

Leveraging NLP to characterize real-world triptan use and adherence from unstructured EHR

Katherine Brown, PhD, MSN, RN; Sarah Platt, MS; Agnes Pastwa, BS; Sunny Guin, PhD, Emily Webber, PhD

Disclosures: All authors are employees of Truveta Inc.

Background

Migraines represent a significant burden across health systems, with high direct and indirect costs due to underdiagnosis and inexact treatments. Triptans, while established as a standard acute therapy, exhibit variable real-world use and patient adherence. Traditional structured electronic health record (EHR) data often miss contextually rich information needed to understand treatment behavior at scale¹.

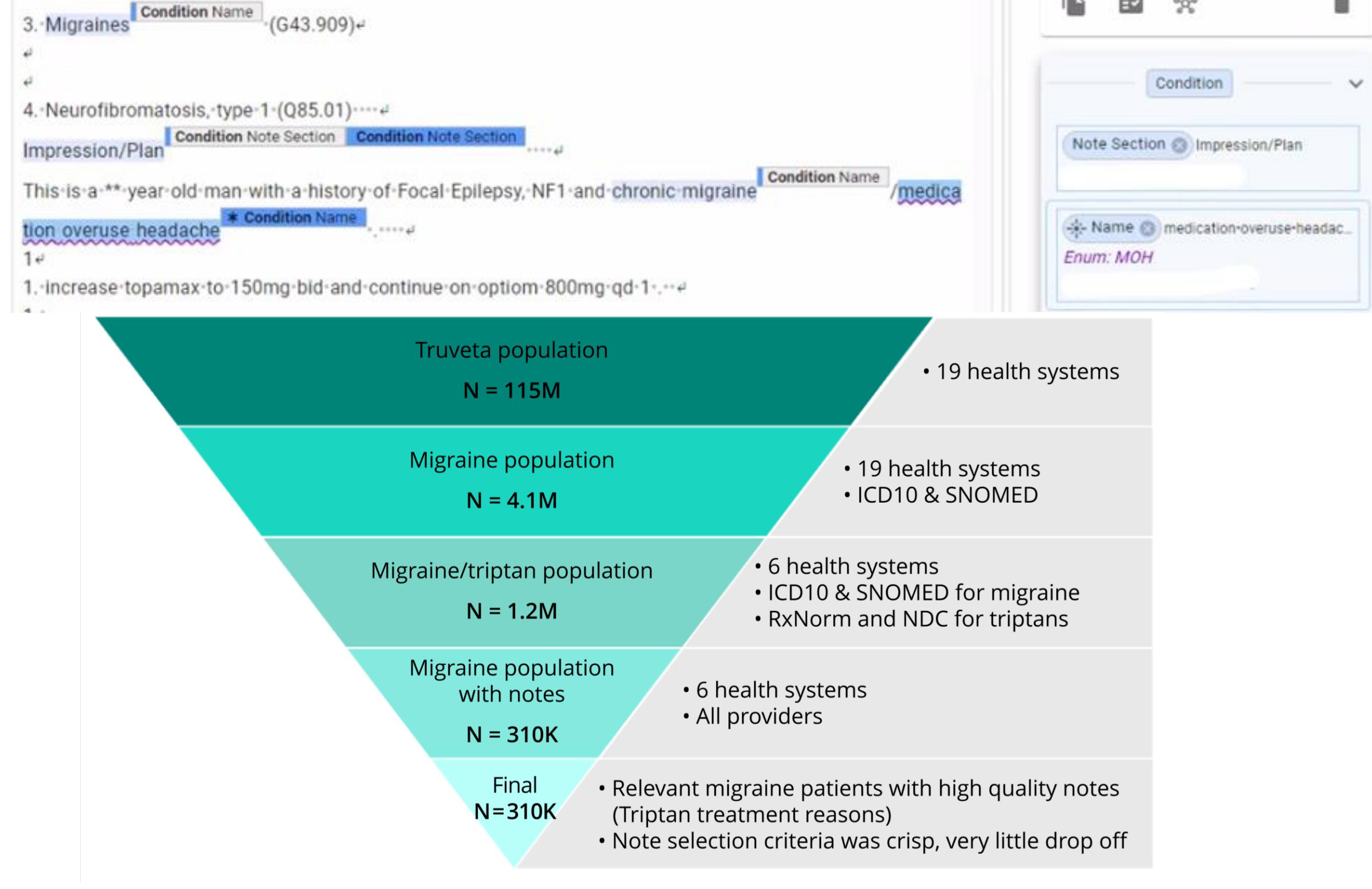
Objective

- Develop a natural language processing (NLP) model to understand triptan use and adherence from unstructured clinical notes.

Methods

- Truveta Data is comprised of **real-world US electronic health record (EHR)** data, which is aggregated, normalized, and de-identified from US health care systems comprising clinics and hospitals.
- Clinical notes associated with migraine or headache diagnoses and referencing triptan medications were selected. Truveta's proprietary language model (TLM) extracted medication mentions, associated treatment attributes, and adherence-related concepts.
- Extracted concepts were mapped to standard clinical ontologies using a zero-shot ontology normalization framework².
- Model performance was assessed using precision, recall, and accuracy, benchmarked against expert annotations.

Note annotation and cohort funnel for patient selection



Results

- The TLM achieved a precision of 86.8%, recall of 86.1%, and accuracy of 76.1%.
- The identified population was predominantly female (82.4%) and aged 30-49, aligning with known epidemiology.
- Notable disparities were observed in older and racially diverse populations, underscoring potential gaps in documentation and treatment access.
- Regarding triptan, 26% patients were found to have stopped treatment with ineffective (55%) and side effects (28%) cited by clinicians as the reasoning.

Conclusions

- This study highlights the value of applying advanced NLP to unstructured clinical data for real-world adherence insights.
- By enabling large-scale, high-fidelity capture of treatment behaviors, this approach offers new opportunities to inform health policy, optimize migraine care, and improve patient outcomes.
- This scalable ontology-aligned methodology supports applications in pharmacovigilance, patient segmentation, and population health management.

Results

Demography of patient's medication adherence and migraine/headache

		Triptan medication statement (Notes) N = 312,559	All migraine/headache patients N = 10M		
		N	%	N	%
Age	0-17	21,718	6.95%	537K	6.00%
	18-29	67,962	21.74%	1.7M	19.40%
	30-39	74,998	23.99%	1.7M	18.60%
	40-49	68,588	21.94%	1.6M	17.90%
	50-59	46,316	14.82%	1.3M	14.60%
	60-69	24,033	7.69%	1.1M	11.80%
	70-79	7,795	2.49%	721K	8.00%
	80+	1,143	0.37%	347K	3.70%
	Unknown	6	0.00%	0	0.00%
Sex	Female	257,382	82.35%	7.2M	69.80%
	Male	54,511	17.44%	3.1M	30.00%
	Other	62	0.02%	0	0.00%
	Unknown	604	0.19%	27K	0.20%
Race	American Indian or Alaska Native	2,571	0.82%	72K	0.70%
	Asian	8,785	2.81%	415K	4.00%
	Black or African American	28,169	9.01%	1.6M	14.30%
	Native Hawaiian or Other Pacific Islander	1,098	0.35%	61K	0.60%
	White	237,153	75.87%	6.7M	66.10%
	Declined to answer	15,767	5.04%	0	0.00%
	Other	19,016	6.08%	1.5M	14.30%
Ethnicity	Hispanic or Latino	36,980	11.83%	1.6M	15.00%
	Not Hispanic or Latino	258,870	82.82%	8.1M	77.70%
	Unknown	16,709	5.35%	761K	7.30%

Overview of triptan adherence status extracted from notes

	N Patients	%	N Rows	%
Total	312,559	--	2,390,155	--
Taking	294,047	94.08%	1,896,379	79.34%
Stopped	83,644	26.76%	488,221	20.43%
Not taking	1,598	0.51%	4,342	0.18%
On hold	433	0.14%	1,163	0.05%
Unknown	16	0.01%	50	0.00%

Triptan medication adherence reason extracted from notes

	N Patients	%	N Rows	%
Total	312,559	--	2,390,155	--
NULL (Taking or Not Available)	295,108	94.42%	1,912,834	80.03%
Doctor stopped drugs - ineffective	32,657	10.45%	165,288	6.92%
Drugs not taken/completed	17,054	5.46%	76,295	3.19%
Doctor stopped drugs - side effect	16,153	5.17%	86,436	3.62%
Medication stopped - ineffective	13,846	4.43%	52,456	2.19%
Medication discontinued	9,962	3.19%	26,672	1.12%
Drug declined by patient	7,692	2.46%	20,661	0.86%
Drug declined by patient - side effects	7,098	2.27%	20,130	0.84%
Medication stopped - contra-indication	3,525	1.13%	12,482	0.52%
Drug discontinued - too expensive	1,941	0.62%	5,443	0.23%
Drug not taken - side-effects	1,150	0.37%	4,123	0.17%
Treatment stopped - alternative therapy undertaken	751	0.24%	1,666	0.07%
Medication declined by patient - cannot pay script	547	0.18%	1,405	0.06%
Doctor stopped drugs - medical aim achieved	402	0.13%	863	0.04%
Drug declined by patient - patient beliefs	393	0.13%	884	0.04%
Drug treatment stopped patient ran out of medication	318	0.10%	747	0.03%
Drug not available for administration	214	0.07%	510	0.02%
Drug not taken - dislike form	156	0.05%	352	0.01%
Drug treatment no longer needed	102	0.03%	226	0.01%
Drug treatment stopped - medical advice	87	0.03%	197	0.01%
Drug declined by patient - dislike taste	81	0.03%	198	0.01%
Drug discontinued - patient fear/risk	41	0.01%	100	0.00%
Drug declined by patient - alternative therapy	24	0.01%	55	0.00%
Drug course completed	23	0.01%	50	0.00%
Drug not taken - patient lost tablets	13	0.00%	32	0.00%
Drug declined by patient - problem swallowing	8	0.00%	16	0.00%
Inappropriate medication stopped	5	0.00%	12	0.00%
Doctor stopped drugs - avoid interaction	4	0.00%	8	0.00%
Drug declined by patient - inconvenient	2	0.00%	4	0.00%
Drug not taken - dislike taste	2	0.00%	4	0.00%
Drug discontinued - reason unknown	1	0.00%	2	0.00%
Doctor stopped drugs - inconvenient	1	0.00%	2	0.00%
Drug declined by patient - reason unknown	1	0.00%	2	0.00%

Truveta has combined structured EHR data with clinical insight from patient notes using advanced NLP to better understand treatment behavior of migraine patients to improve outcomes.

