8947

# Natural Language Processing Of Imaging Reports to Identify Intracerebral Hemorrhage>1cm And Amyloid-related Imaging Abnormalities

Scan the QR code for a list of all Lilly content presented at the congress. Other company and product names are trademarks of their respective owners.

**Laura L. Schott[1], Zhun Cao[1], Sunday I. Ikpe[1], Ben Davis[1], Ann Hartry[2], Krista Schroeder[2]**

**[1]Premier, Inc., Charlotte, NC, USA; [2]Eli Lilly and Company, Indianapolis, IN, USA**

**Sponsored by Eli Lilly and Company**

## OBJECTIVES

1. Use imaging report narratives from US electronic medical records to develop NLP algorithms to identify:
   - Intracerebral hemorrhage (ICH) >1 cm
   - ARIA-E (edema)
   - ARIA-H (microhemorrhages)
   - ARIA-H (superficial siderosis)
2. Estimate algorithm performance using manual review of imaging reports.

## CONCLUSIONS

■ NLP algorithms for ICH >1 cm, ARIA-E and ARIA-H have been developed for use with real-world EMR imaging reports.

■ Validation work supports the utility of the ICH>1cm algorithm for use in ATT safety surveillance.

■ Additional validation work using a larger validation sample will occur in the coming months with the goal of including more positive ARIA cases.

■ This work will enable post-marketing real-world safety assessment of ATT treated patients that is not currently possible using structured codes alone.

Copyright ©2026 Eli Lilly and Company. All rights reserved.

**Alzheimer's Association International Conference 2026; London, United Kingdom; July 12-15, 2026**

## BACKGROUND

Patients on amyloid-targeting therapies (ATTs) require ongoing monitoring for three key safety outcomes detected on brain imaging:

■ Amyloid-related imaging abnormalities: edema/effusion (ARIA-E)

■ Amyloid-related imaging abnormalities: microhemorrhage(s) or superficial siderosis (ARIA-H)

■ Intracerebral hemorrhage (ICH) >1 cm — a clinically meaningful size threshold for radiologically severe intracerebral hemorrhage

In donanemab clinical trials, ARIA-E occurred in 15.6% of treated participants, ARIA-H in 25.5%%, and cerebral hemorrhage in 0.5%.[1] Real-world incidences are unknown and cannot be measured directly in large real-world databases because:

■ ICD-10-CM has no specific diagnosis codes for ICH >1 cm, ARIA- E, or ARIA-H.

■ These outcomes are recorded only in free-text report narratives.

■ Manual chart review does not scale to the populations needed for safety surveillance.

**Validated algorithms** to identify these outcomes in US Electronic Medical Records will facilitate safety surveillance of large populations of ATT treated patients in the real world.

[1]Wang H, et al. *The Journal of Prevention of Alzheimer's Disease*. Volume 12, Issue 8, September 2025, p100266

## METHODS

### Data Source:

■ Healthix Health Information Exchange (HIE) — largest US HIE; >9,000 healthcare facilities; 21M+ patients in the New York downstate region.

■ Study time period: 01 Jan 2022 – 28 Feb 2025

■ Independent samples used for development vs. validation steps.

### Algorithm Development and Validation Populations:

■ ARIA: Brain MRI reports from patients treated with ATTs during the study time period who were 65+ years of age.

■ ICH >1cm: Head CT + brain MRI reports from patients with Alzheimer's Disease or Related Dementias (ADRD) or Mild Cognitive Impairment (MCI) diagnoses during the study time period who were 65+ years of age.

- Enriched for patients with non-traumatic ICH ICD-10-CM codes in their structured EMR record.
- ATT use was not required.

### Reference Standard Generation:

■ Reference standard built by double, independent manual annotation of radiology narratives with adjudication of discrepancies by a physician radiologist subject-matter expert

■ Annotators were trained using clinical examples not included in the validation sample and were blinded regarding patient treatment status (ATT vs. no ATT). MRI reports were assessed for ARIA and ICH>1cm. CT reports were assessed for ICH>1cm.

■ Both outcome label and supporting text were captured in electronic Case Report Form

## RESULTS

### Reference Standard Validation Populations:

■ ARIA-E / ARIA-H: 68 patients, 140 imaging reports

■ ICH>1cm: 172 patients, 625 imaging reports

### Reference Standard Considerations

■ Generated using data through 28 Feb 2025

■ ICH>1cm positive cases (ADRD+MCI population): 82

■ ARIA positive cases (ATT population): 2 for ARIA-E, 9 for ARIA-H microhemorrhage, 0 for ARIA-H superficial siderosis

■ An additional year of data is actively being de-identified to execute a more robust validation (with the aim of including more ARIA positive cases).

**Table 1. NLP Algorithm Performance* (Preliminary - using data through Feb 2025)**

| Outcome | Se | Sp | PPV | NPV |
|---|---|---|---|---|
| ICH >1 cm | 98% | 99% | 93% | 100% |
| ARIA-E | 50% | 100% | 100% | 99% |
| ARIA-H micro-hemorrhage(s) | 67% | 100% | 100% | 97% |
| ARIA-H superficial siderosis | N/A | 100% | N/A | 100% |
| ARIA-H overall | 67% | 100% | 100% | 97% |

Se = Sensitivity, Sp = Specificity, PPV = Positive predictive value, NPV = Negative predictive value; N/A due to no positive cases identified.

*Adjudication label categories of "no", "unknown" and "size unknown" all were rolled up to "no" for algorithm performance assessment.

## Methods (continued)

### Algorithm Development Approach:

■ Keyword-based rules (e.g., hemorrhage, edema terms); see Table 2

■ Contextual modifiers (e.g., "no evidence", "suggestive of")

■ Anatomical and modality logic (e.g., brain regions, CT vs MRI)

### NLP + LLM (large language model) Framework

■ Model: Llama3 Instruct; Version: 3.3-70B; Prompt structure: Single shot with classification rules

■ A clinical NLP pipeline extracts structured information from text

■ The LLM applies rule-based logic to classify outcomes based on keywords, contextual modifiers and anatomical/modality logic.

■ Context-aware: Incorporates negation, uncertainty, and clinical context (e.g., "no evidence of", "possibly", "history of")

■ Considers both numerical and adjective descriptive of ICH size

■ Keywords identified using substring matching, allowing for misspellings and nuances in terms (e.g., hemorrhage, hemorrhagic).

■ Outputs are standardized as binary (Yes/No)

### Key Classification Rules

■ Conservative design: Requires explicit evidence for positive classification

■ ARIA evaluated only on MRI (CT scans default to No)

### Iterative Development with Clinical Expertise

■ Algorithms refined through multiple iterations

■ Clinical experts identified appropriate positive, negative, and modifying keyword terms

■ "Human-in-the-loop" process where mis-categorizations were reviewed, and algorithm was refined

**Disclosures:** K. Schroeder and A Hartry are employees and shareholders of Eli Lilly and Company. L.L. Schott, Z. Cao, S.I. Ikpe are employees of Premier, Inc., which received funding from Eli Lilly and Company to conduct this study. B. Davis was an employee of Premier Inc.

**Table 2: Natural Language Processing (NLP) Terms and Assignment**

| Outcome | Keywords (NLP Terms) |
|---|---|
| ARIA-E | vasogenic edema, vasogenic oedema, parenchymal edema, cerebral edema, sulcal effacement, edema, effusion, swelling, oedema, ARIA related edema, amyloid-related imaging abnormality, high signal on T2 weighted, hyperintensity T2, new T2 or sulcal effusion, FLAIR imaging, stable edema, no change in edema |
| ARIA-H with microhemorrhage | microhemorrhage, micro-hemorrhage, micro hemorrhage, microbleed, micro-bleed, cerebral microbleed, microbleeds, punctate hemorrhage |
| ARIA-H with superficial siderosis | superficial siderosis, superficial siderotic, hemosiderin deposits, hemosiderin deposition, hemosiderosis, hemosiderotic, subarachnoid hemorrhage (SAH) |
| ICH >1cm | ICH, intracerebral hemorrhage, intracavitary hemorrhage, hematoma, cortical hemorrhage, cerebellar hemorrhage, cerebellum hemorrhage, intracranial hemorrhage, intraparenchymal hemorrhage, parenchymal hemorrhage, hemorrhagic infarction, intra-axial hemorrhage, thalamic hemorrhage, aneurysmal subarachnoid hemorrhage, hemorrhagic stroke, brain hemorrhage, subcortical hemorrhage, gyriform regional hemorrhage, intercurrent hemorrhage, midline vermian hemorrhage, consistent with hemorrhage, hemorrhagic component, hemorrhagic blood products AND >1 cm, >10 mm, macro, medium, large |

### Label categories for adjudication:

- **Yes** = Outcome is present, clearly indicated in the free-text narrative
- **No** = Narrative notes state that the outcome is not present, does not meet criteria, is absent, or has cleared/resolved
- **Unknown** = Information regarding the outcome is not provided or is unclear
- **Size unknown** [ICH>1cm only] = Evidence of hemorrhage, but no indicator of size (numerical or descriptive) is present.

"No" and "Unknown" options distinguish when an outcome is "definitively no" versus "assumed no" because the outcome is not specifically mentioned, or it is referenced but without enough information