Content Evaluation Guide

Umut Eser
Chief AI Officer

When evaluating content—whether produced by humans or AI—it is crucial to adopt a systematic and structured approach to ensure an objective and consistent assessment.

We have identified several categories that encompass the key aspects of technical writing for regulatory submissions related to therapeutics. Within each category we provide example metrics and a scoring rubric to promote consistency across reviewers. While these example metrics are not exhaustive and may not apply directly to all types of content (e.g., CMC), they are intended to be semi-quantitative and explainable. This rubric is most effective when applied to the smallest logical "unit" of writing, such as an individual study summary, and measured for numerous examples.

There is no universally "correct" method to measure the quality of technical content, but this framework serves as a valuable starting point. It is essential to adapt it, especially the specific metrics, to best align with your domain and needs. Ultimately, the most important aspect is to identify the criteria that are important to you and to articulate clear and objective means of measurement.

Score

Criteria	1	2	3	4
Accuracy	All written statements (sentences) were consistent with the report text or an explanation for the deviation was provided.	A few written statements (≤10%) were not consistent with the report text.	Several written statements (10-25%) were not accurately transcribed.	Many written statements (≥25%) were not accurately transcribed.
Completeness	All endpoints described in the study design were described in the results.	A few endpoints (≤10%) described in the study design were not presented in the results.	Several endpoints (10-25%) described in the study design were not presented in the results.	Many endpoints (≥25%) described in the study design were not presented in the results.
Redundancy	Details of the study design, results, or interpretations were presented only once.	Details of the study design, results, or interpretations were presented more than once, but appropriately (such as connecting related results from different endpoints or in a conclusion-type paragraph).	Details of the study design, results, or interpretations were presented more than once without adding clarity or emphasis (not counting a concluding statement).	N/A
Conciseness	No extraneous information (e.g., details irrelevant to the overall conclusion) was presented.	Little extraneous information was presented, but it did not distract from the overall emphasis or understanding.	Some extraneous information was presented, and distracted from the overall emphasis but not from the overall understanding.	Too much extraneous information was presented, and it confounded the overall emphasis or understanding.
Clarity	Language was direct, with few unnecessary words.	Language was mostly direct, but contained some unnecessary words, although it did not distract from the overall emphasis or understanding.	Language was fairly direct, but contained unnecessary words, which distracted somewhat from the overall emphasis or understanding.	Language was not direct and contained unnecessary words which distracted notably from the overall emphasis or understanding.
Emphasis	The level of detail emphasized the most pivotal data and conveyed an accurate interpretation of the key benefits or risks of the drug product.	The level of detail did not emphasize the most pivotal data and an accurate interpretation of the key benefits or risks of the drug product was not conveyed.	N/A	N/A