

What Smokers Who Switched to Vapor Products Tell Us About Themselves

Presented by

Julie Woessner, J.D. – CASAA National Policy Director



The CASAA Consumer Testimonials Database

- Collection began in 2013 through the CASAA.org website
- As of April 2018, the database contained 11,236 testimonials
- All submissions are voluntary
- Entries are unstructured free-text – whatever the user writes
- The only moderation is for spam content which is removed
- Testimonials entries are uniquely coded to allow extraction
- Stored as part of the CASAA website's database

Study Design and Data Collection

- Initial study design by Dr. Brian Carter using small dataset and traditional topic extraction by independent human coders
- Proof-of-concept needed rapid data extraction without inter-rater reliability training
- Natural Language Processing (NLP) methods are able to do key topic extraction from a large corpus of information at least as accurately
- Database extraction was done using SQL methods to assure random sample
- Records were removed only for unrecognizable characters or empty entries
- Dataset size for NLP is 994 records of 1000 randomly extracted

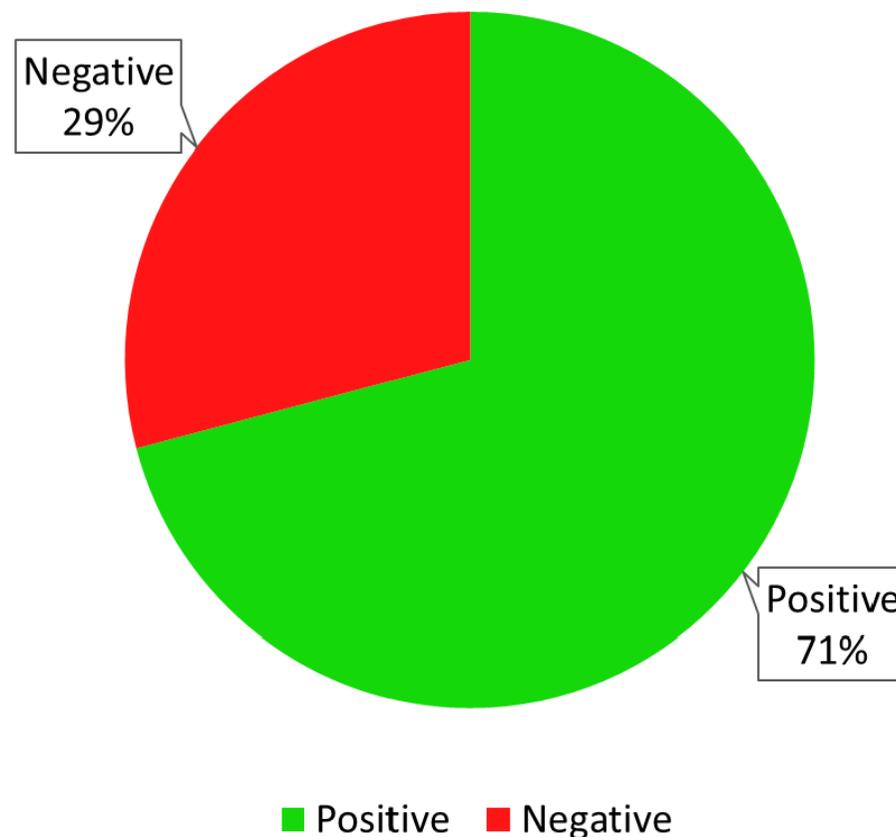
Natural Language Processing Techniques

- Mathematic analysis yields “key phrase in context” mapping of free-form text
- First technique is Sentiment analysis – based on word types used
- Second technique is Topic analysis – based on similarity to test phrase
- Essentially ideas are expressed in clusters of words
- NLP allows us to explore those clusters with mathematical precision
- Results were then processed in Excel to extract overall topics and detailed results
- Test phrases are included in the result presentations

Overall Results

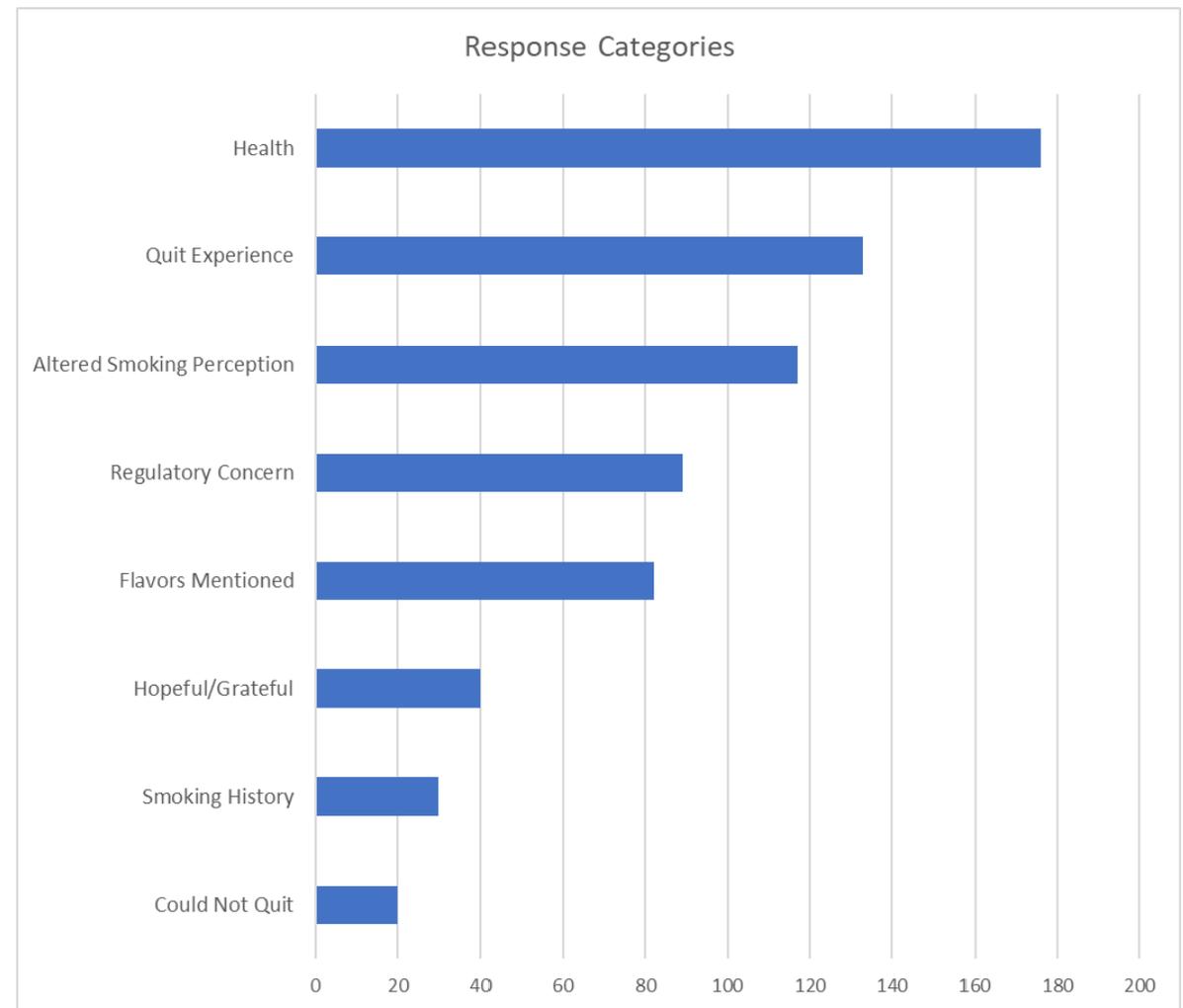
- Sentiment analysis looked at the overall tone of each testimonial
- Positive sentiment has more affirmative statements than negative sentiments
- Unsurprisingly, respondents were mostly posting with a positive sentiment

Composite Sentiment Analysis



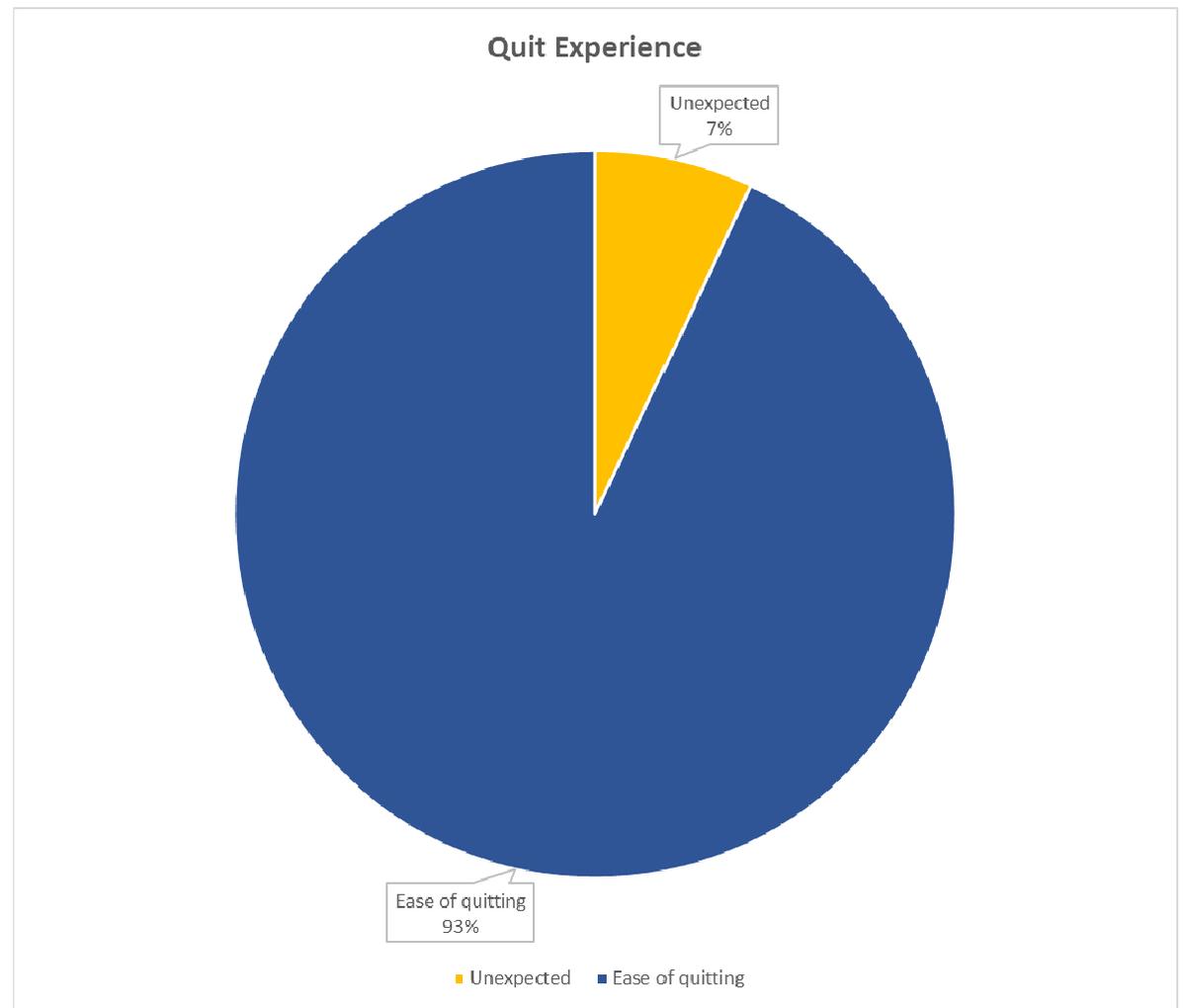
Overall Results

- Responses were gathered into broad categories of topics
- Health and Quit Experience were the most popular topics
- Altered Smoking perception was the 3rd popular topic
- Regulatory Concerns and Flavors were 4th & 5th



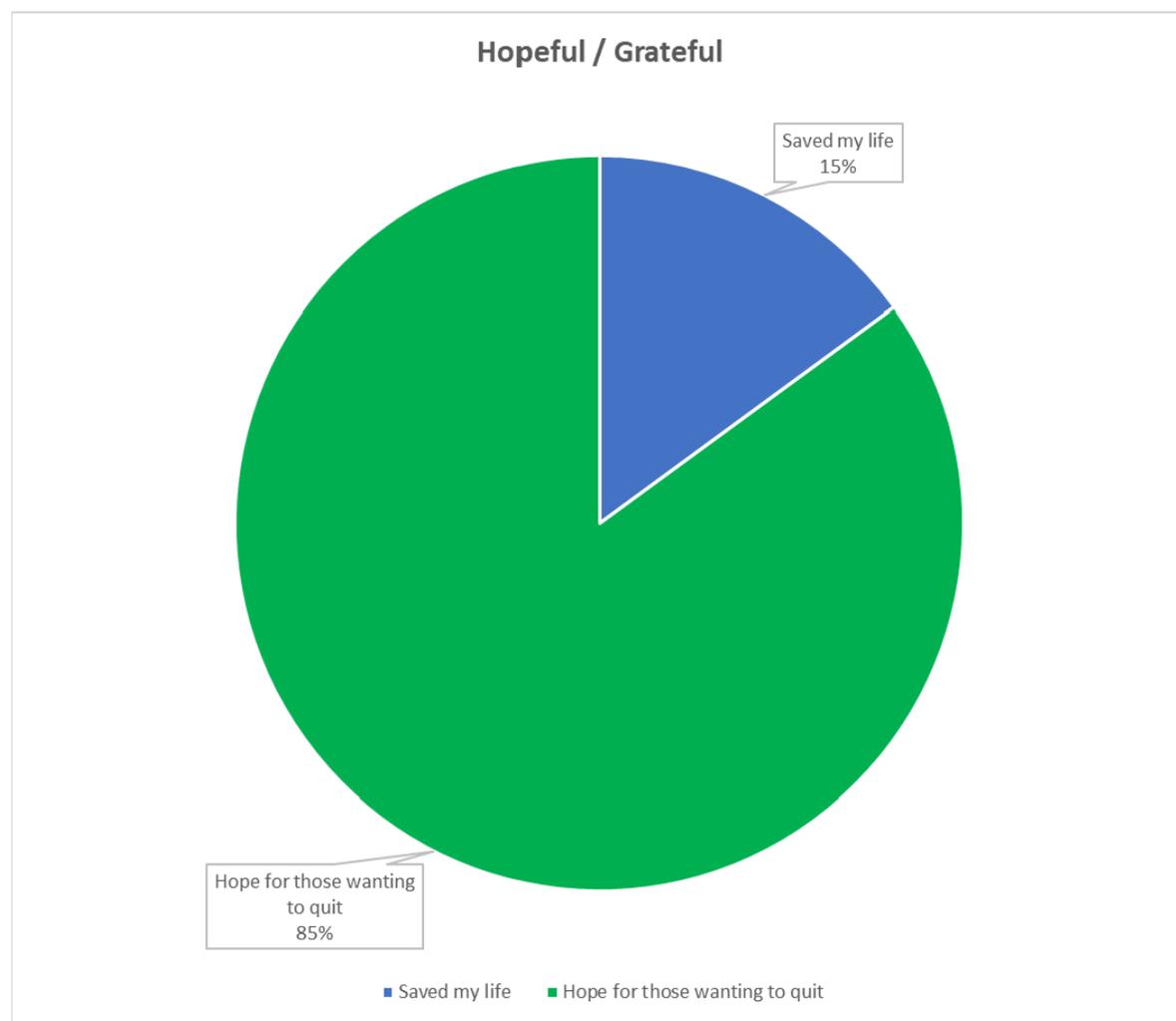
Interesting Details

- Drilling down into the Quit Experience were two topics
 - Quitting was unexpected
 - Quitting was easy
- Of those who spoke on their quit experience, 7% were unexpected quit results!
- Is there any other quit method that produces “accidental quitters”?



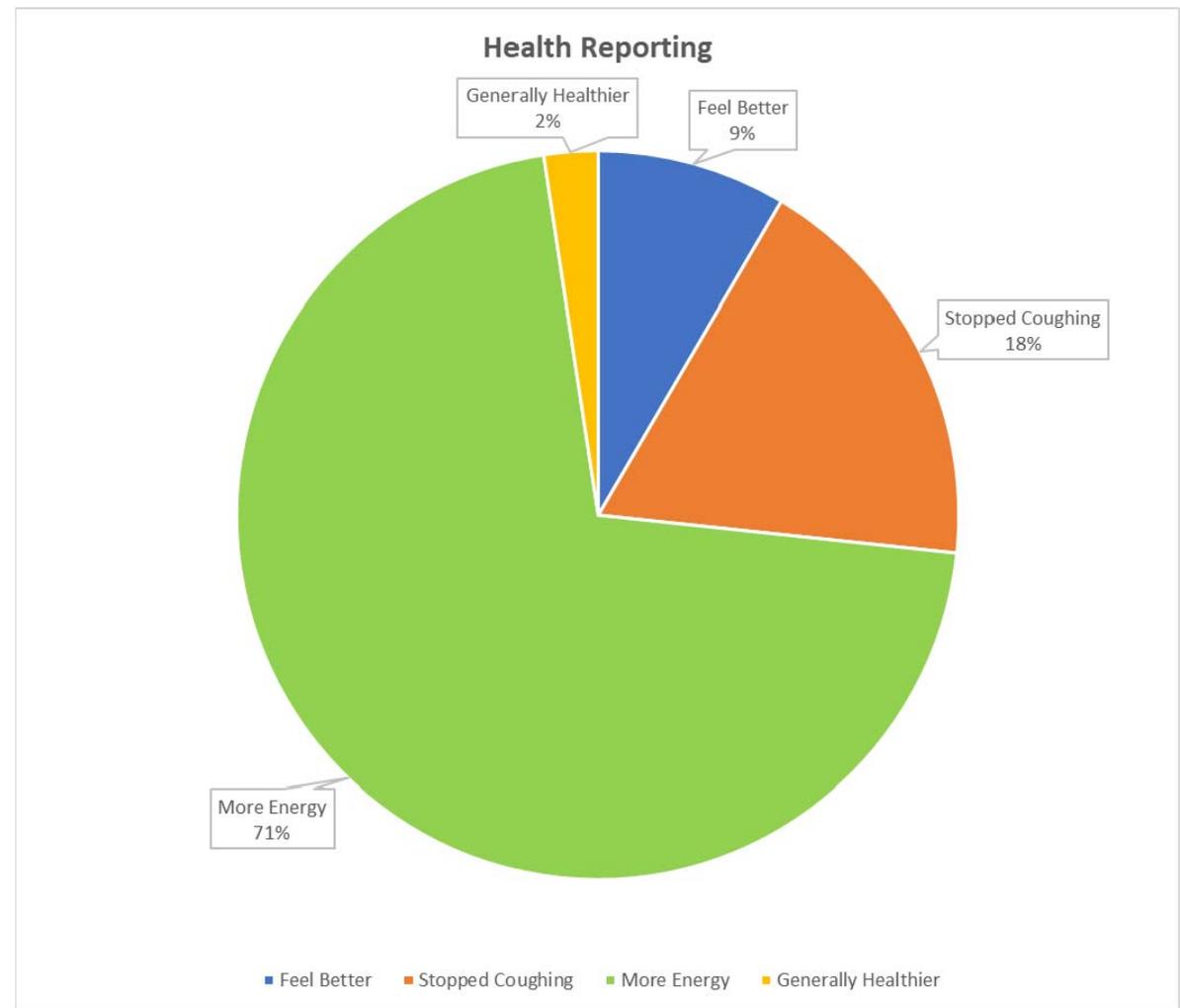
Interesting Details

- One topic explored with our proof-of-concept queries was to get at the underlying causes of positive sentiment
- While vapers have often said “Vaping saved my life” the real positive sentiment was overwhelmingly hope for others wanting to quit



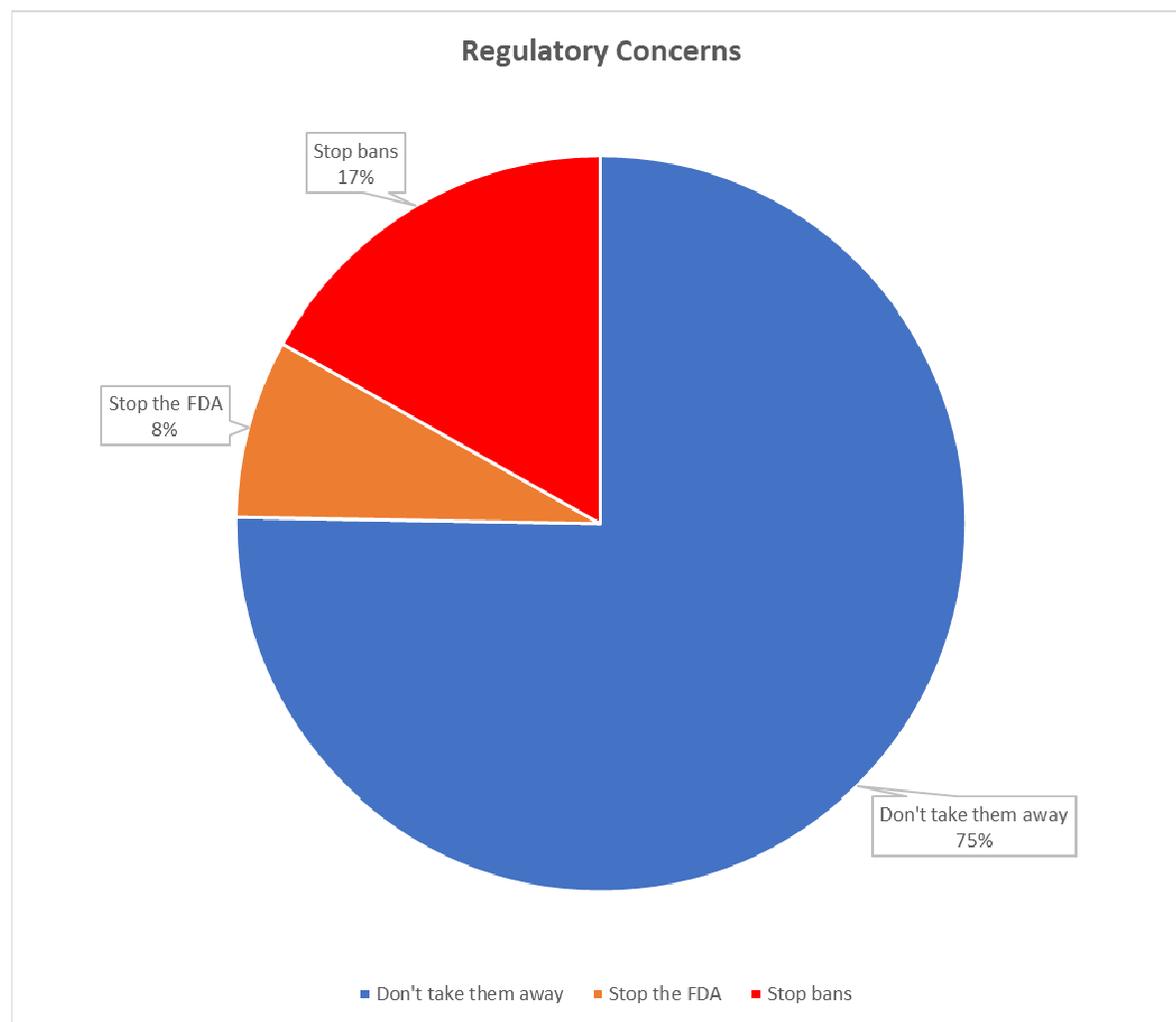
Interesting Details

- Health perception is an important reinforcement to any quit attempt
- The most commonly expressed health topic was “More Energy”
- Stopped Coughing came in 2nd at 18%
- Compare this with other quit methods



Interesting Details

- Drilling into the Regulatory concerns to explore negative sentiments
- Over 75% of those expressing concern was generally “Don’t take them away”
- “Stop the bans” was 2nd at 17%



Future Work

- The current work is a primitive “proof-of-concept” design
- Further work to refine the correlation using better models is possible
- Topic selection was ad-hoc and simplistic
- Better results could be possible with more directed queries
- Dr. Brian Carter’s upcoming work will lead to these more refined queries and allow model accuracy to be even further refined
- Collaboration of the two study methodologies can create a rapid and accurate analysis of a large body of texts – including the entire CASAA testimonials database, Twitter, Facebook, and other collections

Final Remarks

- Listening to vapers is an important, and often dismissed as anecdotal, resource to guide efforts to reduce the harm from smoking
- Turning these anecdotes into data is possible with coordinated research across multiple problem domains.
- The current effort represents a non-invasive attempt to listen to what smokers who switched to vaping say about themselves