Smoke Signals

How audio analytics can help life insurers detect undisclosed tobacco use

Contributors
Shane De Zilwa
Erik Edwards
Nicholas Irwin
Meyeneobong Inyang, MD

Society has come a long way from smoke-filled airplanes, offices, and restaurants.

Yet even as smoking has been pushed toward the margins of public life, it remains a potent public health scourge. Today, 13.7 percent of U.S. adults are smokers, according to the Centers for Disease Control and Prevention.[1] While this number is down from nearly 21 percent in 2005, smoking is still the leading cause of preventable death in the United States.[2] Indeed, while other forms of tobacco consumption, such as smokeless tobacco or vaping, have grown in recent years, smoked tobacco still constitutes the overwhelming majority of consumption—a full 1.1 billion users around the world.[3]

Life insurers naturally have a keen interest in this vice because diseases related to tobacco use are responsible for killing 7 million people globally per year.[4] Self-declared smokers are estimated to comprise about 4 percent of a U.S.-based life insurer's term life book of business, according to Verisk research. In addition, many smokers choose not to self-declare. Smokers have a clear financial incentive to withhold their smoking status on a life insurance application: Life insurance rates can be up to 200 percent higher for smokers compared to nonsmokers.[5] In fact, studies have shown that as many as 47 percent of smoker life applicants don't disclose their smoking status, and instances of nondisclosure increase as the cost of the policy rises.[6]

These hidden smokers can drive up premiums for nonsmokers because insurers are invariably forced to spread the cost of insuring these nondisclosed risks across their entire book. They also constitute a significant source of premium leakage. Verisk research suggests this nondisclosure costs the life insurance industry $3.4 billion in lost premiums annually.

Not surprisingly, "catching lying smokers" was one of the top five challenges cited by life insurers designing automated underwriting systems.[7]

Life insurance companies don't need to pump the brakes on accelerated underwriting just because some applicants are not self-declaring their smoking habits. Instead, they can deploy sophisticated screening tools that are more frictionless for their customers and more effective at rooting out undisclosed risk.

Enter audio analytics.

$10 billion

in lost premiums over the next 3 years from tobacco nondisclosure.

13.7%

of U.S. adults are smokers according to the Centers for Disease Control and Prevention.[1]

1.1 billion

people smoke tobacco around the world.

Up to 47%

of smoker life applicants don't disclose their smoking status.

7 million

deaths globally each year due to tobacco-related diseases.[3]

Establishing a link between smoking and vocal changes

Inhaling combustible tobacco irritates the mucosal surfaces of the mouth, upper airway, and lungs.[11] Additionally, tobacco smoke can cause damage to the vocal cords.[12] This damage can include:

- erythema, or swelling of the vocal cord region
- asymmetry in the vocal cords
- abnormal vocal cord movements
- the growth of lesions, including polyps, cysts, and nodules
- polypoid degeneration, also known as Reinke's edema, which is a buildup of edema fluid in the mucous membranes of the vocal cords[13]

Collectively, these impairments have been linked to changes in or degradations of a person's voice—a condition known as dysphonia. Dysphonia helps us draw a clear distinction between the voice of a smoker and a nonsmoker.[14]

The anatomy of voice

Dysphonia's markers aren't simply physiological and perceptual. Dysphonia makes a discernable imprint on the audio wave produced by vocal cords.[15] Vocal cords produce audio waves that have certain measurable characteristics. There are several elements of the audio wave produced by human speech that serve as markers for smoking-induced dysphonia.
These include:

- fundamental frequency, or the lowest frequency produced in an acoustic wave
- range, or the variety of pitches a person can produce
- jitter, a measure of instability in the audio wave's frequency
- shimmer, a measure of the instability of the amplitude (or distance between the high and low points) of the vocal wave
- tremor
- maximum phonation time[16]

In addition to these individual metrics, there are well-established collective indices that are used frequently by the medical and scientific community in the study of voice disorders, such as the Dysphonia Severity Index and the Voice Handicap Index, that aggregate several different variables linked to voice quality and physiological impairments.

Urine analysis
Insurers don't always accept an applicant's self-declared nonsmoker status without additional verification. However, some may choose to only pursue a subset of this self-declared population for closer scrutiny—such as those with higher sum policies (as low as $100,000), and/or those with a significant medical history. In such cases, insurers typically employ cotinine urine testing to verify smoking status.[8]

Cotinine is a metabolite of nicotine that can appear in a urine analysis for roughly a week, or less, after smoking.[9]

Aside from its time-limited efficacy, a cotinine test can also return positive results for nonsmokers who chew nicotine gum or wear a nicotine patch to aid quitting.[10] For life insurers, cotinine tests can add time, expense, and customer irritation during the application process. Sidestepping, or at least minimizing, the frequency of urine tests that a life insurer must order for applicants is a critical objective for those insurers seeking to modernize the underwriting process.

Studies describing the measurable differences in the voices of smokers versus nonsmokers serve as the medical basis for our tobacco voice analytics product. (While there is evidence to suggest that routine marijuana smoking causes vocal cord damage similar to combustible tobacco, the literature on this topic is not as well established. Therefore, we have currently restricted the Verisk voice analytics model to the effects of combustible tobacco.)

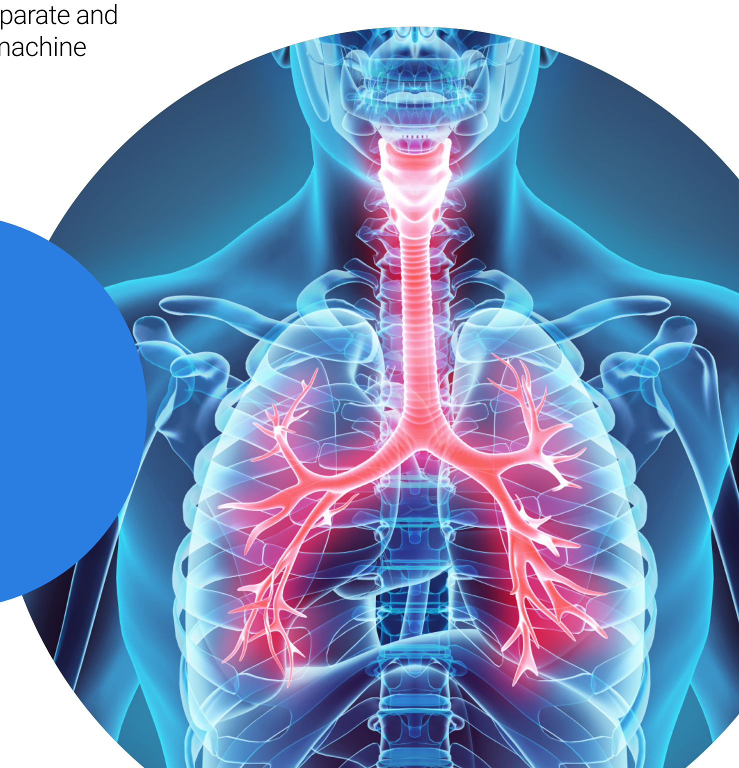
For instance, one study found that both male and female smokers had a lower fundamental frequency than nonsmokers.[17] Another study, conducted on relatively early-stage smokers, also found that several other acoustic wave parameters, such as tremor and jitter, were significantly altered in a consistent fashion compared to nonsmokers.[18] This study also revealed that smokers had a 1.8x higher risk for self-reported voice problems than nonsmokers.

A 2017 study of 220 young adult females (110 smokers and 110 nonsmokers) found a statistically significant difference for fundamental frequency, the sustained voicing of the /a/ sound, and the percentage of jitter between smokers and nonsmokers.[19] Another study, which included water pipe or hookah smoking of tobacco in addition to conventional cigarettes, found a lower mean fundamental frequency in cigarette smokers than nonsmokers.[20] An additional study showed that smoking caused a decrease in fundamental frequency, increase in jitter, and an increase in shimmer for the vowels /i/, /u/, /a/ (as in "got"), and /æ/ (as in "cat"). All these effects were stronger in smokers who had smoked longer.[21] Collectively, these studies confirmed empirically what may already sound intuitively correct: Smoking damages your throat and vocal cords, and that damage, in turn, changes your voice.

Finding the (smoke) signal in the noise

Armed with the knowledge that smoking makes a discernable impact on the voice, Verisk set out to develop machine learning methods for detecting the signatures of smoking in voice signals captured during a recorded phone interview with a life insurance applicant and delivering a probability of smoking status during the underwriting process. While there's a rich body of medical literature supporting the notion that smoking causes dysphonia, there has been scant work to date on using statistical and machine learning models to analyze digital voice signals for dysphonic symptoms linked to smoking.[22] This is due, in part, to the tremendous difficulty involved in creating such systems:

They require clearly labeled audio files to train on, a robust method to separate and process the smoker's voice from the interviewer's, and statistical and machine learning model development to isolate elements of the audio file that may reveal the presence of dysphonia. After extensive research, Verisk developed a first-of-its-kind, proprietary, and patent-pending system that employs three distinct subsystems that work in parallel to analyze voice data and deliver predictions on smoking habits.



The automatic analysis of voice can assign smoker probability to a high degree of accuracy.

Acquiring the training data

To properly develop our machine learning model, we needed to train it on labeled data—files whose contents are clearly identified so that the model can recognize subsequent appearances of the same or similar data. In this case, that means we needed data that clearly labeled the voices of smokers versus nonsmokers. One of the challenges in using machine learning for smoker identification is the lack of properly labeled data from which to train our models.

Although there is one open data set with smoker labels and voice recordings, it only has audio clips captured during open phonation (the classic "aaahh" you intone while the doctor peers into the back of your throat).

The Verisk team licensed a database of telephone interviews from an international non-profit supporting language education research with about 600 speakers, of whom approximately 20 percent were smokers, further supplemented by new data collection efforts and partnerships. This allowed us to train and test a variety of models on the smoker identification question. The performance of our current voice analytics model is consistent with the hypothesis that the automatic analysis of voice can assign smoker probability to a high degree of accuracy.

The detection system in action

Figure 1 illustrates generally how the detection system works.

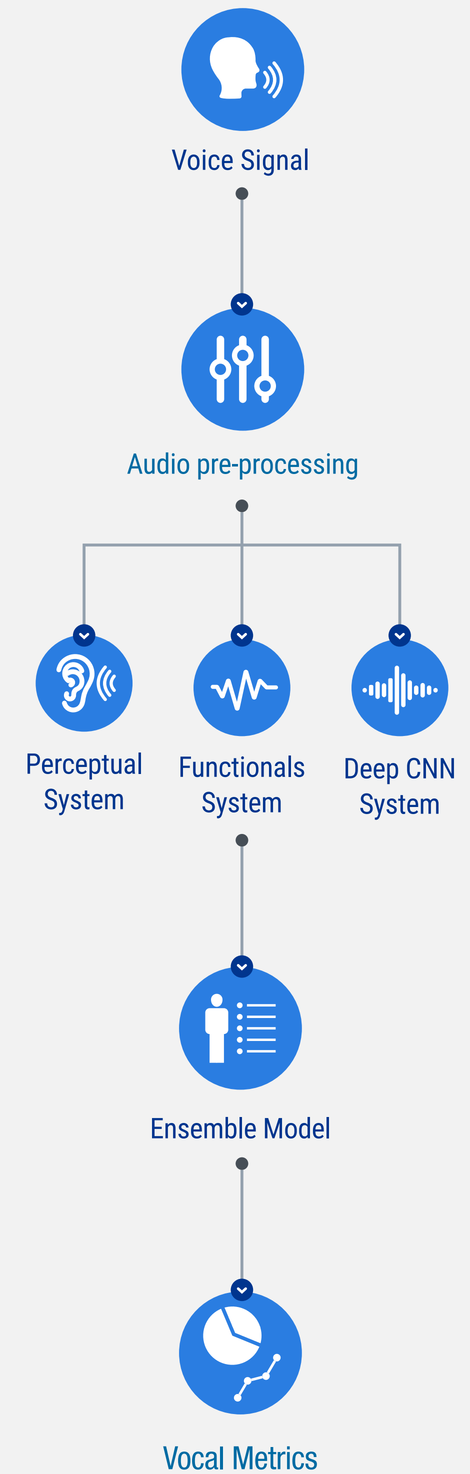
Stage one: Acoustic feature extraction
In this stage, we ingest the audio of the telephone interview with the life applicant and perform several operations, including digital signal processing and audio segmentation. We also perform speaker diarization, or the process of splitting the voice signal into two components: the voice of the interviewer and the voice of the applicant.

The final process in this stage is speaker identification, which enables us to pinpoint which of the two voices is the applicant's. Since our audio model parses speech during an interview process, it's critical to isolate the intervals where the life insurance applicant is speaking from the interviewer's dialogue so that we can be confident we're analyzing the correct voice.

Stage two: Subsystem processing
After the audio processing stage is complete, voice data is passed in parallel to three subsystems that employ three distinct audio modeling techniques to derive provisional probabilities about the likelihood that the voice being analyzed is indeed a smoker's voice. By using three methodologically diverse systems that are nonetheless individually accurate, we can deliver a more robust final prediction than if we simply applied a single system to the problem.

Subsystem one: Perceptual system
The first subsystem, dubbed the perceptual system, uses methodology motivated by well-established principles of human auditory perception combined with classical statistical methods. The system works by extracting voice features from individual snippets, or frames, of the audio signal to create a signal vector representative of the voice signal. The features are then averaged over all of the frames in the audio file, and the result is compared to an average class of nonsmoker audio signals. At a high level, this subsystem can quantify differences in the signal according to metrics that align with how humans would perceive differences in sound.

Figure 1: The detection system



Subsystem two: Functionals system

The functionals system uses a methodology that has emerged in the last decade in the voice analytics literature, particularly in emotion recognition. Where the first subsystem tackled extremely brief snippets of audio, this subsystem is designed to analyze longer time scales and target vocal characteristics, like pitch and prosody, that could be missed when analyzing shorter periods.

The functionals system starts with the raw audio data and extracts numerous features from it (features in this context refer to short-term characteristics in the audio data, such as noise ratios, frequency patterns, etc.). From these initial features, the system derives a huge number of functionals (more than 6,000), which are various mathematical transformations and summarizing statistics over longer time scales. Then, machine learning methods are used to select the most useful features, which helps us purge redundant data while creating a median voice point for the speaker. These are analyzed using a pattern recognition algorithm to identify functionals that could be most indicative of smoking-induced dysphonia.

Subsystem three: The deep CNN system

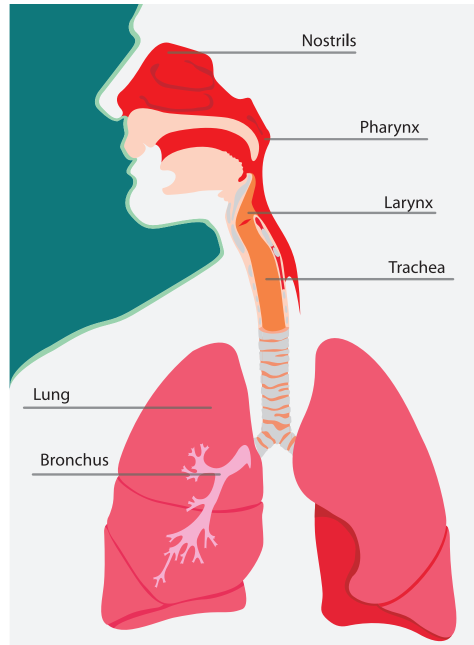
The third parallel system is the deep convolutional neural network (CNN) system. This system digests the raw audio signal data and transforms it into a pictorial representation called a spectrogram. The spectrogram, or image, represents voice frequencies over a given time period.

When the spectrogram is built, state-of-the-art machine vision algorithms compare it against similar images representing the frequency patterns of a smoker's voice and the frequency patterns of a nonsmoker's voice. Depending on how it matches up, the voice file (i.e., the life applicant) is assigned a probability as to the likelihood that the individual is a smoker.

Stage three: The ensemble

All three subsystems return a probability score indicating the likelihood that the voice analyzed is that of a smoker. These subsystems also return a "confidence score" that reflects the individual model's confidence that its probability score is accurate.

In the final stage, all of the probabilities and confidence scores from the subsystems are fed into an ensemble model that algorithmically blends them into a final prediction.



All three subsystems return a probability score indicating the likelihood that the voice analyzed is that of a smoker.

Further research

Working off its initial training data, our model can correctly identify smokers 85 percent of the time.*

We plan to expand and enrich this training data with real-world tele-med interviews. As training on these new data sets continues, we expect the performance of the model to improve, reflecting the input of additional data.

There are also several areas of ongoing research as we continue to refine the model. Among them are efforts to ensure that ambient and environmental noises recorded during the interview process don't interfere with signal processing and analysis.

There are some outstanding questions, including:

1. **Is the model accuracy impacted if someone has a cold, throat cancer, or other ailment?** Generally, different pathologies result in differing changes to voice patterns and our model is unlikely to be impacted by the presence of other conditions, but we do not have enough data to make a definitive statement on this issue at this time.

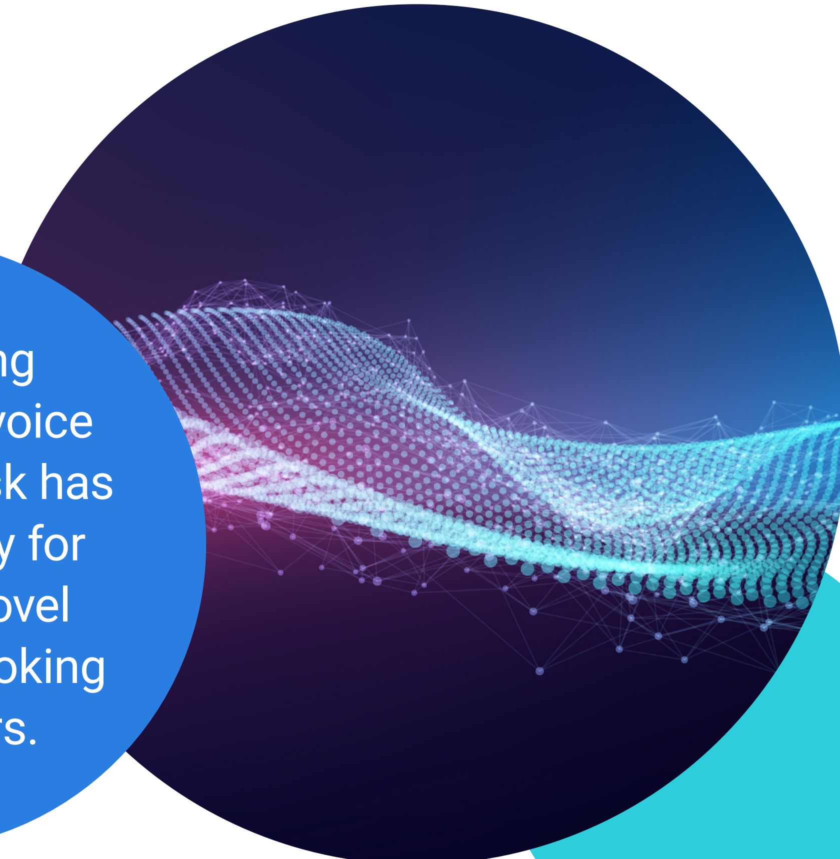
2. **Is the model less accurate for people who have been smoking for only a few years?** In our data we have both recent smokers and long-time smokers and we have not noticed challenges in flagging recent smokers, but there is not enough data to make a definitive conclusion at this time.

3. **How long does the smoking damage last for someone who quits smoking?** Will they be flagged as a smoker after quitting? The length of time to heal vocal cord damage due to smoking combustible tobacco depends on how long the individual has been smoking and how much prior to quitting. It's possible that long-time smokers who have recently quit will still be flagged by our model.

Unlocking the power of voice
Smoking is just one of a broad range of personal and clinical variables that may be detectable using voice analytics. From the existing medical literature, we find that age, gender, body size, hydration, dementia, depression, reflux, Parkinson's, and other conditions and physiological characteristics have acoustic markers that may be amenable to the same analytical modeling we've applied to smoking.[23] We're currently working to generalize our system so that it may be applied to additional physiological variables of interest to life insurance underwriters.

Working off its initial training data, our model can correctly identify smokers 85 percent of the time.*

* Based on Verisk research, only 4 percent of a typical U.S. life insurer's term-life book consist of self-declared smokers. Therefore, when examining the model's output, we focused on the top 4 percent of results returned with a high probability that the voice was that of a smoker. Of this 4 percent, the model accurately identified 85 percent of the smokers (calculated using cross validation). An accuracy rate calculated on self-declared smokers on a life insurance application will be roughly



By unlocking the power of voice analytics, Verisk has paved the way for an entirely novel method of smoking out smokers.

A novel path to simpler workflows

We've found this audio analytics model effective at deriving a percentage likelihood that an applicant is a tobacco user. Life underwriters can use this information to help decide whether to subject an applicant to more traditional underwriting measures, such as lab tests, or to pass them through an accelerated application process. Audio analytics can help significantly reduce the number of cotinine tests an insurer needs to order by testing only those individuals who are likely not disclosing their actual smoker status.

Life insurers may not even need to conduct a lengthy tele-med interview to acquire voice data for analysis. Online applications could feature a "record" function that prompts applicants to record a short audio snippet—perhaps verifying that the information they're submitting in the rest of their application is truthful. This short snippet could then be subjected to analysis.

By unlocking the power of voice analytics, Verisk has paved the way for an entirely novel method of smoking out smokers.

Notes

1. “Current Cigarette Smoking Among Adults in the United States,” Centers for Disease Control and Prevention.
2. “Current Cigarette Smoking Among Adults in the United States.”
3. Kavitha Hariharan, et al., Insurer Perspectives on Smoking Risks, Marsh & McLennan Advantage, 2019.
4. “Smoking and Tobacco Use: Fast Facts,” Centers for Disease Control and Prevention.
5. Ibid.
6. Brian Lanzrath, et al., “Applicant Medical and Smoking History Nondisclosure in the Life Insurance Marketplace,” Contingencies, November/December 2016.
7. Verisk survey of 27 carriers.
8. Kavitha Hariharan et al.
9. Nicotine Testing – Common Questions, Alere Wellbeing, 2012, p. 4, accessed on March 16, 2020.
10. Ibid.
11. Thomas Fitzpatrick, et al., “Upper Airway Complications of Smoking,” Clinics in Chest Medicine, March, 2000.
12. Vocal Cord Disorders, Johns Hopkins Medicine, accessed on January 29, 2020.
13. “Polypoid Degeneration,” Laryngopedia, accessed on January 29, 2020.
14. Dogan Pinar, et al., “Investigating the Effects of Smoking on Young Adult Male Voice by Using Multidimensional Methods,” Journal of Voice, November, 2016.
15. Haewon Byeon, “The association between lifetime cigarette smoking and dysphonia in the Korean general population: findings from a national survey,” PeerJ, April 28, 2015.
16. Linda Carroll, The Voice Lab: Is It Just Numbers?, American Speech-Language-Hearing Association, 2011.
17. Marie Reine Ayoub, et al. “The Effect of Smoking on the Fundamental Frequency of the Speaking Voice,” Journal of Voice, September, 2019.
18. Byeon.
19. Dionysios Tafiadis, “Effects of Early Smoking Habits on Young Adult Female Voices in Greece,” Journal of Voice, November 2017.
20. Marie Reine Ayoub, et al.
21. Irena Vincent, et al., “The Effects of Cigarette Smoking on the Female Voice,” Logopedics Phoniatrics Vocology, April 2012.
22. Verisk review of published research material.
23. Erik Edwards, et al., “Analysis of Voice Characteristics for Detection of Smoking and Other Risk Factors,” Verisk, December 2019.

+1.800.888.4476 / info@verisk.com / verisk.com/life