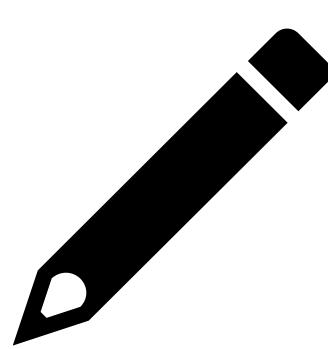




INTRODUCTION



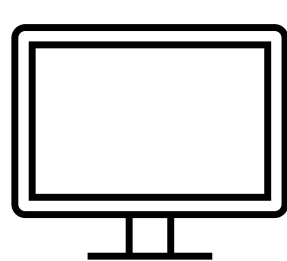
Quality of life (QOL) is a patient-centered outcome, providing insight into the physical, emotional, and social functional impact of both disease and treatment that extends beyond clinical endpoints such as survival and response rates.



The Functional Assessment of Cancer Therapy–Leukemia (FACT-Leu) scale is a widely used patient reported outcome (PRO) instrument. Developed and validated in 2012, FACT-leu offers a standardized, validated framework for assessing symptom burden and well-being in patients with leukemia.



Advances in leukemia treatment – including targeted agents, novel combination regimens, and changes in treatment duration and toxicity profiles – may affect the relevance of existing PRO tools.



Large language models (LLMs) offer a novel approach to assessing QOL tools. Our group has developed a quantitative analysis method to evaluate the potential for LLMs to update PRO tools and tested this for FACT-leu



AIM

1

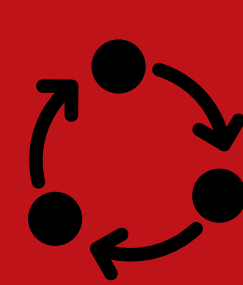
Evaluate the content and structure of the FACT-LEU using publicly available LLMs.

2

Assess LLM feedback on the FACT-Leu. Use qualitative analysis to review LLM analysis of the FACT-leu including issues related to the current patient experience with evolving leukemia treatment paradigms.

3

Establish a proof-of-concept methodology for using LLMs in the evaluation of disease-specific QOL instruments.



METHODS

1

Five publicly available LLMs were selected:



2

Each LLM received a standardized prompt to assess the FACT-Leu across four predefined domains: **Strengths of the instrument, Content gaps, Item redundancy and Suggested revisions**

3

Each LLM assigned a global quality rating on a **0–10 scale** to summarize its overall assessment of the instrument.



4

Responses were coded by two independent reviewers to identify **recurring themes, points of convergence/divergence** across models, and **notable insights** within each domain.



KEY FINDINGS

Model	Rating	Key Strengths	Key Gaps / Weaknesses	Recommendation
Google Gemini	8/10	- Holistic plus leukemia-specific “Additional Concerns” items (bleeding, bruising, fevers) 7-day recall suits rapid clinical change; simple, comparable scoring across cancers.	- Lacks financial-toxicity and cognitive (“chemo fog”) domains - Intimacy items risk high non-response or feeling intrusive.	Use as-is, adding a supplementary financial-toxicity module and minor wording updates to functional items for a broader age range.
Open Evidence	6/10	- Strong psychometrics - Validated across AML, CLL, CML, ALL, hairy-cell 10–15 min completion	- COSMIN-GRADE review rated content validity “inconsistent” - Missing cognitive, musculoskeletal, oral/mucosal, dermatologic, and financial-toxicity domains.	Use with modifications. Supplement with FACT-Cog and FACIT-COST for cognitive and financial-toxicity domains not currently covered.
ChatGPT	7/10	- Comprehensive multidomain structure with leukemia-specific symptoms - Short recall period - Strong legacy validation - Practical length and feasibility.	- Limited cognitive-function coverage insufficient treatment-burden granularity - Underrepresents financial toxicity/access - Possible ceiling/floor effects.	Use with targeted additions for cognitive function, financial toxicity, treatment-burden specificity, and work productivity
Anthropic Claude	7/10	- Comprehensive coverage plus 17-item leukemia subscale - Established psychometrics across subtypes	- No cognitive-functioning domain (“chemo brain”) - Single generic side-effect item can’t discriminate toxicity profiles - No financial-toxicity or care-access items.	Use with modifications: remove 4 redundant items, revise 5, and add 5 new items (cognition, finances, body image, adherence).
Perplexity	7/10	- Broad QOL coverage plus leukemia-specific symptom items - Validated psychometrics - Feasible 44-item length across acute and chronic leukemia	- Functions more as an HRQoL than a severity index - Missing financial toxicity, cognition, neuropathy, fertility, and caregiver-burden domains.	Use with modest modifications adding financial-toxicity, cognitive, treatment-logistics, and fertility/sexuality items.

Illustrative Domain Gaps



Cognitive burden

Fatigue, “chemo brain,” memory issues



Financial toxicity

Treatment costs, insurance challenges



Modern therapies

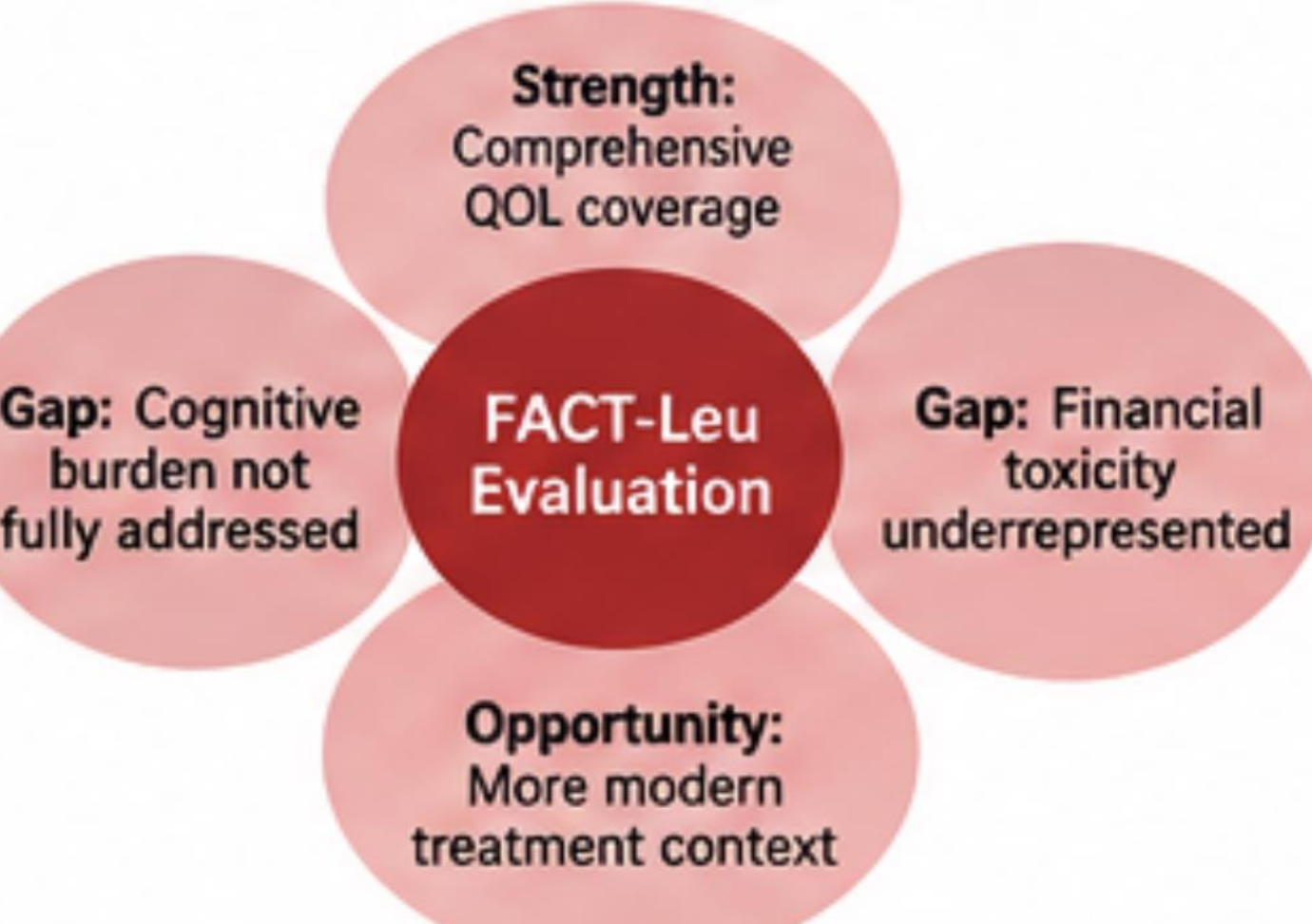
Targeted agents, combination regimens



Long-term survivorship

Quality of life beyond active treatment

Common Themes Across LLMs



CONCLUSIONS

- LLM-based evaluation identified potential areas for revision of the FACT-Leu, including greater consideration of **cognitive burden** and **financial toxicity**.
- Different LLMs may attend to different aspects of instrument content making a multi-model approach able to capture a broader range of potential concerns than any single model alone.
- These findings are exploratory and should not be interpreted as evidence for immediate modification of the instrument.
- Future studies incorporating patient perspectives and clinician input are needed to validate these findings and contextualize LLM-identified gaps before any instrument revision.

REFERENCES

- Functional Assessment of Chronic Illness Therapy (FACIT.org). FACT-Leu – Functional Assessment of Cancer Therapy – Leukemia. Available at: <https://www.facit.org/measures/fact-leu>
- Cella D, Jensen SE, Webster K, Du H, Lai JS, Rosen S, Tallman MS, Yount S. Measuring health-related quality of life in leukemia: the Functional Assessment of Cancer Therapy – Leukemia (FACT-Leu) questionnaire. Value Health. 2012;15(8):1051-1058. doi:10.1016/j.jval.2012.08.2210
- Cancer-related cognitive impairment: a downside of cancer treatment. Front Oncol. 2024;14:1387251. doi:10.3389/fonc.2024.1387251