

Documentation of ENDS, Nicotine Pouch, & Combustible Tobacco Product Use in Electronic Health Record Clinical Notes

Abstract

We conducted a comprehensive assessment of the documentation and distribution of terms related to ENDS, nicotine pouches, and combustible tobacco use in unstructured electronic health records (EHR) clinical notes to assess the feasibility of conducting real-world evidence studies using retrospective data from Optum’s EHR and Clinical Notes Databases (2007-2023). There were 4.4 million patients with ≥1 diagnosis code for combustible tobacco use history and ≥1 clinical note containing a term for ENDS, nicotine pouches, or combustible tobacco, with 66.1 million notes. The terms used evolved over time concurrently with the dramatic increase in their documentation in recent years. Extensive ENDS use data are available in unstructured EHR clinical notes that can be extracted for research; however, documentation of nicotine pouch use was very limited during the study period. When combined with additional data sources, these data may enable robust studies of the health impacts of ENDS use. Standardization and consistent documentation practices will further enhance the utility of these data for research.

Introduction

Real-world data from healthcare claims and EHR may enable a new approach to assess the health impact associated with the use of ENDS (1-4). However, due to the lack of standardized codes, the documentation of ENDS use is non-existent in healthcare claims data (5) and severely limited in the structured fields of EHR data (6, 7). Previous studies using data from EHR reported that documentation of ENDS use has increased substantially in recent years in the clinical notes and the terms used to describe ENDS are diverse and changed over time (8, 9). This study aimed to describe the documentation of ENDS, nicotine pouches, and combustible tobacco products use in the unstructured clinical notes among a population of patients with a history of combustible tobacco use.

Methods

This was a retrospective study using data from January 2007 through August 2023 (study period) from Optum® de-identified EHR data set (10) which is derived from the electronic health records of a network of healthcare provider organizations across the United States. The Optum EHR Database currently has more than 107 million unique patients across the United States, with an average of 56 months of observed data per patient with clinical notes database comprising a subset of more than 45 million patients (11). The study population consisted of all patients who met the following criteria: Patients with ≥ 1 EHR record with an International Classification of Diseases (ICD) diagnosis code for the history of combustible tobacco use and ≥ 1 clinical note with a term of interest related to ENDS, nicotine pouches, or combustible tobacco during the study period.

Results

There were 4.4 million patients with ≥1 clinical note containing a term for ENDS, nicotine pouches, or combustible tobacco, with 66.1 million notes. ENDS documentation was identified in 27.6% of patients and 11.7% notes, nicotine pouch documentation was identified in 0.03% of patients and 0.004% of clinical notes, and combustible tobacco documentation was identified in 78.8% of patients and 73.3% of clinical notes. Of the 77 ENDS-related terms identified, 31 were variations of “e-cigarette” (e-cig variant terms) and 46 were variations of “vape” (vape variant terms). An additional 27 ENDS brand terms were identified. The terms used evolved over time concurrently with the dramatic increase in their documentation in recent years. There were 8 terms and 10 brand names for nicotine pouches. 118 combustible tobacco-related terms were identified, with 49 and 34 indicating current and former cigarette smoking, respectively: 21 for cigar and 14 for water pipe use.

Conclusion

Extensive ENDS use data are available in the unstructured EHR clinical notes that can be extracted for research; however, nicotine pouch use documentation was very limited during the study period. When combined with additional data sources (structured EHR, claims, and survey linkage), these data may enable robust studies on the health impacts associated with the use of ENDS. Standardization and consistent documentation practices will further enhance the utility of these data for research.

Extensive ENDS use data are available in unstructured EHR clinical notes and, when combined with additional data sources (structured EHR, claims, and survey linkage), may enable robust studies on the health impacts associated with ENDS use.

Table 1. Patient Selection and Final Study Population Derived from Patients/Notes in the Full EHR/Notes Data Asset (2007-2023)

		PATIENTS** (n)	NOTES** (n)
EHR Population*	≥ 1 ICD code for combustible tobacco use	13,583,236*	
EHR Notes Population	≥ 1 clinical note	5,194,000	862,491,151
Study Population	≥ 1 note containing a term for ENDS, nicotine pouches, or combustible tobacco	4,371,006	66,101,820
	≥ 1 combustible tobacco term**	3,443,072 (78.8%)	48,432,066 (73.3%)
	≥1 ENDS term OR ≥1 ENDS brand term**	1,207,884 (27.6%)	7,738,198 (11.7%)
	≥1 nicotine pouch term OR ≥1 nicotine pouch brand term**	1,217 (0.03%)	2,653 (0.004%)

*EHR population (first row) is among the full set of patients/notes in the full EHR data asset. Each subsequent row is a subset of the prior row
**Subset of Study Population; patient and note counts are unique within each row; a patient/note may appear in more than one row

Table 2. Patient and Note Counts for ENDS-related Terms among the Study Population (2007-2023)

Description	Patients*	Notes*
≥1 Vape term	1,114,448	7,270,473
≥1 E-cig term	718,117	3,265,350
≥1 ENDS brand term	5,458	16,544
≥1 Vape term OR ≥1 E-cig term	1,205,838	7,727,143
≥1 Vape term OR ≥1 ENDS brand term	1,116,643	7,281,820
≥1 E-cig term OR ≥1 ENDS brand term	721,137	3,279,077
≥1 Vape term OR ≥1 E-cig term OR ≥1 ENDS brand term	1,207,884	7,738,198

*patient and note counts are unique within each row; a patient/note may appear in more than one row

Figure 1. Stacked Patient Count with E-Cigarette Variant Terms in Notes (2007-2023)

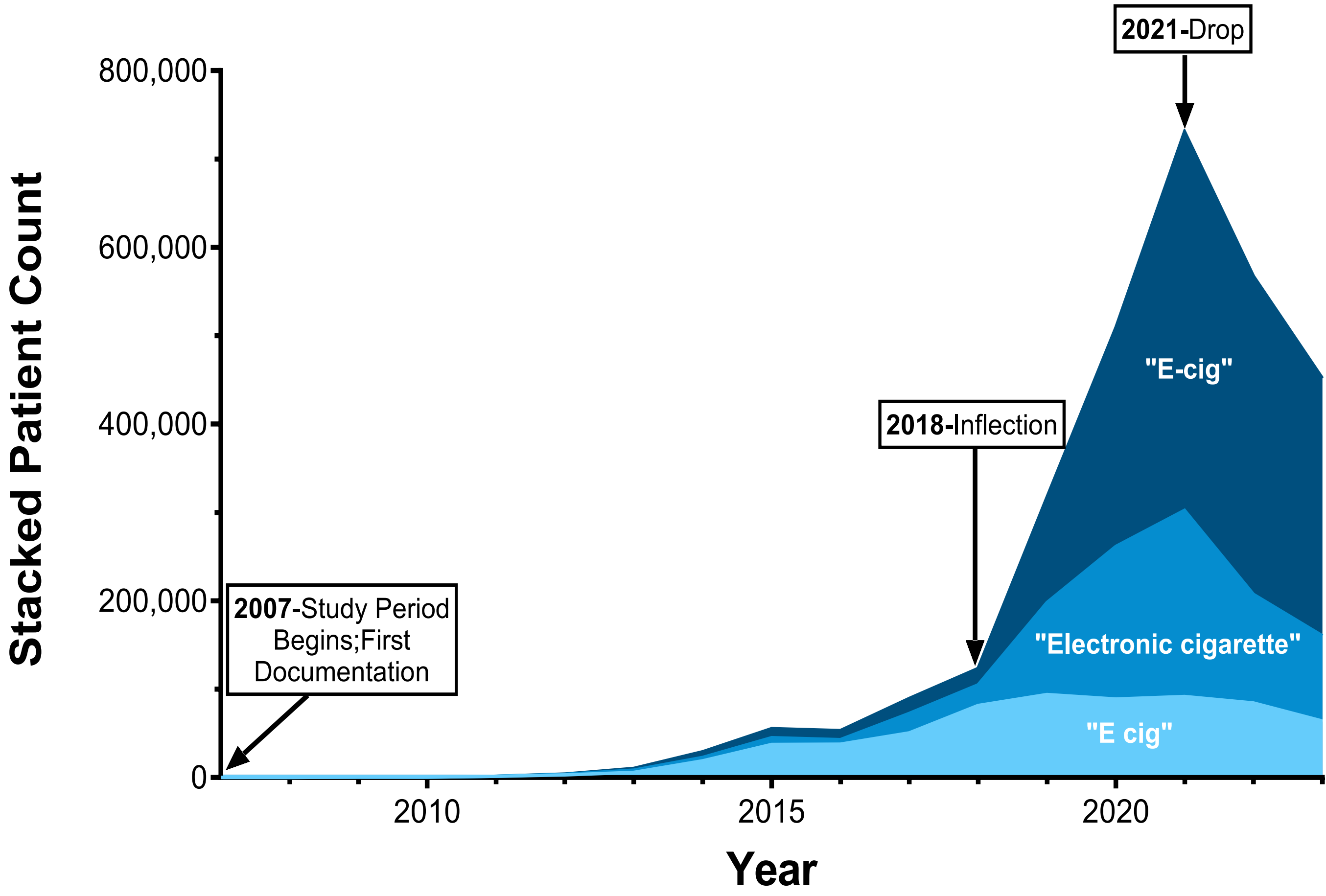
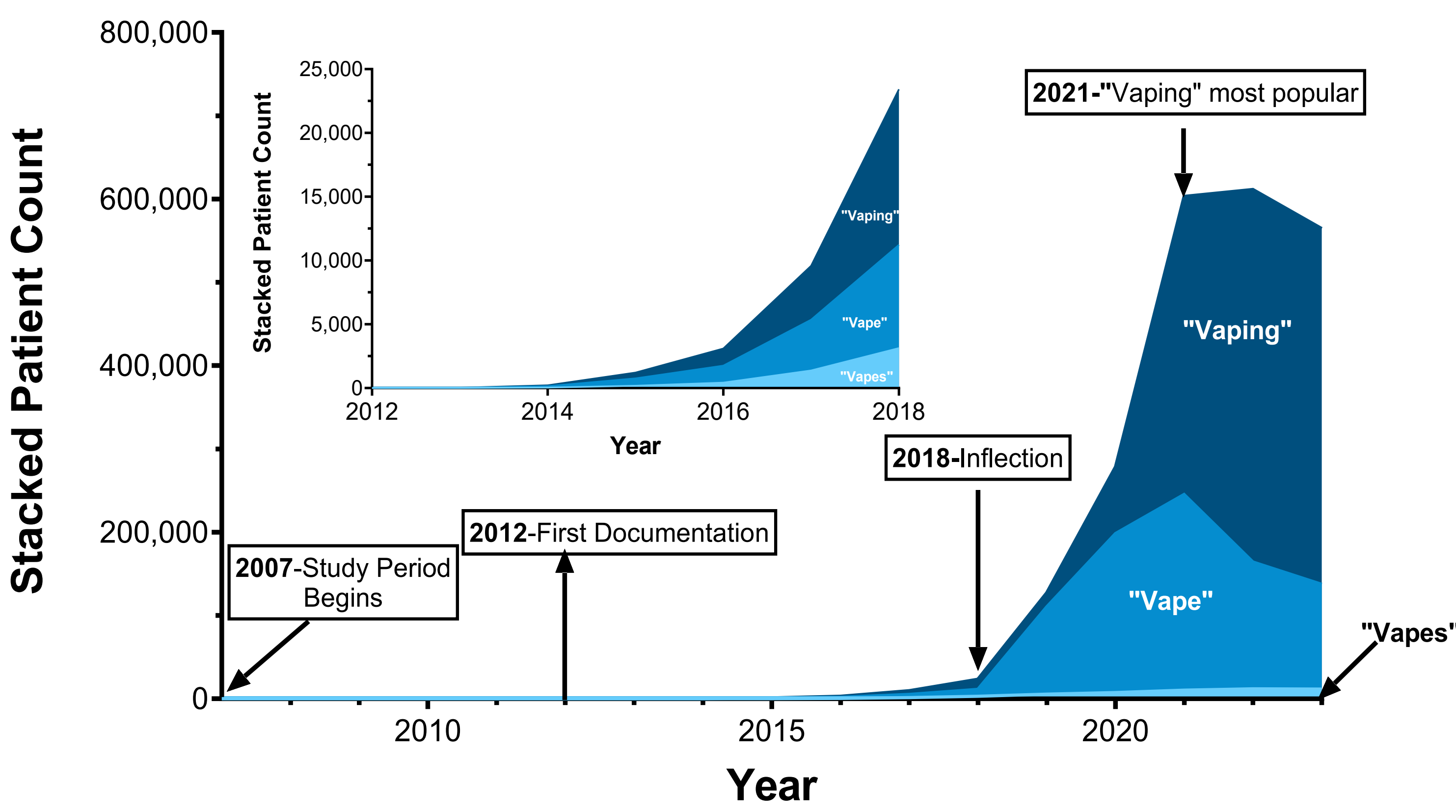


Figure 2. Stacked Patient Count with Vape Variant Terms in Notes (2007-2023)



References

1 Cohen G & Cook S. iScience. 2024.
2 Asfar T et al. Tobacco control. 2024;33(3):373-82.
3 Jackson SE et al. PLoS Mental Health. 2024;1(1):e0000007.
4 Zhang M et al. ISPOR 2023.
5 Zhang M et al. Harm Reduction Journal. 2024;21(1):227.
6 Jose T et al. Int J Environ Res Public Health. 2020;17(16).
7 Winden TJ et al. AMIA Jt Summits Transl Sci Proc. 2015;2015:199-203.
8 Conway M et al. AMIA Annu Symp Proc. 2021;2021:343-51.
9 Conway M et al. Am J Prev Med. 2019;56(3):474-5.
10 Optum. Gain visibility into patient care and outcomes with EHR data. 2024.
11 Guerra Currie A et al. Natural Language Processing Methods. 2023.

