AI-Assisted Approaches for Defining Lines of Therapy in Oncology Real-World Data: An Exploratory Analysis in Chronic Lymphocytic Leukemia

MSR218

Eric Chen, Catherine Fu, Keri Yang

BeOne Medicines, Inc., San Carlos, CA, USA

CONCLUSIONS

- LOT outputs remain sensitive to operational definitions and require expert adjudication, regardless of derivation approach
- AI-assisted LOT derivation offers a scalable and transparent alternative to manual rule-based methods, substantially reducing time burden and enabling systematic testing of methodological assumptions
- Accurate AI performance depends on clear specification of prompts and foundational regimen information. Differences in detecting treatment-switch events were a notable contributor to line-count discrepancies
- Findings reinforce the need for standardized LOT frameworks to enhance reproducibility and comparability of oncology real-world data analyses

BACKGROUND

- Accurate derivation of lines of therapy (LOT) is foundational for real-world oncology research, informing treatment pattern analyses, comparative effectiveness evaluations, and burden-of-illness studies
- LOT determination remains inconsistent and highly variable in practice, with no universally accepted standard algorithm
- AI-assisted approaches offer a potentially scalable and transparent alternative to labor-intensive manual rule-based methods

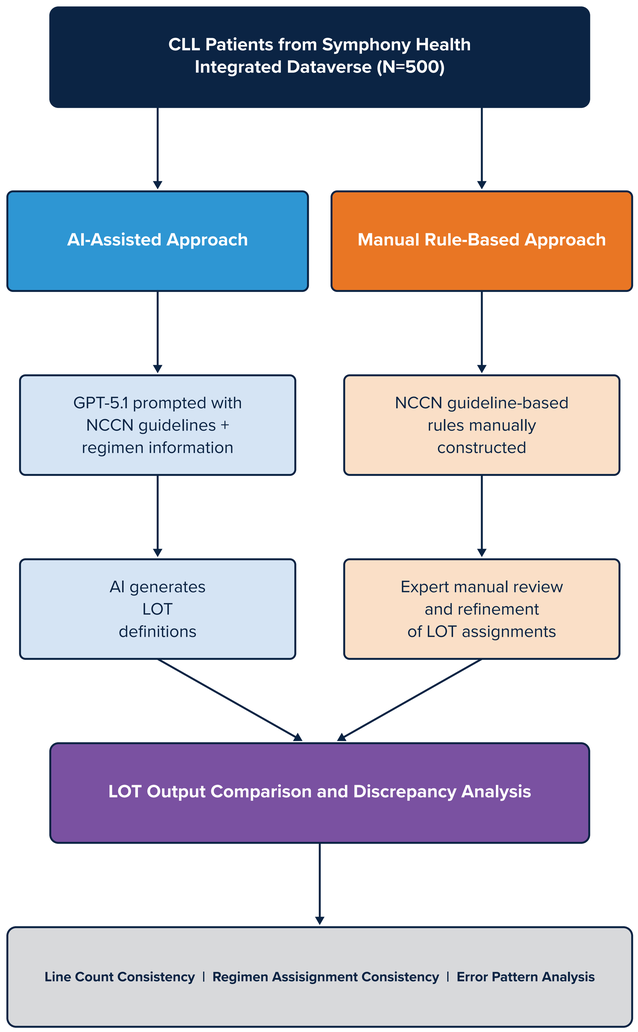
OBJECTIVES

- Examine AI-assisted and traditional manual LOT derivation approaches to understand how methodological choices impact LOT results
- Assess consistency in LOT assignments and treatment regimen classification between two independent approaches
- Identify and characterize sources of discrepancy between AI-assisted and manually curated LOT outputs

METHODS

- An exploratory comparison of two LOT approaches was applied to 500 randomly selected CLL patients from Symphony Health Integrated Dataverse®, a nationally representative US open-claims database
  - AI-Assisted Approach: GPT-5.1 prompted with NCCN guidelines and treatment regimen information to generate LOT definitions
  - Manual Rule-Based Approach: NCCN guideline-based rules manually constructed and further refined through expert review
- LOT outputs from both approaches were compared at the patient level for line-count consistency and regimen-level agreement

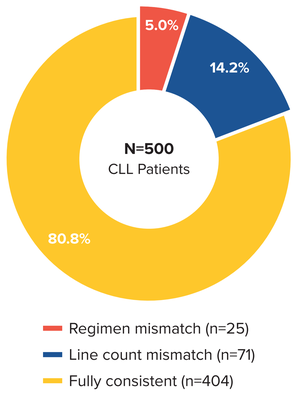
Figure 1. Study design and analytical framework



RESULTS

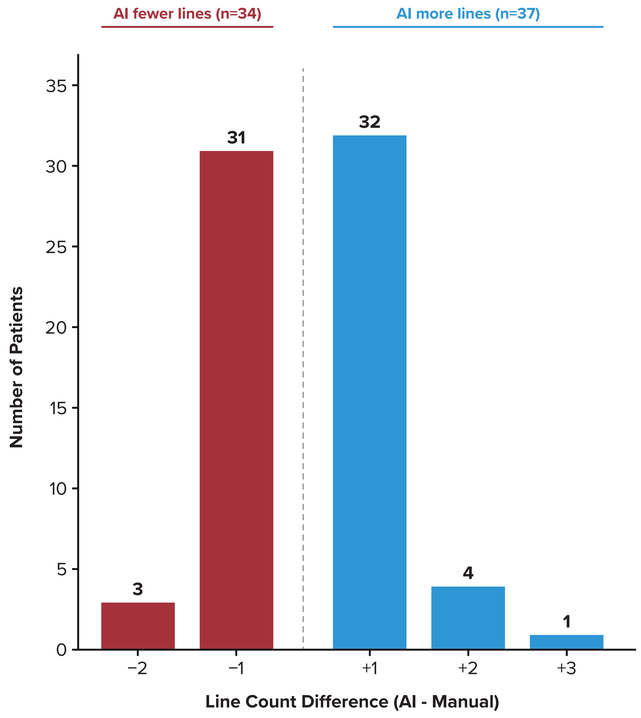
Overall Consistency

Figure 2. Patient-level LOT consistency classification (N=500)



Line Count Mismatches

Figure 3. Distribution of line count differences between AI and manual approaches (n=71 patients with mismatches)



Discrepancy Sources

Figure 4. Breakdown of discrepancy sources across mismatch categories

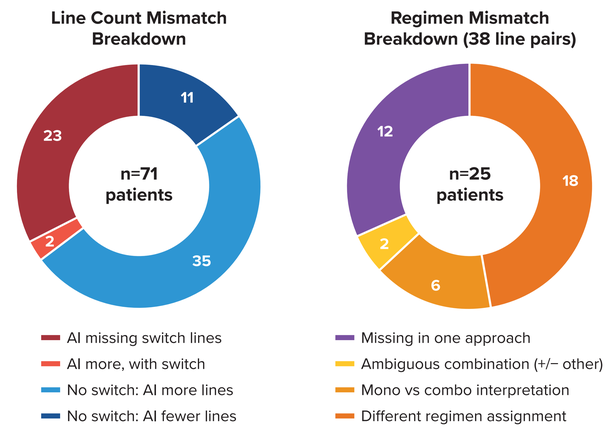


Table 1. Most common regimen mismatch patterns (generic names shown; 25 patients, 38 mismatched line pairs)

Manual (Ground Truth)	AI Assignment	n
Chlorambucil +/− other	Obinutuzumab + chemo	4
Zanubrutinib monotherapy	[Missing]	4
Venetoclax monotherapy	Venetoclax +/− other	2
[Missing]	Rituximab monotherapy	2
[Missing]	Ibrutinib monotherapy	2
[Missing]	Zanubrutinib monotherapy	2

REFERENCES

1. National Comprehensive Cancer Network. NCCN Clinical Practice Guidelines in Oncology: Chronic Lymphocytic Leukemia/Small Lymphocytic Lymphoma. Version 1.2026.
2. Symphony Health Integrated Dataverse. Accessed 2025.

DISCLOSURES

EC, CF, and KY are employees of BeOne Medicines and may hold stock or stock options.

ACKNOWLEDGMENTS

We would like to thank the investigators, site support staff, and especially the patients for participating in this study. This study is sponsored by BeOne. Medical writing support was provided by AMICULUM USA, with funding provided by BeOne.

CORRESPONDENCE: Catherine Fu, catherine.fu@beonemed.com

Presented at the International Society for Pharmacoeconomics and Outcomes Research Annual Meeting; Philadelphia, PA, USA; May 17–20, 2026