFs



Automated Version-to-Version Delta Analysis

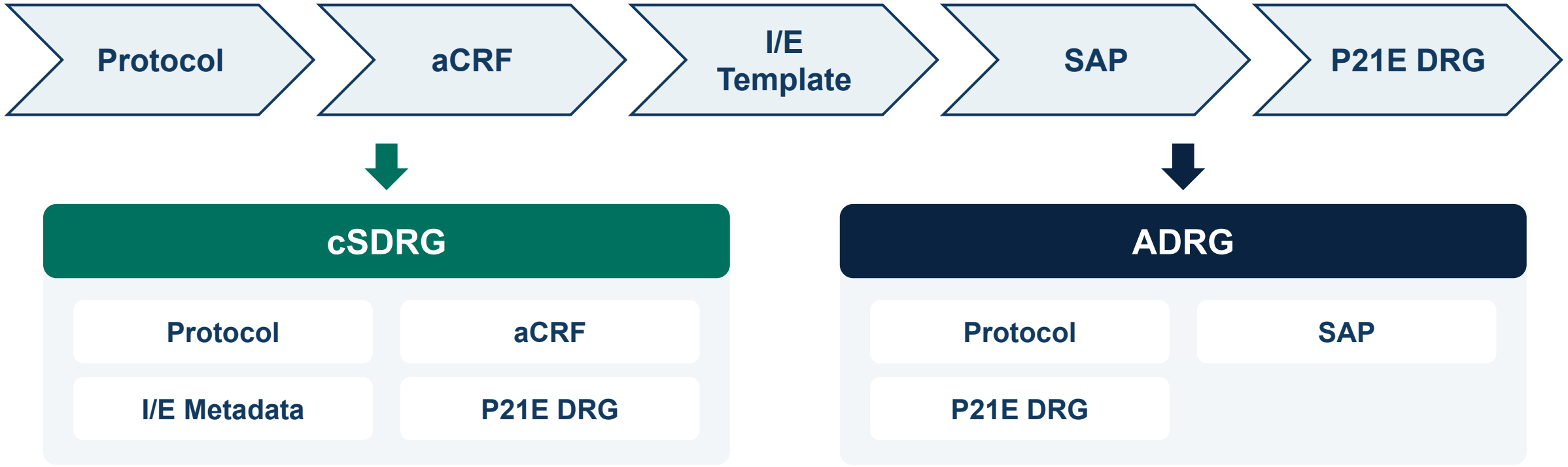


Interactive Q&A Panel for Note Preparation

From source data to first draft – five tasks the engine handles end to end.

Automating the cSDRG and ADRG

Authoring Setup



Five core inputs feed two standardized guides, generated automatically.

Authoring Process

cSDRG

TEMPLATE

Template Type*

DRG

Template Sub Type*

cSDRG

REFERENCE FILES

Pinnacle 21 cSDRG*

Select file

Protocol*

Select file

aCRF*

Select file

I/E Excel*

Select file

ADRG

TEMPLATE

Template Type*

DRG

Template Sub Type*

ADRG

REFERENCE FILES

Pinnacle 21 ADRG*

Select file

Protocol*

Select file

SAP*

Select file

Conditions

If date imputation was performed, were there rules that were used in multiple analysis datasets? No

If study has intermediate analysis datasets? No

cSDRG Automation

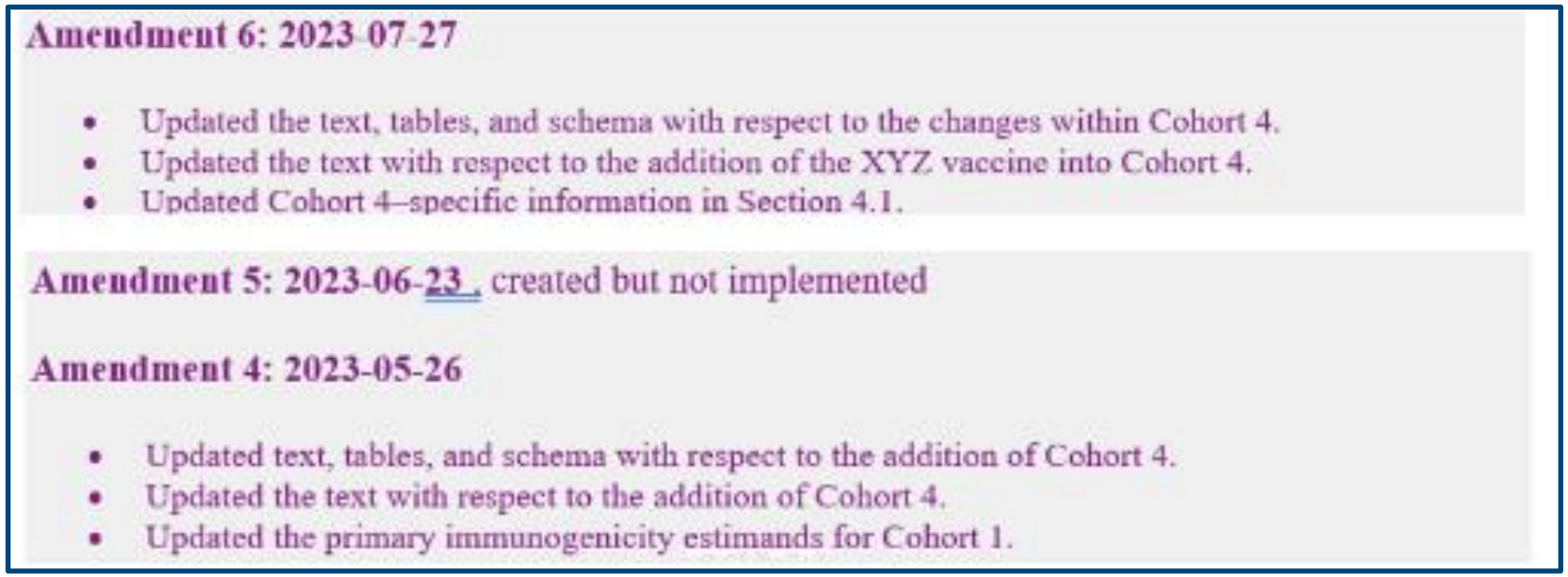
cSDRG Section	Method	What Happens
2.1 Protocol Number and Title	**Auto-Updated**	Protocol versions extracted from the Protocol
2.2 Protocol Design	**Auto-Updated**	Analyzed and summarized from the Protocol
2.3 Trial Design Datasets	**Auto-Updated**	Populated from the provided reference documents
3.1 Screen Failures Table	**Auto-Updated**	Formatted and generated automatically
3.3 Annotated CRFs	**Auto-Updated**	Reads the aCRF for “NOT SUBMITTED” pages and fields
3.4, 4.1 and 4.2 Subsections	**Auto-Updated**	Reference documents ingested to compose content
Appendix I (Inclusion/Exclusion)	**Auto-Updated**	Pulled from upstream data management metadata

ADRG Automation

ADRG Section	Method	What Happens
2.1 Protocol Number and Title	**Auto-Updated**	Copied from the Protocol title page and Document History tables.
2.2 Protocol Design in Relation to ADaM	**Auto-Updated**	Copied directly from the SAP (Overall Design).
3.5 Imputation/Derivation Methods	**Auto-Updated**	Standard text points to define.xml, plus the relevant SAP section.
4 Analysis Data Creation and Processing Issues	**Auto-Updated**	Standard text, or a generated table where intermediate datasets exist.
7 Submission of Programs	**Auto-Updated**	Program names pulled from the repository; non-compliant names flagged.

Outcomes

Automated Output: Section Drafts



Amendment 6: 2023-07-27

- Updated the text, tables, and schema with respect to the changes within Cohort 4.
- Updated the text with respect to the addition of the XYZ vaccine into Cohort 4.
- Updated Cohort 4–specific information in Section 4.1.

Amendment 5: 2023-06-23, created but not implemented

Amendment 4: 2023-05-26

- Updated text, tables, and schema with respect to the addition of Cohort 4.
- Updated the text with respect to the addition of Cohort 4.
- Updated the primary immunogenicity estimands for Cohort 1.

What to notice

Auto-Updated

Every line here is generated straight from the protocol – no manual re-keying.

Section drafts are completed, retaining the source document format.

Automated Output: Data Tables

3.3 Annotated CRFs

Collected fields and pages that have not been tabulated have been annotated as "Not Submitted". Pfizer Inc. collects certain data elements to facilitate operational processes including data cleaning and dynamically creating additional forms in the electronic data capture system. These fields and pages that have been annotated as "Not Submitted" and meet this database operational criterion above are not included in the table below.

Explanation of data fields [Not Submitted]

aCRF page Numbers	*Data Collection Field*	*Explanation of why [NOT SUBMITTED]*
11	*Comparison Term*	
12	*Comparison Term*	

What to notice

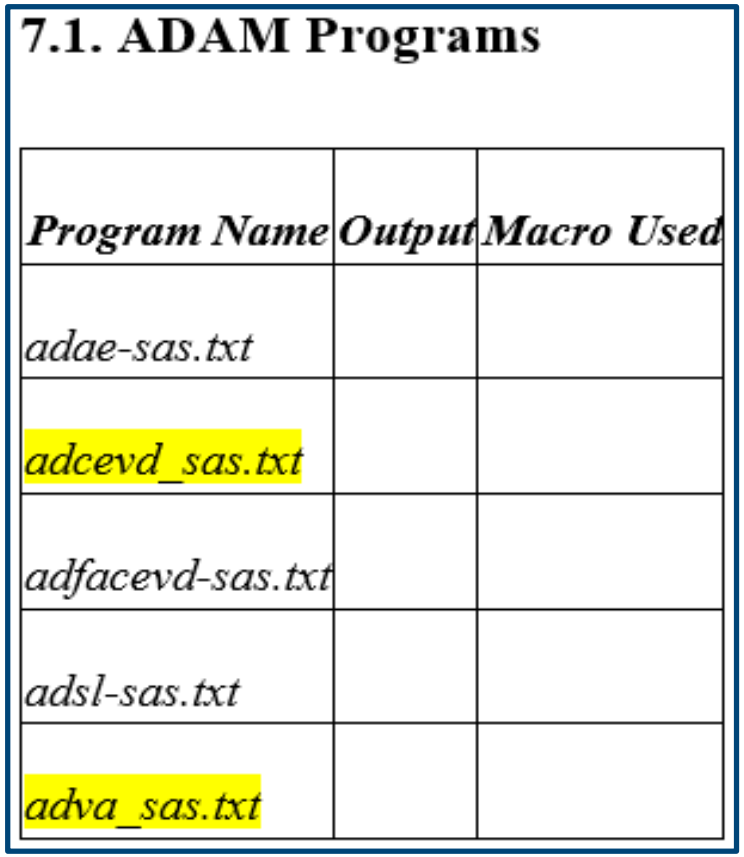
Auto-Updated

Field names and pages line up with no manual transcription.



Tables are populated from the source – ready for review.

Automated Output: Review Cues



7.1. ADAM Programs

Program Name	*Output*	*Macro Used*
adae-sas.txt		
adcevd_sas.txt		
adfacevd-sas.txt		
adsl-sas.txt		
adva_sas.txt		

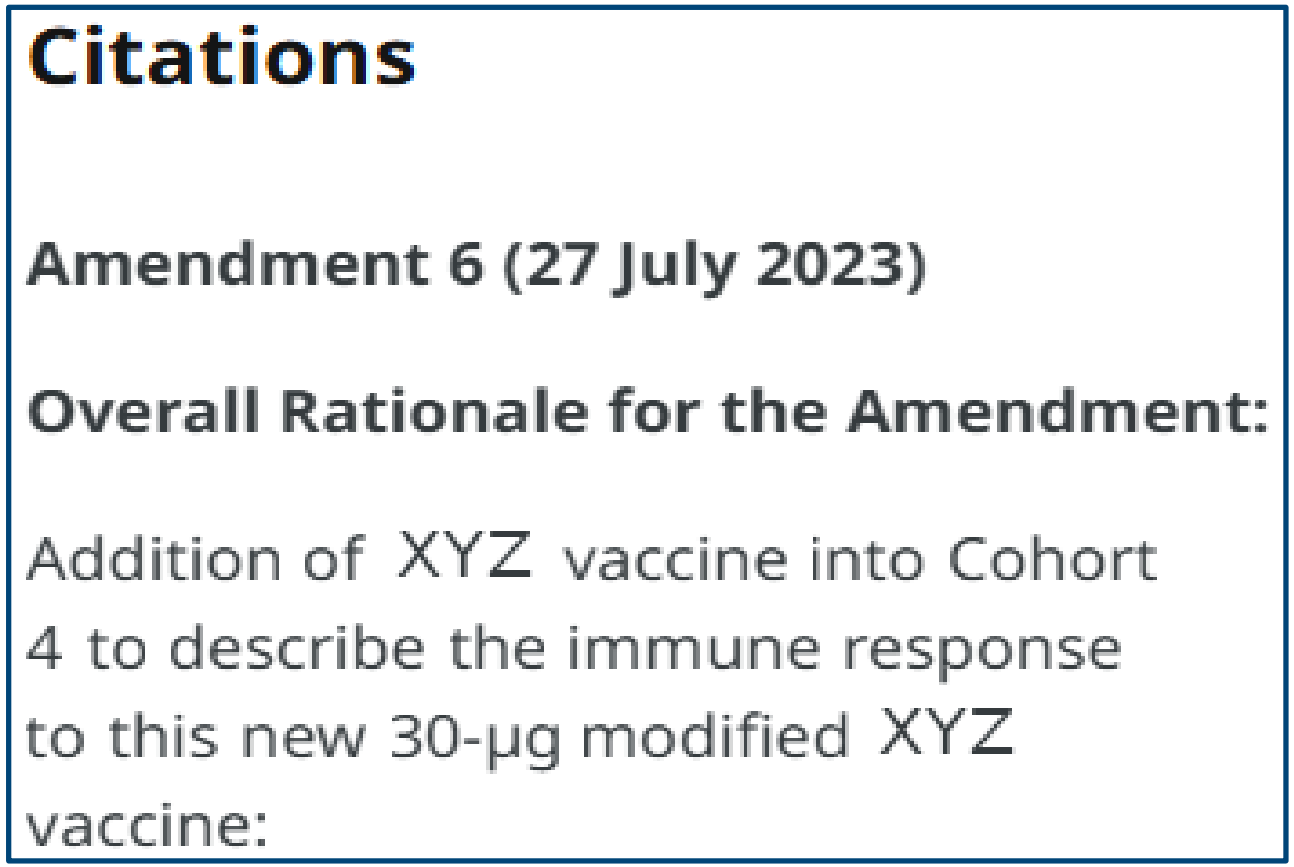
What to notice

Review Cue

Items needing a human decision are flagged inline.

The draft tells the author exactly where to look.

Automated Output: Source Citations



Citations

Amendment 6 (27 July 2023)

Overall Rationale for the Amendment:

Addition of XYZ vaccine into Cohort 4 to describe the immune response to this new 30-µg modified XYZ vaccine:

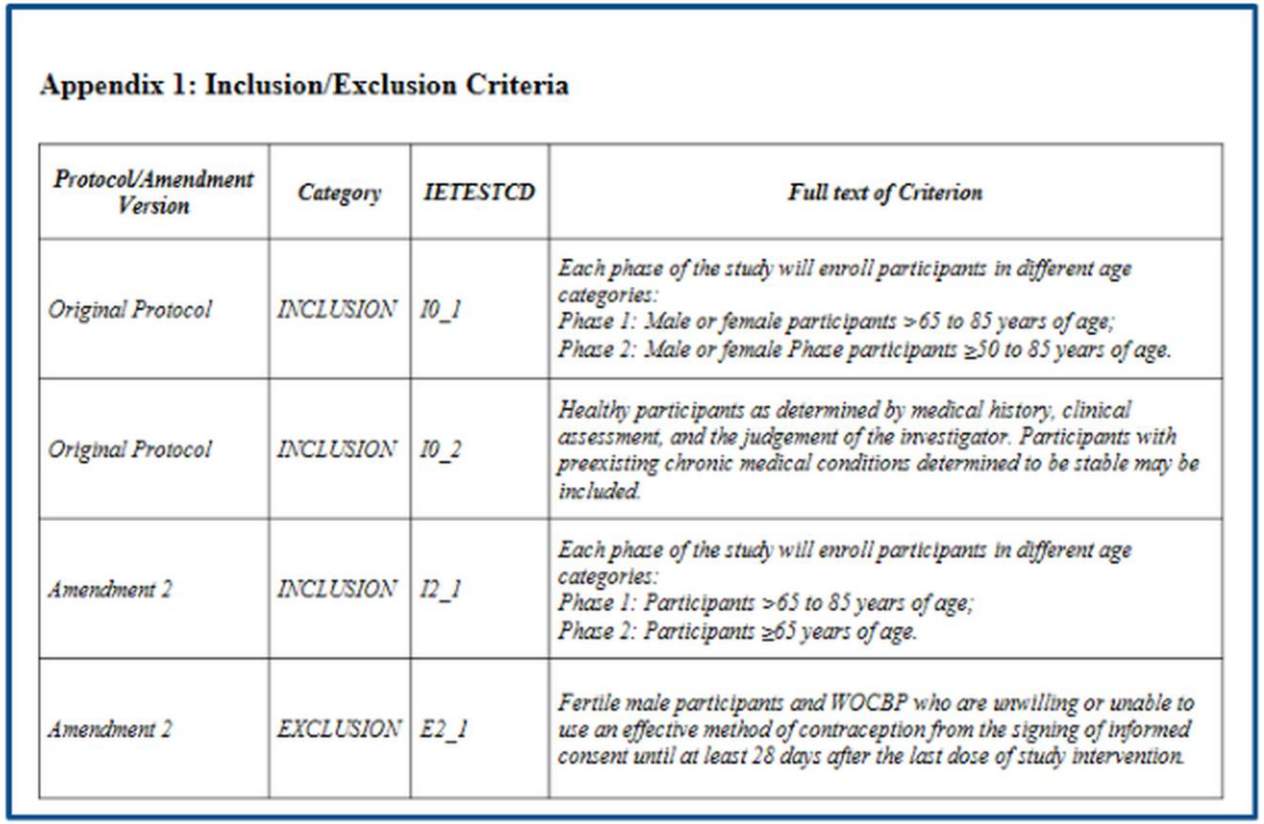
What to notice

Traceable

Each statement carries a link back to its source.

Every claim in the guide can be traced to its origin.

Automated Output: Auto Update



Appendix 1: Inclusion/Exclusion Criteria

Protocol/Amendment Version	*Category*	*IETESTCD*	*Full text of Criterion*
Original Protocol	*INCLUSION*	*I0_1*	*Each phase of the study will enroll participants in different age categories:* *Phase 1: Male or female participants >65 to 85 years of age;* *Phase 2: Male or female Phase participants ≥50 to 85 years of age.*
Original Protocol	*INCLUSION*	*I0_2*	*Healthy participants as determined by medical history, clinical assessment, and the judgement of the investigator. Participants with preexisting chronic medical conditions determined to be stable may be included.*
Amendment 2	*INCLUSION*	*I2_1*	*Each phase of the study will enroll participants in different age categories:* *Phase 1: Participants >65 to 85 years of age;* *Phase 2: Participants ≥65 years of age.*
Amendment 2	*EXCLUSION*	*E2_1*	*Fertile male participants and WOCBP who are unwilling or unable to use an effective method of contraception from the signing of informed consent until at least 28 days after the last dose of study intervention.*

What to notice

Auto-Updated

Criteria tables refresh with each protocol amendment.

The whole I/E appendix - original plus every amendment - is ingested from the source template, not built by hand.

Trust, Citations, and Human-in-the-Loop

Trust, Verifiability, and Citation Mapping

01 | TRUST



Authors verify every output and keep full control of the workflow.

02 | TRACE



Citation tags map each section back to its source data.

03 | FIND



One click opens the reference document at the exact page.

04 | FOCUS



Missing facts are flagged for the author, never guessed.

05 | SPEED



A complete first draft is ready in minutes.

Every generated statement stays traceable to its source, with the author in control.

Challenges, Impact and Takeaways

Challenges and How GenAI Handles Them

The Hard Problems	The Mitigations
Scrambled tables – Complex, nested tables in protocols that standard PDF parsers routinely scramble.	**Structured ingestion –** Complex tables come through intact, not scrambled, converted to structured text before indexing.
Page limitations – Study protocols exceeding 200 pages, past older model context-window limits.	**Hierarchical chunking –** Only the passages that matter are retrieved, however long the protocol.
Risk of Fabrication AI may fabricate details or values when information is not found.	**Strict fact-grounding –** Missing context is reported back, never invented.

Each known failure mode has a specific control - none are left to chance.

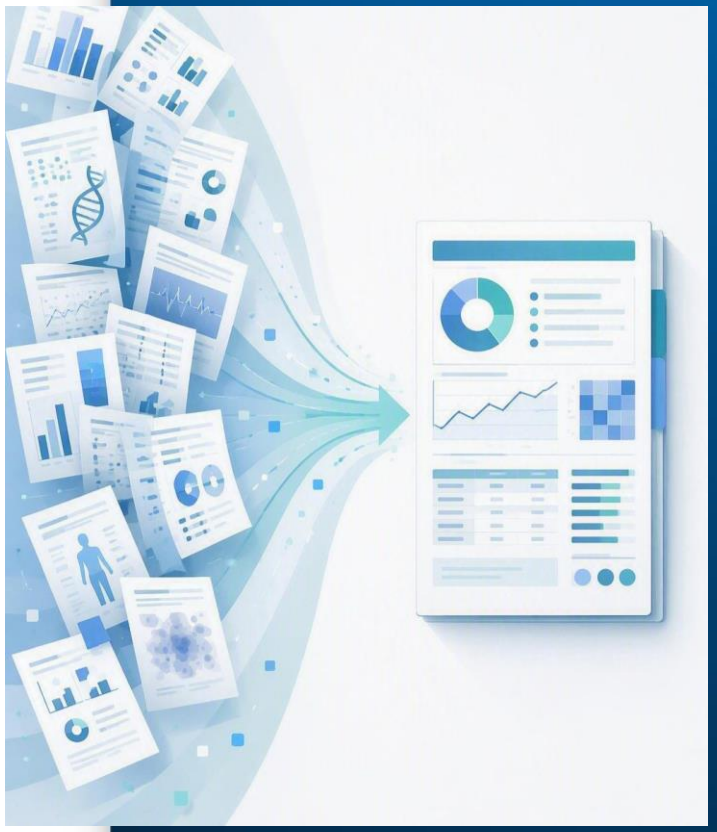
AI-Accelerated, Human-Governed

Author-approved – GenAI drafts pre-defined sections; authors accept or reject every proposed passage.

Copied text flagged – Purple text marks content lifted from source files, as seen on the section draft.

Nothing auto-final – No section reaches the submission without an author's review.

GenAI drafts - the author decides.



Where This Leaves Us

Proven on real studies – cSDRG and ADRG sections drafted end to end from the actual source data.

Effort reclaimed – Days of manual assembly collapse into a reviewed first draft.

Built to scale – The same pipeline extends to any study using standardized CDISC sources.

88-94%

less manual effort

~65%

less drafting time

Minutes

to a reviewable version

Contact Information

Sirisha Davuluri

Pfizer Inc.

Sirisha.Davuluri@pfizer.com

Valerie Cadorett

Pfizer Inc.

Valerie.Cadorett@pfizer.com

Questions and follow-ups welcome

Acknowledgements

Christine Rossin

Christina Scienski

Helen Yang

Liping Zhang

Twaha Bukenya



Questions?