Trial-Level AI-Generated Informed Consent Document Evaluation Summary Report (T-AICD)

Study and AI Parameters

Date: 01 Jan 2025

Study Title: ABC Trial in adults with moderate to severe lupus erythematosus

Protocol Number: ABC-XYZ-123

Version 1.0

Reviewers/Approvers:

Diane Spritz (Statistical Sciences Director) - Reviewer

John McAfee (Clinical Sciences Director) -Reviewer

Margaret Feder (Regulatory Sciences Director) - Reviewer

James Smart (Patient Advocacy Director) – Reviewer

AI Model(s) Used: Claude 2.0; Llama 2.0

Fine-Tuned or Base Model? Base models with RAG pipeline that integrates retrieval module indexing regulatory guidelines and past informed consent documents with a transformer-based generation module.

AI Supplemental Data Source(s):

RAG:

Acme Container 1.3.4

Acme Container 2.5.7

Inputs: Study details, patient population, historical consent documents, and regulatory documents.

Evaluation Purpose: Initial implementation

Final Intended Audience: Patients, clinicians

AI Model Risk Assessment

Model Influence

Medium:

Since the AI system generates the complete informed consent document its influence is high. However, human review ensures any discrepancies or ambiguities are addressed.

Decision Consequence

Medium:

Errors or omissions in the consent document could lead to inadequate participant understanding or ethical concerns. Rigorous human oversight mitigates this risk.

Overall Model Risk Rating

Medium:

While the system is solely responsible for drafting the initial informed consent document, the inclusion of comprehensive human review processes reduces the overall risk to medium.

Limitations and Potential Bias

There may be under representation of certain study types or participant demographics. Bias detection measures are outlined in the evaluation criteria.

AI models can produce erroneous information or fail to include important information for proper informed consent. Human review and oversight is required.

2.0 Evaluation

Assess the informed consent form for **readability, regulatory compliance, accuracy, and bias**.

Informed Consent Document (ICD) sections included in this review:

- Introduction
- Study Purpose & Procedures
- Risks & Benefits
- Alternative Treatment Options
- Confidentiality & Data Protection
- Voluntary Participation & Right to Withdrawal

Table 1.0 Categories, Metrics, and Allowable Thresholds

Category	Sub-Category	Metric	Criteria	Threshold	Result
Internal Consistency	Readability	Flesch-Kincaid cross-sectional comparison	Comparison across sections	6-9th grade across all sections	Confidentiality requires reduced complexity
	Factual	Named Entity Recognition	Drug names, trial phases, subject numbers	100% factual	Sample size missing; randomization and endpoint mismatch
	Conflicting medical terms & modifiers		"always", "never", "rarely" contradict earlier statements; The drug is safe for children vs. Children under 12 should not use this drug	0% contradictions	Pass
Compliance	FDA	Textual parsing	Does the content align with FDA guidelines and standards?	100%	
Accuracy	Scientific information check clinicaltrials.gov	SummaC	Unsupported claims	.90-1.0 indicating no factual issues detected and that the claim is well-	

				supported by evidence.	
Bias & Fairness	Objectivity of risks & benefits	Sentiment Analysis	Benefit-related versus risk-related sentiment distribution comparison		
Compliance		Textual parsing	Does the content align with FDA, EMA, and other relevant guidelines and standards (FDA, EMA, GxP)?		
Stylistic similarity to reference					
Semantic similarity	Content adequacy	BertScore		>.90	Requires

Table 2.0 Initial review results [add a table for each responsible party and remove the iteration documentation aspect]

Category	Category	Metric	Criteria	Reviewer/Approver Initials	Reviewer Notes
Internal Consistency	Readability	Flesch-Kincaid cross-sectional comparison	Comparison across sections		
	Factual	Named Entity Recognition	Drug names, trial phases, subject numbers		
	Conflicting medical terms & modifiers		"always", "never", "rarely" contradict earlier statements; The drug is safe for children vs. Children under 12 should not use this drug		
Accuracy	Scientific information check clinicaltrials.gov	FactCC			

	Text similarity to protocol	Cosine similarity	FILL OUT TABLE		
Bias & Fairness	Objectivity of risks & benefits	Sentiment Analysis	Benefit-related versus risk-related sentiment distribution comparison		
Compliance		Textual parsing	Does the content align with FDA, EMA, and other relevant guidelines and standards (FDA, EMA, GxP)?		

---

1. AI Model Performance Summary (add table for notes in appendix. Make one column for each function vertically.

Align with measures above	AI-Generated Document	Reference Document	Reviewer Notes
Readability Score			
Scientific Accuracy			
Compliance Pass Rate			
Bias/Coercion Flagged			
Hallucination Rate			

2. Reviewer/Approval Signatures

Version	Name	Title	**Function**	Signature	Date of Signature
1.0					**DD-MM-YYYY**

Python Code

Ensure Python 3.7+ is installed (mine 3.11.5)

#create environment

python -m venv myenv

#activate environment

myenv\Scripts\activate

#upgrade pip

#upgrade pip, setuptools, and wheel

pip install --upgrade pip setuptools wheel

#install spaCy

pip install spacy --only-binary :all:

#install English model

python -m spacy download en_core_web_sm

3. Key Evaluation Metrics and Review Process

Category	Evaluation Criteria	Quantitative Metric	Threshold for Acceptance	Reviewer(s)
Readability & Comprehension	Ensure patient-friendly language, avoiding jargon	Flesch-Kincaid Grade Level	Grade Level 6-8	Patient advocate, ethics committees
Scientific Accuracy	Alignment of study procedures, risks, and benefits with protocol	NLP similarity score with study protocol	>0.85 (high similarity with protocol)	Clinical SME, IRB reviewers
Regulatory Compliance	Adherence to FDA, EMA, ICH-GCP, and HIPAA	AI-assisted checklist completion rate	100% key compliance criteria met	Regulatory Affairs, Ethics board
Risk-Benefit Balance	Clarity and neutrality in presenting risks vs. benefits	Sentiment Analysis (neutral tone balance)	<10% sentiment bias	Ethics Committee, Legal review
Alternative Treatments	Clear disclosure of alternative treatment options	AI-extracted mentions of alternatives	At least on alternative mentioned	Medical writing, ethics
Bias & Fairness	Avoids coercive language, ensures fair representation	Bias detection model score	<5% deviation from historical trial demographics or diversity standards	ethics

Hallucination Rate	Instances where AI generates incorrect/unverifiable claims	% of fabricated content flagged by reviewers	<2% hallucination rate	Medical Writing, Clinical SME
Informed Decision-Making	Emphasizes voluntary participation and withdrawal rights	AI-check for presence of “voluntary” and “withdraw”	Both terms must be explicitly present	IRB, legal, patient advocate

Summary of AI-Generated Content

Protocol Section	AI Contribution (%)	Review Status	Comments
Study rationa			