Evaluating Pain Scores and Detecting NSCLC Using Named Entity Recognition and Agentic AI in Clinical Texts

Somani M, Nayyar A, Daral S, Roy A, Markan R, Khan S, Verma V, Sachdev A, Goyal R, Seligman M, Brooks L

Objective

This study aimed to develop a Named Entity Recognition (NER) model to accurately detect Non-Small Cell Lung Cancer (NSCLC) and related concepts, as well as evaluate pain scores from unstructured clinical texts. Additionally, Agentic AI was used to extract and analyze biomarker data and recommend personalized treatment plans, thereby supporting improved patient management and outcomes.

Methodology

- The study utilized Optum's de-identified clinical notes spanning 2016 to 2023 to identify NSCLC diagnoses and associated pain scores.
- NSCLC diagnoses were confirmed using ICD-10 codes in the structured data and textual mentions in the clinical notes **(Figure 1)**.
- The remaining clinical notes were analyzed using an NER model built with a Char CNNs - BiLSTM - CRF architecture, trained to detect medical entities related to NSCLC and pain scores.
- Pain scores were extracted from a small subset of the original cohort using GPT-4 and their correlation across different stages was analyzed.
- Agentic AI using GPT-4 was then applied to extract the biomarker and respective result for NSCLC, as well as to suggest the treatment based on the NSCLC stage and biomarker result.
- NER Model performance was assessed using precision, recall, and F1-score metrics, ensuring robust validation of data extraction processes.

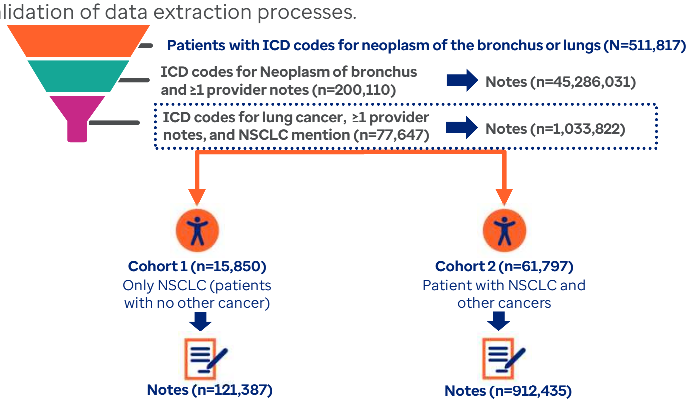

Figure 1. Cohort selection

Results

- The NER model showcased strong performance, achieving a precision of 89%, recall of 95%, and an F1-score of 92% in accurately identifying NSCLC-related entities (**Figure 2**).
- Pain score variability was observed to increase in patients with stage 3 and 4 NSCLC, reflecting greater heterogeneity in pain levels among advanced-stage patients (**Figure 3**). This variability may be influenced by factors such as tumor burden, metastasis, or treatment approaches.
- Additionally, the Generative AI agent leveraged GPT-4 to identify biomarkers and respective results with a 95% confidence score, as well as extract pain scores in the validation dataset.
- The Agentic AI framework was also able to suggest treatment approaches based on the stage and biomarker result of the patient on a smaller subset of data.

Char CNN: Character-level Convolutional Neural Networks
BiLSTM: Bidirectional Long Short-Term Memory
CRF: Conditional Random Field

Results

	Precision	Recall	F1 Score
NSCLC	0.9	0.95	0.92
NSCLC Location	0.81	0.75	0.78
NSCLC Histology	0.88	0.96	0.92
NSCLC Stage TNM	0.93	0.96	0.95
NSCLC Stage Numeric	0.96	0.96	0.96
NSCLC Stage Narrative	0.81	0.82	0.82

Figure 2. Performance metrics of NSCLC Entity Recognition

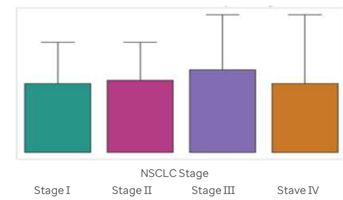

Figure 3. Pain score distribution across NSCLC stages

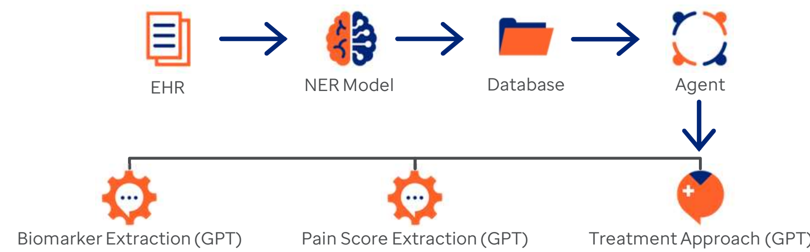

Figure 4. Agentic AI Framework for data extraction

Biomarker Result	Stage	Location	Suggested Treatment Approach from GPT
NA	1	Right lung	Surgical removal if resectable , Active Surveillance
N/A	2	Right lung	Surgical removal if resectable , Active Surveillance
PDL-1 Positive	3	Left Lung Upper Lobe	Concurrent chemoradiation therapy, Immunotherapy (e.g., PD-L1 inhibitors), Adjuvant chemotherapy post-surgery
BRAF positive	4	Both Lungs	Targeted Therapy (using BRAF inhibitors since positive), Palliative radiation therapy for symptoms

Figure 5. Agentic AI based treatment approach

Conclusions

- This study demonstrated the effectiveness of an AI-driven NER model in extracting NSCLC stages and pain scores from clinical texts.
- A preliminary proof of concept using Agentic AI showcased its potential in analyzing biomarker data and suggesting targeted treatments.
- These technologies highlight AI's ability to support personalized care for NSCLC patients. Future work could focus on refining and expanding this approach to other diseases, highlighting AI's transformative role in healthcare.

Optum