Public



Accelerating Submission Readiness:

An AI Driven Framework CSR Automation

Public

Author Introduction

Niharikka Tyagi

Software Developer at Sycamore Informatics with over 2 years of experience advancing the DSW product, focusing on applying software solutions and AI to streamline pharmaceutical processes and accelerate clinical trial outcomes

I am passionate about building tools that positively impact healthcare and research.



Public

Agenda

- 1. Clinical Study Lifecycle
- 2. How CSR is created today
- 3. Challenges in CSR generation
- 4. Our Solution
- 5. How the Tool Works
- 6. Key Features
- 7. Conclusion

Public

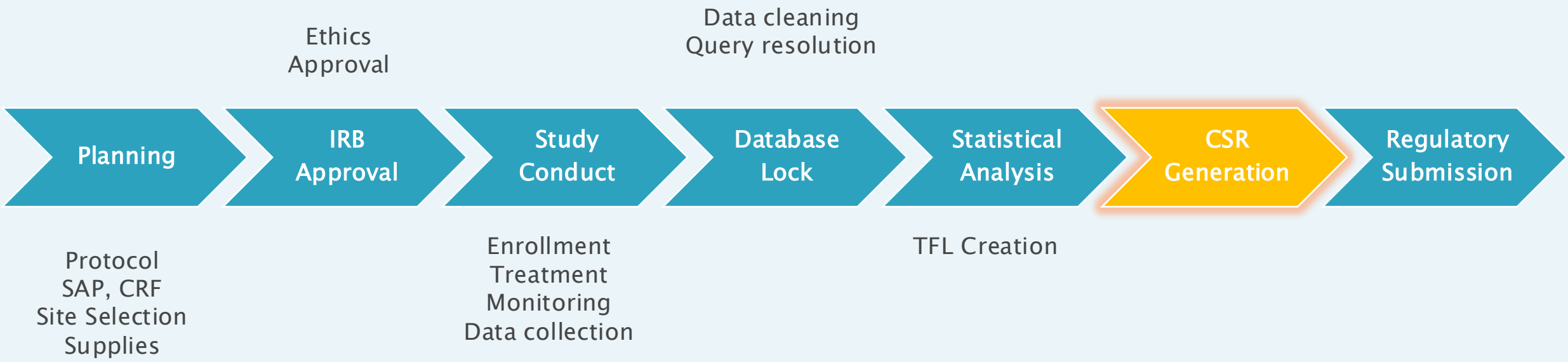


Clinical Study Lifecycle

And where CSR fits in

Public

Clinical Study Lifecycle

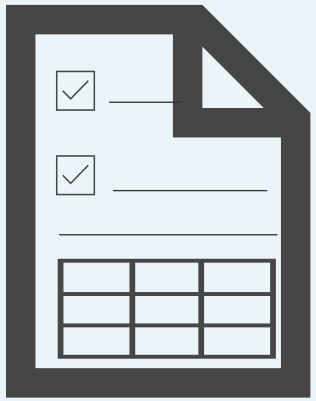


The Clinical Study Report is a critical milestone in the path to regulatory submission—synthesizing months of planning, execution, and analysis into a single, comprehensive document.

Public

Clinical Study Report

A Clinical Study Report is the definitive regulatory document that presents the complete record of a clinical trial—its design, conduct, analysis, and conclusions.

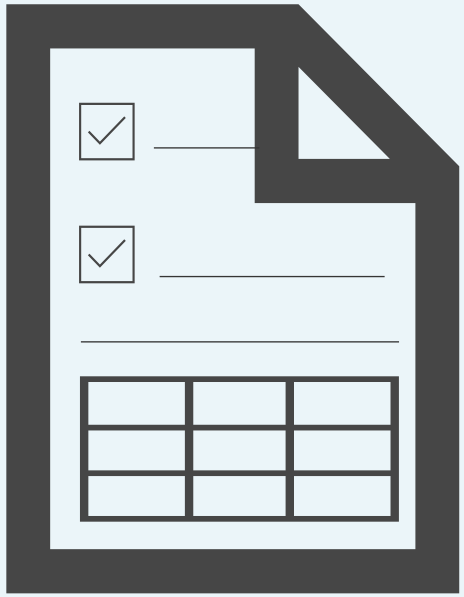


Submitted to regulatory authorities (ex: FDA), it serves as the primary evidence supporting a drug's safety and efficacy for approval.

Public

Clinical Study Report

1. Protocol Content Extraction
2. Tense Conversion
3. Statistical Output Integration
4. Narrative Drafting
5. Multi-Stakeholder Review Cycles
6. Compliance Verification



Manual Generation Of CSRs Today

Public



Challenges in Creating CSRs

Public

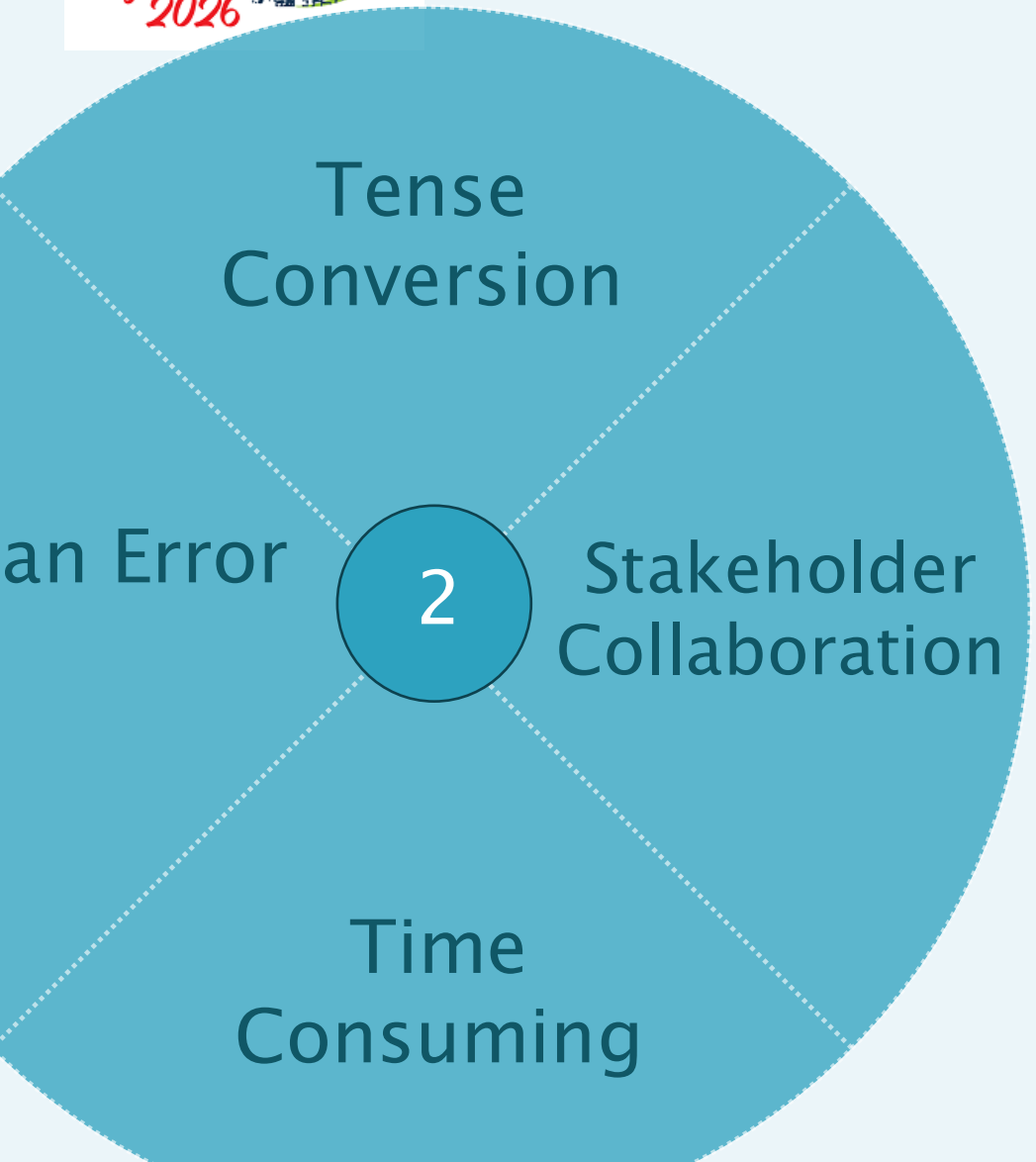
Challenges in Creating CSRs



- CSR creation can take weeks or even months
- Requires manual extraction, structuring, and rewriting of information from various docs
- Regulatory teams must ensure accuracy, consistency, and adherence to guidelines
- Long review cycles
- Inefficient use of resources

Public

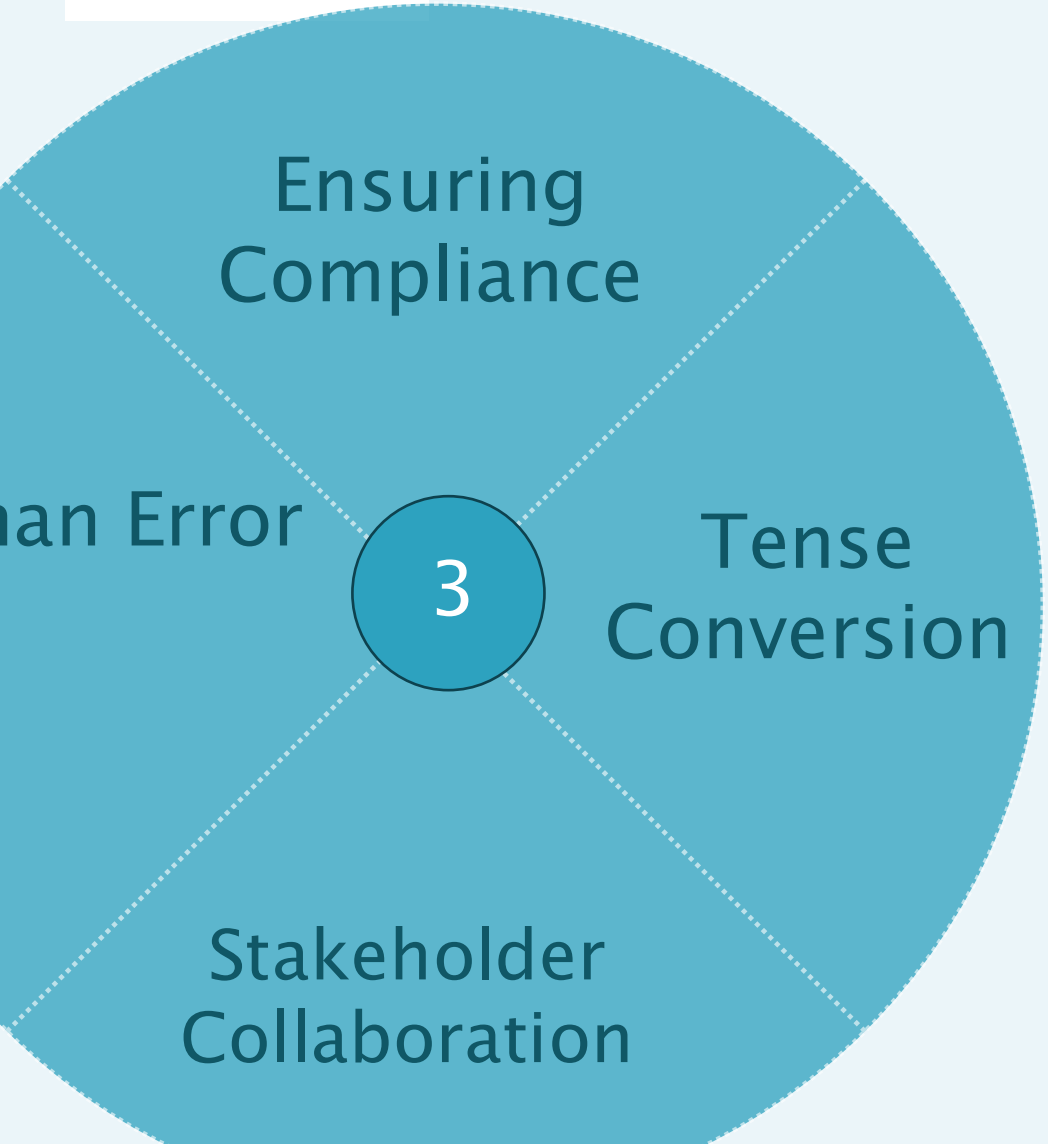
Challenges in Creating CSRs



- The CSR is not created by a single individual
- Requires input from clinicians, medical writers, statisticians, etc.
- Regulatory Coordinating across teams and managing dependencies can be challenging and time consuming

Public

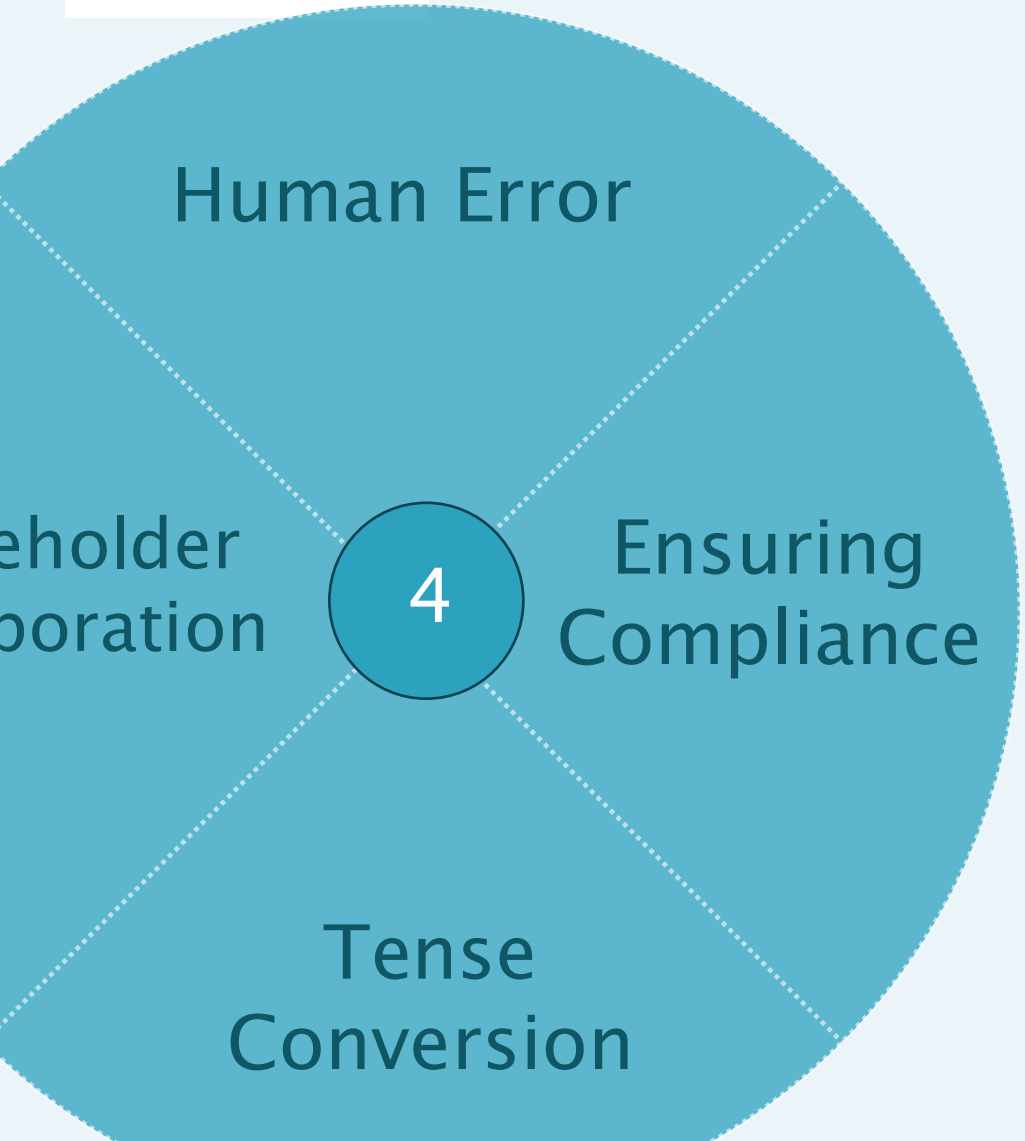
Challenges in Creating CSRs



- Increases time for CSR reviews
- There exists an inconsistency in language of protocol content (future) and CSR content (past) across studies even within same project

Public

Challenges in Creating CSRs



- CSRs must strictly adhere to –
- ICHE3 Guidelines
- Accurate Tense Conversion
- Consistent structure, content and format
- Any inconsistency can lead to delays, rework or even rejection from regulatory bodies

Public

Challenges in Creating CSRs



- Manual drafting increases likelihood of –
- Error prone calculations
- Misinterpretation of in-line tables
- Omission of important data
- Incoherence and inaccuracies

Public



The Solution

Public



The Solution

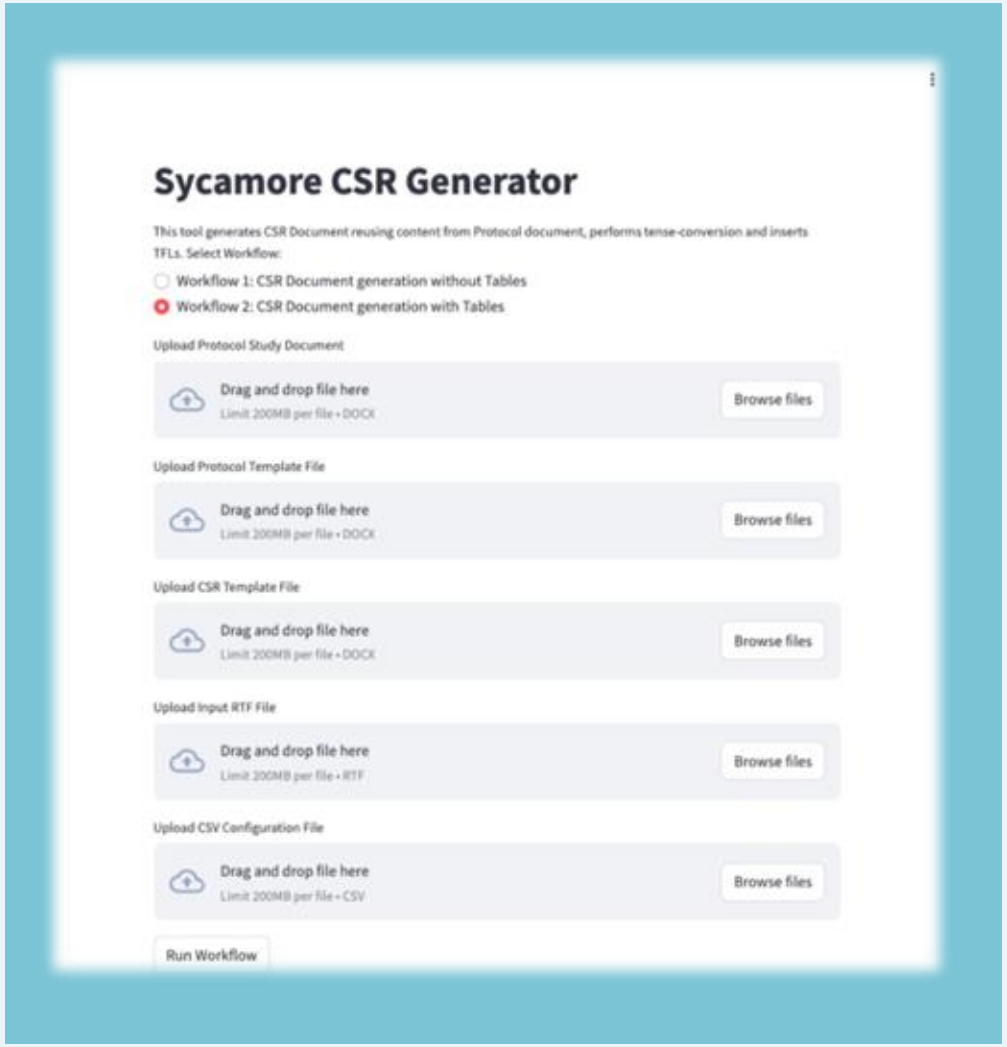
AI Powered CSR Generator

Public

The Solution

A Gen AI solution to reuse clinical content from Protocol, include and summarize the analysis results for the generation of a CSR.

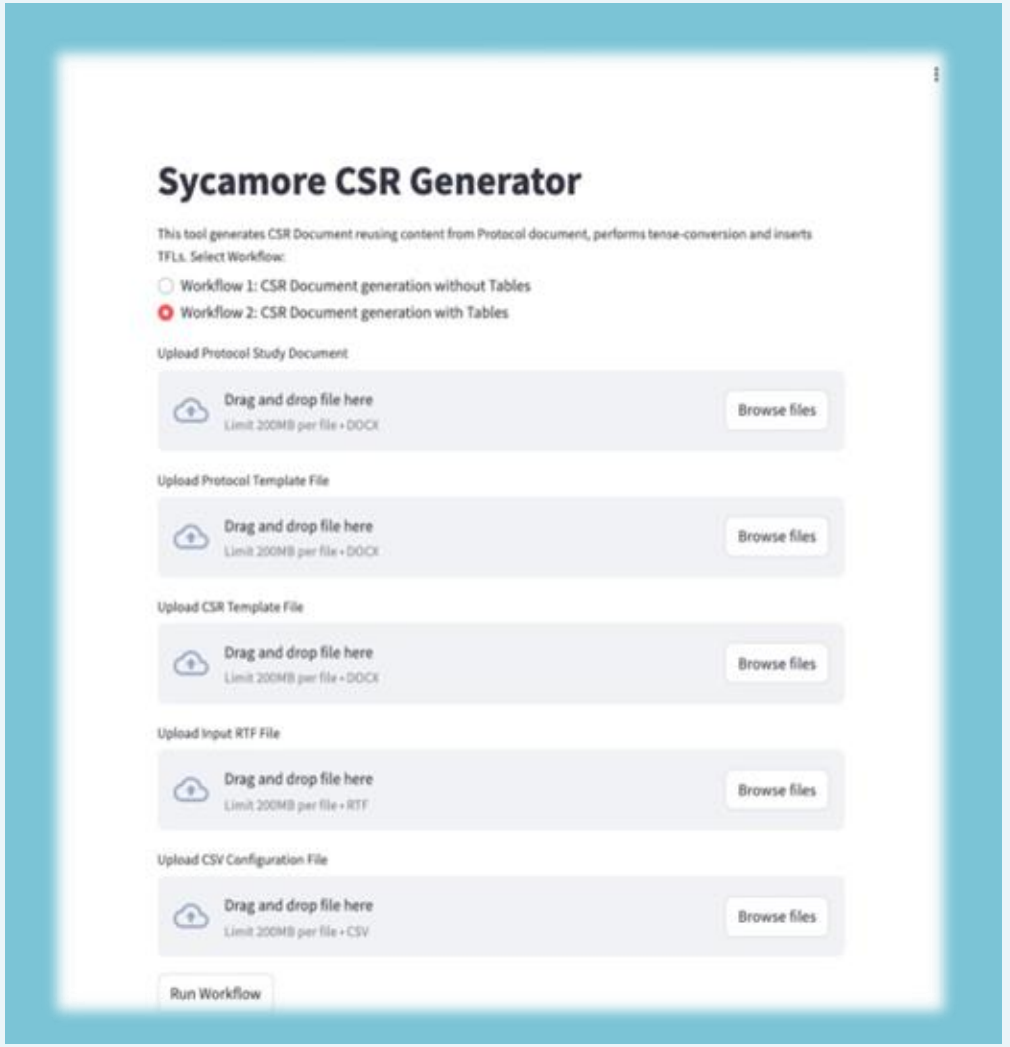
AI Powered CSR Generator



Public

The Solution

- Saves Time
- Saves Manual Effort
- Key Insights
- One Time Set Up



AI Powered CSR Generator

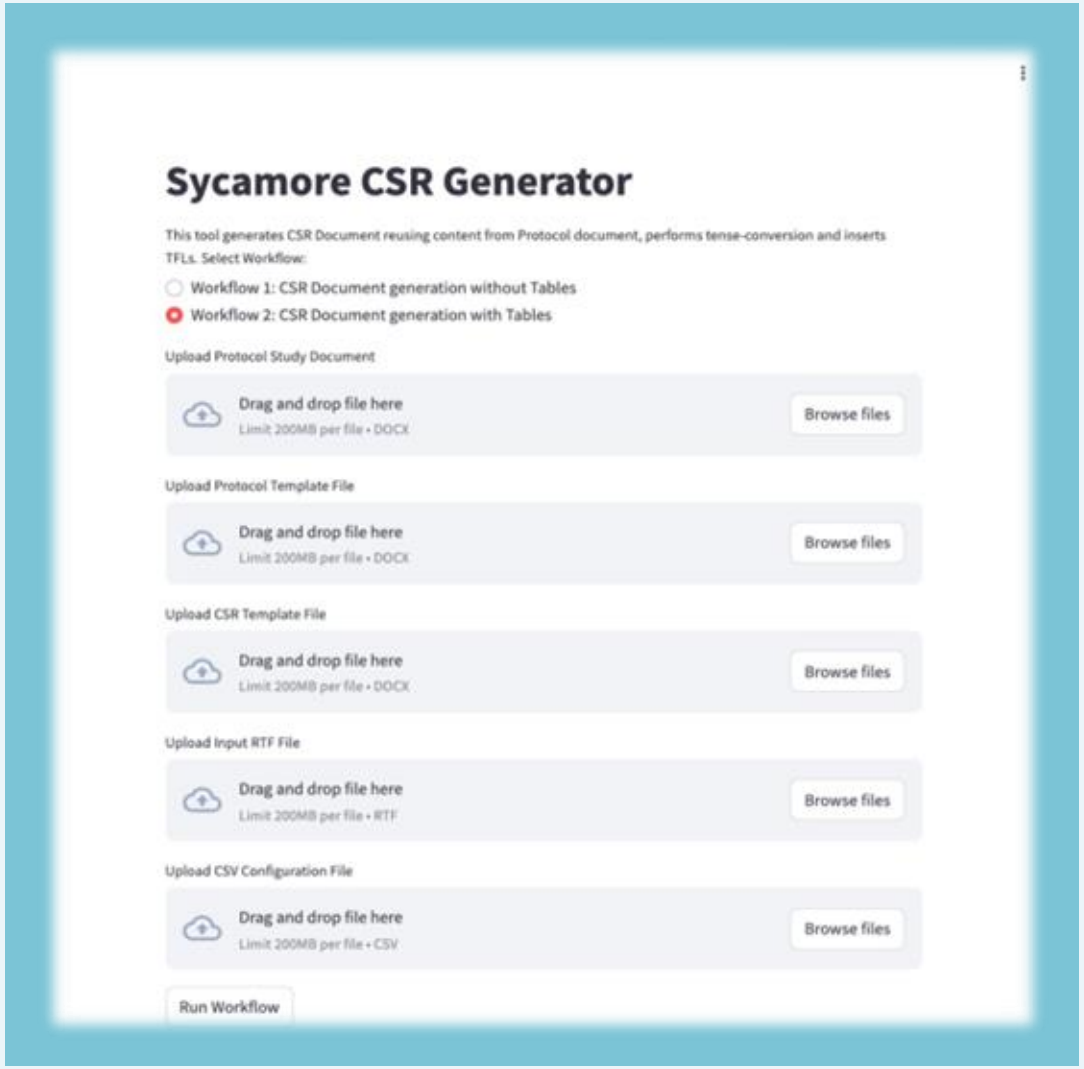
Public

How it Works

Step 1: Prepare Templates

Placeholder Mapping:
Place <sync_xyz> placeholders in Protocol and CSR Templates

Tense-conversion setup:
Use {{past:<sync_xyz>}} syntax in CSR template for tense-conversion



Public

Input Files with Placeholders

2.2.2 Secondary Endpoints

Secondary Endpoint	Summary Measure
<sync_secondary_endpoint>	Number and percentage of subjects overall, by severity, by relatedness, by seriousness, and by action (ie, IP discontinuation)

2.3 Exploratory Objectives and Endpoints

2.3.1 Exploratory Objectives

Instructional Text (exactly same text present in both protocol template and protocol study document) gets omitted automatically from the generated CSR. The exploratory objectives of this study are:

<sync_exploratory_objective>

2.3.2 Exploratory Endpoints

The exploratory endpoints for this study are as follows:

<sync_exploratory_endpoint>

3 Study Design and Oversight

3.1 Overall Design

<sync_overall_design>

3.1.1 Screening Period

<sync_screening_period>

3.1.2 Lead-in Period

<sync_leadin_period>

Protocol Template

8.2.2 Secondary Endpoints

{{past:<sync_secondary_endpoint>}}

8.3 Exploratory Objectives and Endpoints

8.3.1 Exploratory Objectives

{{past:<sync_exploratory_objective>}}

8.3.2 Exploratory Endpoints

{{past:<sync_exploratory_endpoint>}}

9 Investigational Plan

9.1 Overall Study Design and Plan: Description

{{past:<sync_overall_design>}}

9.1.1 Screening Period

{{past:<sync_screening_period>}}

9.1.2 Lead-in Period

{{past:<sync_leadin_period>}}

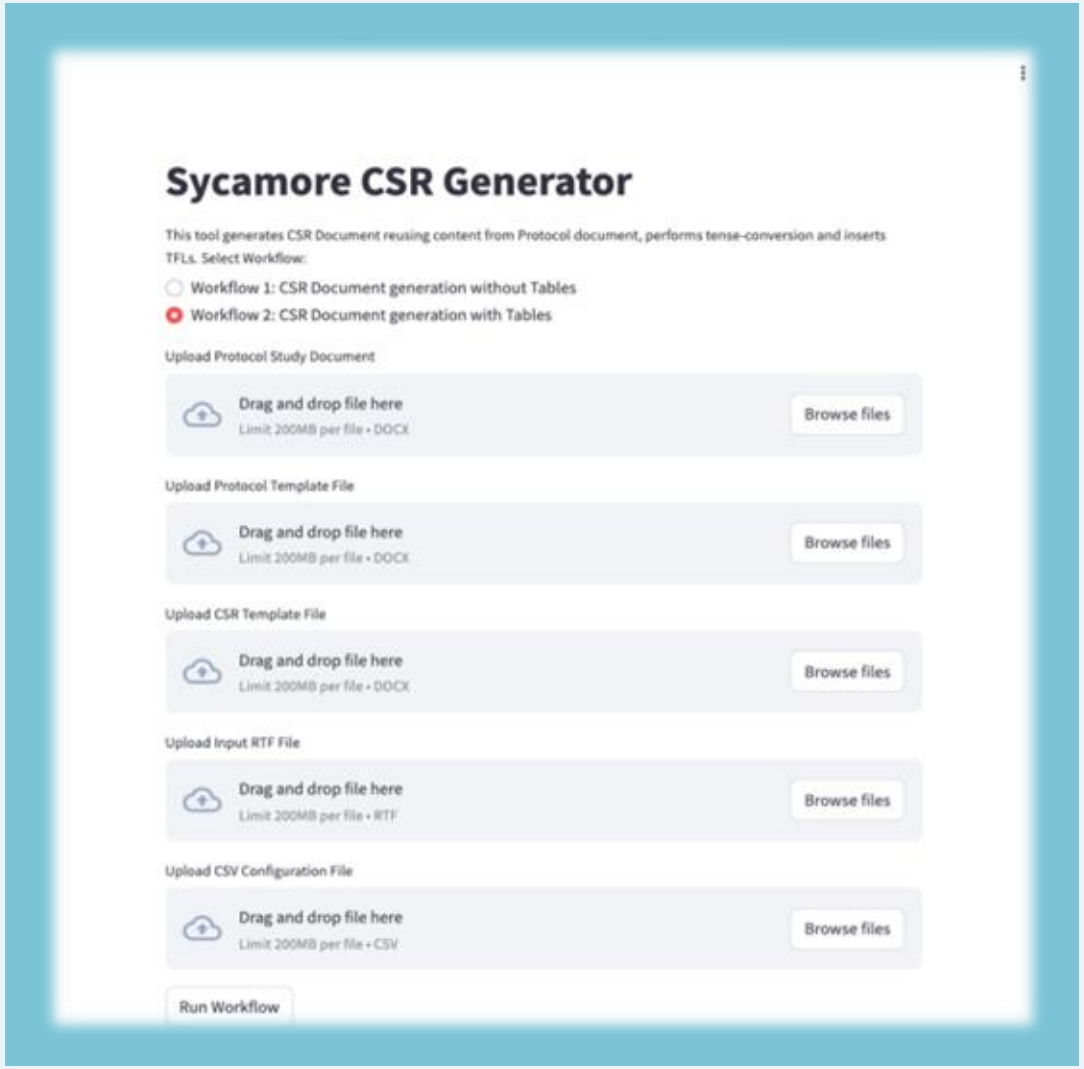
CSR Template

Public

How it Works

Step 2: Upload Required Files

- Protocol Study Document
- Protocol Template
- CSR Template
- Multi-page Tables (RTF) Files
- Summary table Configuration File



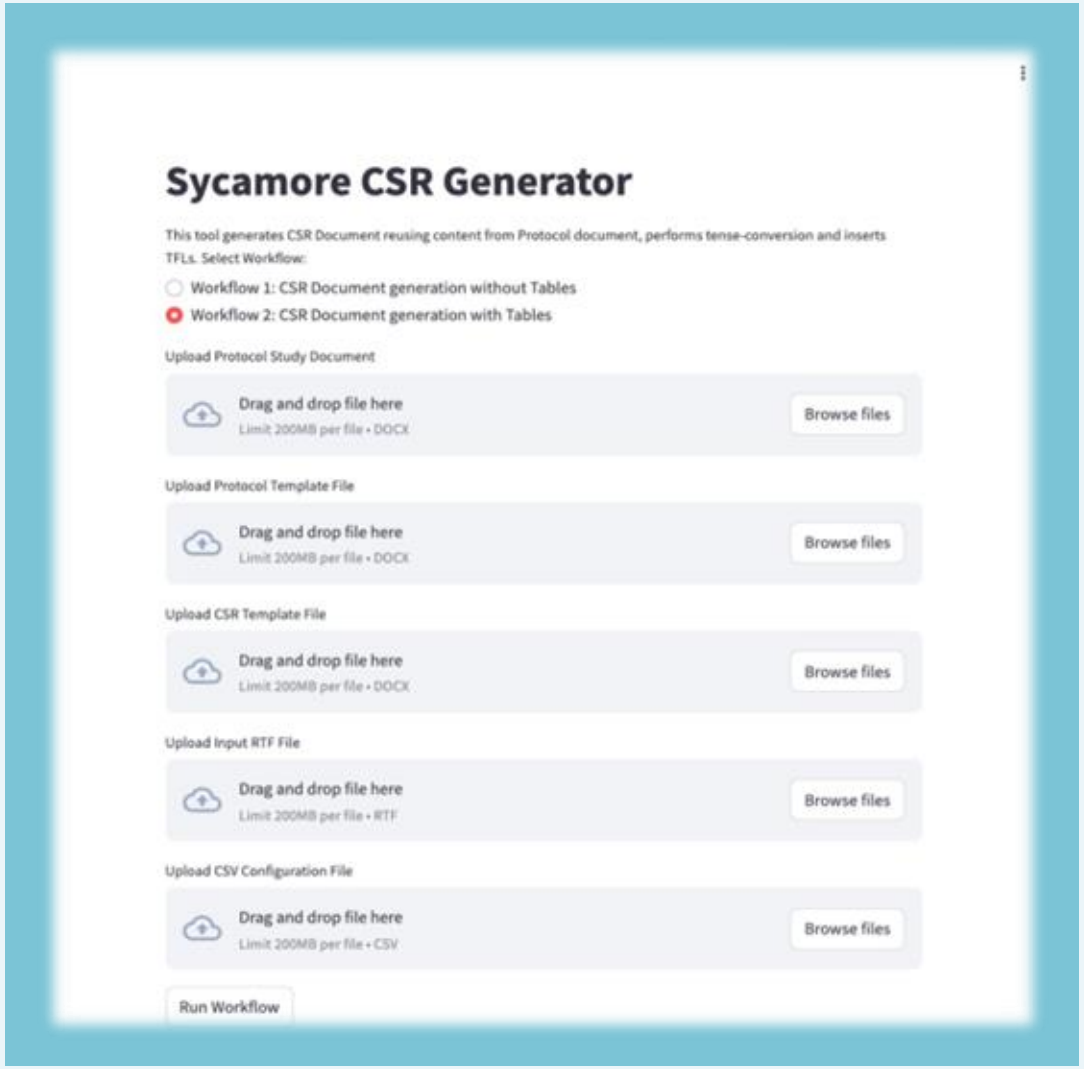
Public

How it Works

Step 3: Tense Conversion

One of the most tedious tasks in CSR drafting is manually converting protocol language from future tense to past tense across hundreds of pages.

Our solution automates this using a secure, enterprise-grade LLM that performs context-aware tense conversion while maintaining regulatory precision.



Public

How it Works

Step 3: Tense Conversion

Protocol Document (Future Tense)

3.1.2 Lead-in Period

During the Lead-in Period, subjects will have a study visit approximately 1 week before randomization (Visit 3, Day –7 ± 4 days). The subject's eligibility to continue into the Treatment Period will be assessed according to Randomization Criteria (Section 4.1.3). Assessments not satisfying the Randomization Criteria can be reassessed during the Lead-in Period if the Investigator has reason to believe the failed criteria(on) may have changed on repeat assessment. If all Randomization Criteria are satisfied, the subject will be eligible for randomization in the Treatment Period (Visit 4, Day 1).

3.1.3 Treatment Period

The Treatment Period will include 12 weeks of approximately monthly (every 4 weeks) administration of IP. Using an interactive response technology (IRT) system, eligible subjects will be randomized (1:1:2) to receive blinded IP (8 mg/kg CSL346, 16 mg/kg CSL346, or placebo).

Each subject's first dose of IP at Visit 4 (Day 1) will be an IV loading dose of 3 mg/kg CSL346 (subjects randomized to 8 mg/kg CSL346), 6 mg/kg CSL346 (subjects randomized to 16 mg/kg CSL346), or placebo (subjects randomized to placebo). All IV loading doses will be in a total volume of 10 mL and infused over the course of approximately 20 minutes. The subject will then be monitored for approximately 2 hours, during which AEs and vital signs will be recorded periodically as described in the Schedule of Assessments and Section 8.1.1. If no clinically relevant AEs are observed during this time, the subject's first SC dose of IP will be administered as an 8 mg/kg CSL346, 16 mg/kg CSL346, or placebo 20 mL SC infusion over the course of approximately 40 minutes.

Subjects will receive 3 subsequent SC infusions at Visit 7 (Week 4), Visit 8 (Week 8), and Visit 9 (Week 12) for a total of 4 SC doses.

"During the Lead-in Period, subjects will have a study visit approx…"

"The Treatment Period will include 12 weeks of approx…"

"The subjects will then be monitored for approximately 2 hours…"

Public

Step 3: Tense Conversion

3.1.2 Lead-in Period

During the Lead-in Period, subjects will have a study visit approximately 1 week before randomization (Visit 3, Day –7 ± 4 days). The subject's eligibility to continue into the Treatment Period will be assessed according to Randomization Criteria (Section 4.1.3). Assessments not satisfying the Randomization Criteria can be reassessed during the Lead-in Period if the Investigator has reason to believe the failed criteria(on) may have changed on repeat assessment. If all Randomization Criteria are satisfied, the subject will be eligible for randomization in the Treatment Period (Visit 4, Day 1).

3.1.3 Treatment Period

The Treatment Period will include 12 weeks of approximately monthly (every 4 weeks) administration of IP. Using an interactive response technology (IRT) system, eligible subjects will be randomized (1:1:2) to receive blinded IP (8 mg/kg CSL346, 16 mg/kg CSL346, or placebo).

Each subject's first dose of IP at Visit 4 (Day 1) will be an IV loading dose of 3 mg/kg CSL346 (subjects randomized to 8 mg/kg CSL346), 6 mg/kg CSL346 (subjects randomized to 16 mg/kg CSL346), or placebo (subjects randomized to placebo). All IV loading doses will be in a total volume of 10 mL and infused over the course of approximately 20 minutes. The subject will then be monitored for approximately 2 hours, during which AEs and vital signs will be recorded periodically as described in the Schedule of Assessments and Section 8.1.1. If no clinically relevant AEs are observed during this time, the subject's first SC dose of IP will be administered as an 8 mg/kg CSL346, 16 mg/kg CSL346, or placebo 20 mL SC infusion over the course of approximately 40 minutes.

Subjects will receive 3 subsequent SC infusions at Visit 7 (Week 4), Visit 8 (Week 8), and Visit 9 (Week 12) for a total of 4 SC doses.

{{past:<copy_section>}}

Tense Conversion Placeholder

9.1.2 Lead-in Period

During the Lead-in Period, subjects had a study visit approximately one week before randomization (Visit 3, Day –7 ± 4 days). The subject's eligibility to continue into the Treatment Period was assessed according to Randomization Criteria. If assessments did not satisfy these criteria, they could be reassessed during the Lead-in Period if the Investigator believed that the failed criteria might have changed on repeat assessment.

If all Randomization Criteria were satisfied, the subject would have been eligible for randomization in the Treatment Period (Visit 4, Day 1).

9.1.3 Treatment Period

The treatment period included 12 weeks of approximately monthly administration of IP. Eligible subjects were randomized (1:1:2) using an interactive response technology system to receive blinded IP, which consisted of 8 mg/kg CSL346, 16 mg/kg CSL346, or a placebo.

Each subject's first dose of IP at Visit 4 (Day 1) was an IV loading dose of 3 mg/kg CSL346 for those randomized to 8 mg/kg CSL346, 6 mg/kg CSL346 for those randomized to 16 mg/kg CSL346, or a placebo for those randomized to the placebo. The IV loading doses were administered in a total volume of 10 mL and infused over approximately 20 minutes.

Following this, subjects were monitored for about two hours during which adverse events (AEs) and vital signs were recorded periodically. If no clinically relevant AEs were observed, the subject's first SC dose of IP was administered as an 8 mg/kg CSL346, 16 mg/kg CSL346, or placebo 20 mL SC infusion over approximately 40 minutes.

Subjects received a total of four SC doses: one at Visit 7 (Week 4), another at Visit 8 (Week 8), and two at Visit 9 (Weeks 11-12). The subsequent SC infusions were administered every 4 weeks.

Protocol Document (Future Tense)

CSR Document (Past Tense)

Public

How it Works

Step 3: Tense Conversion

CSR Document (Past Tense)

*"During the Lead-in Period, subjects **had** a study visit approx…"*

*"The Treatment Period **included** 12 weeks of approx…"*

*"**Following this**, subjects were **monitored** for about 2 hours…"*

9.1.2 Lead-in Period

During the Lead-in Period, subjects had a study visit approximately one week before randomization (Visit 3, Day –7 ± 4 days). The subject's eligibility to continue into the Treatment Period was assessed according to Randomization Criteria. If assessments did not satisfy these criteria, they could be reassessed during the Lead-in Period if the Investigator believed that the failed criteria might have changed on repeat assessment.

If all Randomization Criteria were satisfied, the subject would have been eligible for randomization in the Treatment Period (Visit 4, Day 1).

9.1.3 Treatment Period

The treatment period included 12 weeks of approximately monthly administration of IP. Eligible subjects were randomized (1:1:2) using an interactive response technology system to receive blinded IP, which consisted of 8 mg/kg CSL346, 16 mg/kg CSL346, or a placebo.

Each subject's first dose of IP at Visit 4 (Day 1) was an IV loading dose of 3 mg/kg CSL346 for those randomized to 8 mg/kg CSL346, 6 mg/kg CSL346 for those randomized to 16 mg/kg CSL346, or a placebo for those randomized to the placebo. The IV loading doses were administered in a total volume of 10 mL and infused over approximately 20 minutes.

Following this, subjects were monitored for about two hours during which adverse events (AEs) and vital signs were recorded periodically. If no clinically relevant AEs were observed, the subject's first SC dose of IP was administered as an 8 mg/kg CSL346, 16 mg/kg CSL346, or placebo 20 mL SC infusion over approximately 40 minutes.

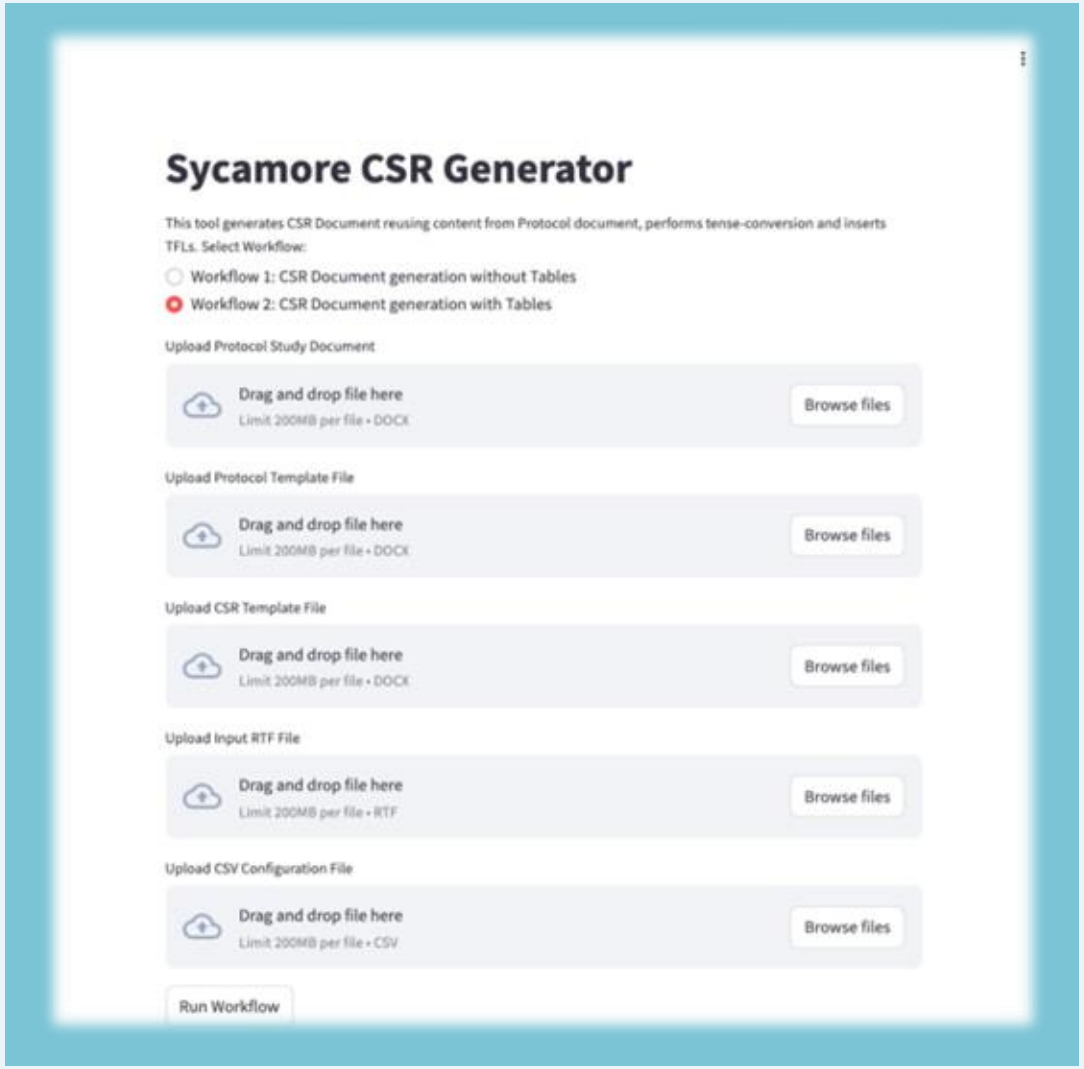
Subjects received a total of four SC doses: one at Visit 7 (Week 4), another at Visit 8 (Week 8), and two at Visit 9 (Weeks 11-12). The subsequent SC infusions were administered every 4 weeks.

Public

How it Works

Step 4: Multi-Page RTF Tables

- The framework intelligently handles complex, multi-page statistical outputs.
- Tables are embedded programmatically into the CSR template—preserving original formatting, column alignment, and pagination without manual intervention.



Public

How it Works

Table 14.1.1.1.1
Summary of Subject Disposition, Screening Period, Lead-in Period, Treatment Period, Follow-up Period
(Screened Analysis Set)

	CSL346 8 mg/kg n (%)	CSL346 16 mg/kg n (%)	Overall CSL346 n (%)	Placebo n (%)	Total n (%)
Study period: Lead-In Period					
Subjects entering lead-in period [1]					123 (37.6)
Lead-in failure [1]					11 (3.4)
Failure to meet randomization criteria [1]					4 (1.2)
Adverse event [1]					0
Pregnancy [1]					0
Death [1]					0
Lost to follow-up [1]					0
Protocol deviation [1]					0
Physician decision [1]					0
Withdrawal by subject [1]					4 (1.2)
Site terminated by sponsor [1]					0
Study terminated by sponsor [1]					0
Other [1]					3 (0.9)
Subjects randomized [1]	29 (8.9)	29 (8.9)	58 (17.7)	56 (17.1)	114 (34.9)
Subjects randomized but not treated with IP [2]	0	0	0	2 (3.6)	2 (1.8)
Subjects entering treatment period and treated with IP [2]	29 (100.0)	29 (100.0)	58 (100.0)	54 (96.4)	112 (98.2)

[1] Percentage based on number of subjects screened.
[2] Percentage based on number of subjects randomized to each treatment group.
[3] Data entered as free text allows identification of discontinuations due to COVID-19.

Source listing: 16.2.1.1, 16.2.1.2.1-2
Data snap shot date: 16DEC2022,/opt/zfs002/prd/csbeh251750/stats/interim/prog/tables/tdisp01.sas/09FEB2023/13:23

Table 14.1.1.1.1
Summary of Subject Disposition, Screening Period, Lead-in Period, Treatment Period, Follow-up Period
(Screened Analysis Set)

	CSL346 8 mg/kg n (%)	CSL346 16 mg/kg n (%)	Overall CSL346 n (%)	Placebo n (%)	Total n (%)
Study period: Treatment Period and Follow-Up					
Subjects entering treatment period and treated with IP [2]	29 (100.0)	29 (100.0)	58 (100.0)	54 (96.4)	112 (98.2)
Subjects discontinued the study medication during treatment period [2]	2 (6.9)	2 (6.9)	4 (6.9)	7 (12.5)	11 (9.6)
Primary reason for premature discontinuation of study medication [2]					
Adverse event [2]	1 (3.4)	2 (6.9)	3 (5.2)	0	3 (2.6)
COVID-19 associated adverse event [2][3]	0	0	0	0	0
Pregnancy [2]	0	0	0	0	0
Death [2]	0	0	0	0	0
COVID-19 associated death [2][3]	0	0	0	0	0
Lost to follow-up [2]	0	0	0	1 (1.8)	1 (0.9)
Physician decision [2]	0	0	0	0	0
COVID-19 associated [2][3]	0	0	0	0	0
Protocol deviation [2]	0	0	0	0	0
Withdrawal by subject [2]	0	0	0	3 (5.4)	3 (2.6)
COVID-19 associated [2][3]	0	0	0	0	0
Site terminated by sponsor [2]	0	0	0	0	0
Study terminated by sponsor [2]	0	0	0	0	0
Pandemic related [2]	0	0	0	1 (1.8)	1 (0.9)
Other [2]	1 (3.4)	0	1 (1.7)	2 (3.6)	3 (2.6)

[1] Percentage based on number of subjects screened.
[2] Percentage based on number of subjects randomized to each treatment group.
[3] Data entered as free text allows identification of discontinuations due to COVID-19.

Source listing: 16.2.1.1, 16.2.1.2.1-2
Data snap shot date: 16DEC2022,/opt/zfs002/prd/csbeh251750/stats/interim/prog/tables/tdisp01.sas/09FEB2023/13:23

These multi-page tables are pasted in apt subsections under Section 14

Public

How it Works

Step 5: a) Generating In-Text Summary Table

These summary tables within the CSR body support narrative claims and enable efficient regulatory review—while detailed outputs provide the full picture.

Our framework automatically generates these in-text summaries from multi-page tables - Extracting key metrics based on configurable parameters and ensuring the right data appears in the right sections.

Public

How it Works

Table 14.1.1.1.1 Summary of Subject Disposition, Screening Period, Lead-in Period, Treatment Period, Follow-up Period(Screened Analysis Set)

	CSL346 8 mg/kg n (%)	CSL346 16 mg/kg n (%)	Overall CSL346 n (%)	Placebo n (%)	Total n (%)
Subjects entering lead-in period [1]					123 (37.6)
Lead-in failure [1]					11 (3.4)
Failure to meet randomization criteria [1]					4 (1.2)
Withdrawal by subject [1]					4 (1.2)
Other [1]					3 (0.9)
Subjects randomized [1]	29 (8.9)	29 (8.9)	58 (17.7)	56 (17.1)	114 (34.9)
Subjects randomized but not treated with IP [2]	0	0	0	2 (3.6)	2 (1.8)
Subjects entering treatment period and treated with IP [2]	29 (100.0)	29 (100.0)	58 (100.0)	54 (96.4)	112 (98.2)
Subjects discontinued the study medication during treatment period [2]	2 (6.9)	2 (6.9)	4 (6.9)	7 (12.5)	11 (9.6)
Adverse event [2]	1 (3.4)	2 (6.9)	3 (5.2)	0	3 (2.6)
Lost to follow-up [2]	0	0	0	1 (1.8)	1 (0.9)
Withdrawal by subject [2]	0	0	0	3 (5.4)	3 (2.6)
Pandemic related [2]	0	0	0	1 (1.8)	1 (0.9)
Other [2]	1 (3.4)	0	1 (1.7)	2 (3.6)	3 (2.6)
Subjects discontinued the study during treatment or follow-up period [2]	2 (6.9)	3 (10.3)	5 (8.6)	7 (12.5)	12 (10.5)
Adverse event [2]	0	2 (6.9)	2 (3.4)	0	2 (1.8)
Death [2]	1 (3.4)	0	1 (1.7)	0	1 (0.9)
Lost to follow-up [2]	0	1 (3.4)	1 (1.7)	2 (3.6)	3 (2.6)
Withdrawal by subject [2]	0	0	0	4 (7.1)	4 (3.5)
Pandemic related [2]	0	0	0	1 (1.8)	1 (0.9)
Other [2]	1 (3.4)	0	1 (1.7)	0	1 (0.9)
Subjects completing the study [2]	27 (93.1)	26 (89.7)	53 (91.4)	47 (83.9)	100 (87.7)

[1] Percentage based on number of subjects screened.

[2] Percentage based on number of subjects randomized to each treatment group.

[3] Data entered as free text allows identification of discontinuations due to COVID-19.

Step 5: a) Generating In-Text Summary Table

In-Text Summary Table Generated from Multipage RTF Tables

Public

How it Works

Step 5: b) Extracting Key Findings from Summary Table

The framework generates key findings—highlighting trends, treatment outcomes, and notable safety or efficacy signals directly from the data.

These AI-identified insights give medical writers a head start, reducing the interpretive burden of synthesizing raw statistics into meaningful narrative descriptions.

Public

How it Works

Step 5: b) Extracting Key Findings From Summary Table

Key Findings for the In-Text Summary Table in the previous slide

The study, which aimed to evaluate the efficacy of an experimental treatment for a specific disease, was conducted during the COVID-19 pandemic. The study identifier and details on subject eligibility are not provided in the given snippet. However, it appears that the sample size may have undergone changes, as indicated by the varying numbers of subjects screened and randomized to each treatment group.

Critical findings from the study include:

* The most common reason for premature discontinuation of the study was adverse events, which affected 6.9% to 3.4% of the subjects depending on the treatment group.
* COVID-19 associated adverse events were reported in a negligible number of cases across both treatment groups.
* Pregnancy and disease relapse were not cited as reasons for premature discontinuation in any of the study groups.
* Death was reported in only one case, affecting 3.4% of subjects in one of the treatment groups, with COVID-19 associated deaths being a rare occurrence.
* Non-compliance with the study drug and lack of efficacy were not significant factors leading to premature discontinuation in either group.
* Lost to follow-up was a concern in both treatment groups, affecting 3.4% and 1.7% of subjects respectively, while withdrawal by subject affected a higher number of participants (7.1%) in one group.
* Withdrawal by parent/guardian or meeting protocol specified withdrawal criteria were not significant factors in either study group.
* Subjects completing the study showed a high completion rate across both groups, with percentages ranging from 93.1% to 83.9%, indicating a good adherence and engagement of subjects throughout the study.

The findings suggest that adverse events, particularly COVID-19 associated incidents, were not a significant concern in this study. However, the impact of COVID-19 on subject retention rates is notable, with higher percentages of lost to follow-up and withdrawal by subject observed in one treatment group compared to another. Despite these challenges, both groups demonstrated good adherence to the study protocol, as reflected in high completion rates across both cohorts. The low incidence of serious adverse events associated with the disease or study-related factors further supports the safe conduct of this trial amidst the pandemic context.

Public

How it Works

Step 5: c) Set Up for Creating Summary Table and Key Findings

Prepare a configuration CSV

	A	B	C	D	E
1	table num	section	description	table	drop 0 rows?
2	14.1.1.1.1	<table_disposition_subjects>	y	y	n
3	14.1.3	<table_protocol_deviations>	y	n	
4	14.1.1.3	<table_itt_analysis>	n	y	
5	14.1.1.3,14.4.1.11	<desc_2_tables>	y	n	
6	14.1.1.1.2	<table_analysis_set>	n	y	y
7	14.1.4.1.1	<table_itt_demog>	y	y	
8	14.1.5.1	<table_med_history>	y	n	

Public

Step 5: c) Set Up for Creating Summary Table and Key Findings

Prepare a configuration CSV specifying :-

table num: *List of tables to summarize*

section: *Placeholder name in the CSR where the in-text summary and description will be inserted*

description (Y/N): *Whether to generate a description for the in-text summary table*

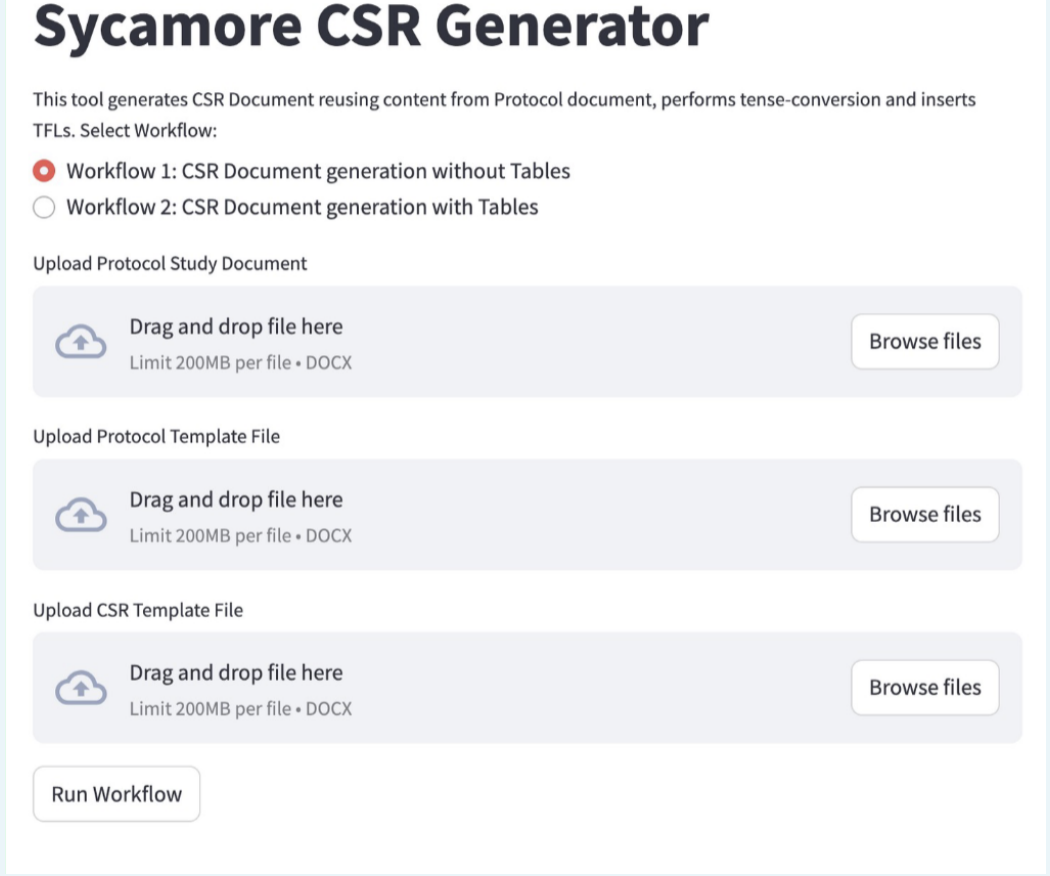
table (Y/N): *Whether to generate a summarized version of the tables*

drop 0 rows (Y/N): *Whether rows with zero counts should be included in the in-text summary table*

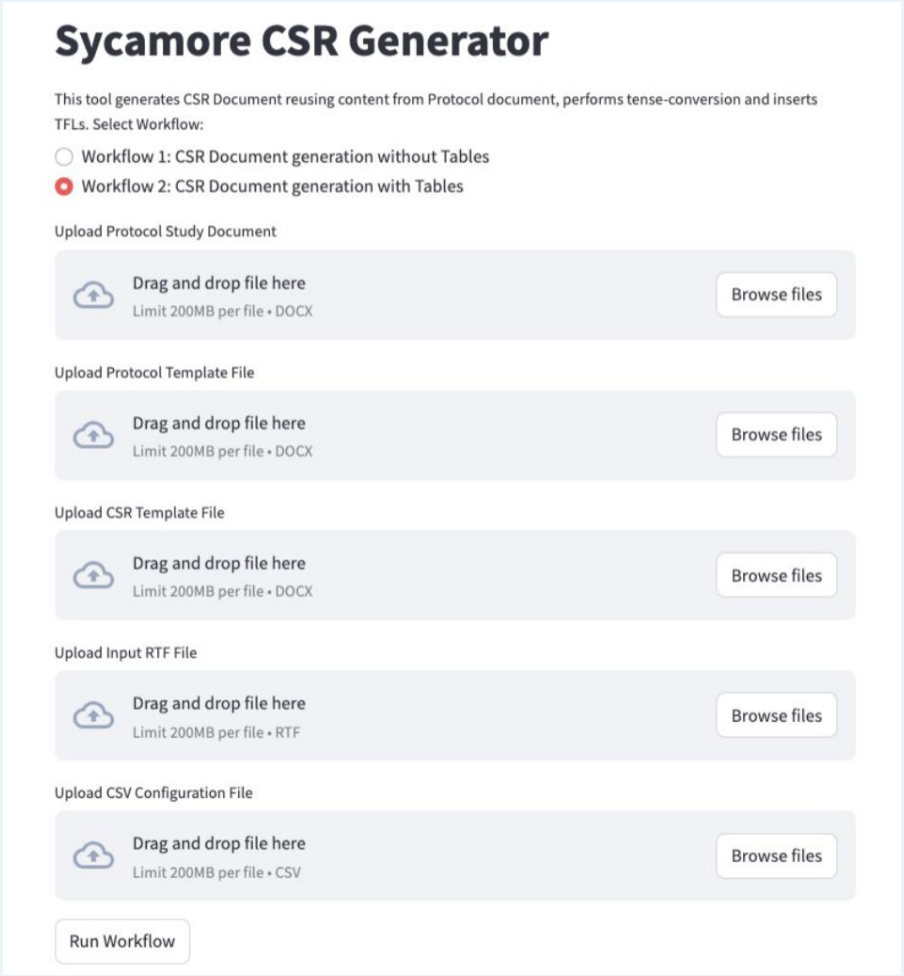
	A	B	C	D	E
1	table num	section	description	table	drop 0 rows?
2	14.1.1.1.1	<table_disposition_subjects>	y	y	n
3	14.1.3	<table_protocol_deviations>	y	n	
4	14.1.1.3	<table_itt_analysis>	n	y	
5	14.1.1.3,14.4.1.11	<desc_2_tables>	y	n	
6	14.1.1.1.2	<table_analysis_set>	n	y	y
7	14.1.4.1.1	<table_itt_demog>	y	y	
8	14.1.5.1	<table_med_history>	y	n	

Public

Workflows



Workflow 1: Without Tables



Workflow 2: With Tables

Public

Workflows

Workflow 1: Generate CSR Without Tables



(One Time Set Up)

(After Tense Conversion)

Required Files:

1. Protocol Study Document
2. Protocol Template
3. CSR Template

Public

Workflows

Workflow 2: Generate CSR With Tables



Required Files:

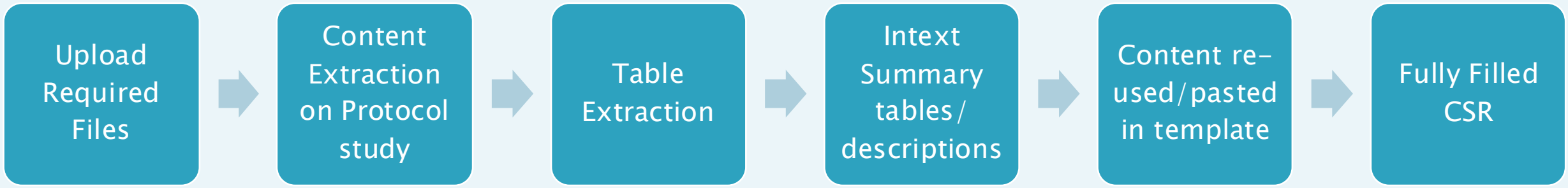
1. Protocol Study Document
2. Protocol & CSR Template
3. Multipage Tables RTF
4. CSV Config File

(After Tense Conversion)

Public

Workflows

Workflow 3: Update the Generated CSR Using Protocol With Tables



Required Files:

1. Pre-filled CSR without tables (Workflow 1)
2. Protocol Study Document
3. Protocol Template
4. Multipage Tables RTF
5. CSV Config File

(After Tense Conversion)

Public

Key Features

- A — Saves Time
- B — Automated Content Extraction
- C — Automated Tense Conversion
- D — Auto-Generated Key Findings
- E — In-Text Summary Table Generation
- F — Seamless multi-page table insertion

Public

Conclusion

An AI-powered CSR generator would transform clinical reporting by reducing weeks of manual work to mere minutes while ensuring accuracy and compliance.

As we continue to further improve our approach & processes, we're excited to bring even greater transparency, reliability, and speed to clinical studies

Public



Thank You!

QnA