Poster number 199

# Identification of Generalized Myasthenia Gravis and Antibody Status Using Rules-Based Natural Language Processing Applied to Neurologist Clinical Notes

Jonathan Darer, MD, MPH[1]; Jacqueline Pesa, PhD, MPH[2]; Zia Choudhry, MD[3]; Alberto E. Batista, PharmD, MS[3]; Purva Parab, PhD[1]; Xiaoyun Yang, MS[1]; Raghav Govindarajan, MD[4]

[1]Health Analytics; [2]Janssen Scientific Affairs, LLC, Titusville, NJ, USA; [3]Janssen Pharmaceutical Companies of Johnson & Johnson; [4]Hospital Sisters Health Systems Medical Group

## INTRODUCTION

- Myasthenia gravis (MG) is a rare chronic autoantibody-mediated neuromuscular disorder affecting approximately 60,000 individuals in the United States[1]
- International MG guidelines from 2020 recommend clinical subtyping to support treatment planning and discussions with patients regarding prognosis[2]
- Evaluation of physician documentation of MG clinical subtypes in patient encounters is not well understood
- Natural language processing (NLP) has been used in analysis of unstructured clinician text,[3] but it has not been applied to the care of individuals with MG

## OBJECTIVE

- To assess the feasibility of identifying meaningful MG clinical subtypes from neurologist clinical progress notes

## METHODS

### Study design

- This was a retrospective, cross-sectional analysis of de-identified neurologist notes

### Data source

- Medical transcriptions from Amplity Insights were used, including the full text from dictated clinical notes covering 150,000 physicians from over 40,000 clinics and hospitals in the United States between 2017 and 2022
- Notes from non-neurologists and notes with non-standard formatting were excluded from the analysis

### NLP model

- A rules-based NLP model, developed using spaCy (an open-source NLP library), was used to create an analyzable dataset
- The performance of the NLP model was assessed against manual annotations using Prodigy

### Analysis

- Calculations were performed of summary rates of individuals with MG who had formal documentation of generalized MG (gMG), versus other forms of MG (ocular, bulbar, unspecified), and antibody status
- Summary rates of relevant MG symptoms that potentially identify gMG status were calculated
- The NLP model's performance was assessed using the F1 score—the harmonic mean of precision and recall (2 × Precision X Recall)/(Precision + Recall)

## RESULTS

- After applying rules to the medical transcription data including excluding notes by providers other than neurologists, 5183 notes were available for analysis (**TABLE 1**)
- A total of 3085 patients with MG were identified; most had unspecified clinical subtype (80.2%), followed by gMG (10.6%), ocular MG (8.2%), and bulbar MG (2.4%; **TABLE 2**) (Note: patients could be labeled with more than one subtype in different notes)
- For patients with unspecified clinical subtype, symptoms provided potential insight in 63% of patients (**TABLE 3**), as evidenced by presence of symptoms indicative of ocular MG or gMG
- The majority of patients did not have a documented MG antibody status (76.0%; **TABLE 2**)

**TABLE 1. Note attrition**

| Description of notes | Notes removed | Number of notes remaining |
|---|---|---|
| **All available notes that were labeled "myasthenia gravis" or "MG" by Amplity Insights** | 0 | 45,414 |
| **Neurologist notes** | | 6557 (100.0%) |
| Notes limited to a single patient and encounter | 37 | 6520 (99.4%) |
| Notes limited to formatting compatible with NLP analysis[a] | 416 | 6104 (93.1%) |
| Notes limited to patients with confirmed diagnosis of MG | 921 | 5183 (79.0%) |

[a]Notes with carriage returns in the middle of sentences or notes configured as a single paragraph (i.e., unformatted) were excluded to avoid misinterpretation.
**MG**, myasthenia gravis; **NLP**, natural language processing.

**TABLE 2. Characteristics of patients with MG**

| Population attribute, n (%) | Patients with MG (N=3085) |
|---|---|
| **Age** | |
| <40 years | 172 (5.6) |
| 40–64 years | 599 (19.4) |
| ≥65 years | 1083 (35.1) |
| Missing | 1231 (39.9) |
| **Clinical subtype** | |
| Unspecified | 2474 (80.2) |
| gMG | 326 (10.6)[a] |
| Ocular | 253 (8.2)[a] |
| Bulbar | 74 (2.4)[a] |
| **MG serology status** | |
| Seropositive for acetylcholine receptor antibody | 441 (14.3) |
| Seronegative | 224 (7.3) |
| MuSK antibody negative | 45 (1.5) |
| Seropositive for MuSK antibody | 29 (0.9) |
| Antibody status unspecified | 2346 (76.0) |

[a]Patients could be identified in more than one clinical subtype.
**gMG**, generalized myasthenia gravis; **MG**, myasthenia gravis.

**TABLE 3. Symptomatology among patients with MG with unspecified clinical subtype**

| Population attribute, n (%) | Patients with unspecified clinical subtype[a] (N=2474) |
|---|---|
| **MG symptoms in patients with unspecified clinical subtype** | |
| Ocular (e.g., diplopia, ptosis) | 1021 (41.2) |
| Bulbar (e.g., dysphagia, dysarthria)[b] | 749 (30.3) |
| Respiratory (e.g., dyspnea)[b] | 609 (24.6) |
| General weakness[b] | 280 (11.3) |
| Appendicular (e.g., lower extremity weakness)[b] | 205 (8.3) |
| Axial (e.g., neck weakness)[b] | 78 (3.2) |
| At least one MG symptom | 1567 (63.3) |

[a]Clinical subtype refers to the degree of spread of weakness to different parts of the body, as opposed to antibody status, which refers to the type of antibodies found on diagnostics (e.g., anti-acetylcholine receptor, anti-MuSK); individuals could have more than one MG symptom.
[b]Symptoms associated with gMG.
**gMG**, generalized myasthenia gravis; **MG**, myasthenia gravis.

## LIMITATION

- The NLP model was customized to the existing dataset, and the model's performance may differ for other datasets

## CONCLUSIONS

- A rules-based NLP model applied to neurologist clinical notes was able to identify MG clinical subtypes for >20% of patients
- Using identified MG symptoms, it was possible to infer a clinical subtype for an additional 63% of patients with unspecified MG
- Antibody status was documented in almost 25% of patients with MG in the neurologist clinical progress notes
- MG treatment is increasingly being determined by clinical subtype and antibody status. Symptom documentation can identify patients with gMG and more severe disease manifestations

## REFERENCES

**1.** Salari N, et al. *J Transl Med.* 2021;19:516.
**2.** Narayanaswami P, et al. *Neurology.* 2021;96:114–22.
**3.** Crema C, et al. *Front Psychiatry.* 2022;13:946387.

## ACKNOWLEDGMENTS

Janssen Scientific Affairs, LLC sponsored this research. Health Analytics, LLC received funding to conduct the study.

## DISCLOSURES

**JP**, **ZC**, and **AB** are employees of Janssen Scientific Affairs, LLC, **JD**, and **XY** are employees of Health Analytics. **PP** was an employee of Health Analytics at the time the study was conducted. **RG** is an advisor for Janssen Scientific Affairs, LLC, and is a speaker and advisor for Argenx, UCB, and Alexion.

Presented at MGFA AANEM 2023 | Phoenix, AZ, USA | November 1–4, 2023

AUTOANTIBODY: MG