WHITE PAPER

The coming revolution of voice-based digital biomarkers

It’s time to understand the opportunities, challenges and potential of these technologies to diagnose and monitor health conditions.

By Todd Greenwood and Nathalia Nunes

Impact where it matters.®

The COVID-19 crisis has accelerated the timeline for disruptive technologies, which many incumbents had until then approached in a guarded wait-and-see mode rather than through active listening, trendspotting and preparing. The pandemic has caused trends to evolve even faster than they otherwise would have, and the new normality has underscored the importance of keeping an ear out for signals that are bound to usher in smarter, more efficient and more user-centric ways of doing research and business.

ZS, which perpetually scans the horizon for new patterns and trends, views voice-based digital biomarkers as the bellwether of a coming revolution in biomedical research and healthcare. Voice quality is both a telltale sign of patient health and a novel frontier for health assessment, and technologies that enable its capture on everyday consumer devices and then perform symptom detection and prediction have tremendous potential for both research and development (R&D) and the clinical environment. Yet because these biomarkers are so novel, questions and challenges regarding how to bring them to market and put them to efficient use abound.

This white paper is intended as a guide to biopharmaceutical companies, clinical researchers, physician entrepreneurs and digital health innovators engaged in or with interest in developing, operationalizing and integrating voice-based digital biomarkers. It is meant to help these stakeholders and understand the opportunities this new frontier presents, assess their needs and capabilities if they plan to embark on this journey and get the strategic advice they need as they traverse this barely charted terrain.

Voice-based digital biomarkers: The coming revolution

What if we could detect disease through the careful analysis of speech and voice? Speech has long been used by neurologists and speech pathologists to detect changes in certain health conditions. With the advent of connected technologies, it's the new frontier for digital biomarkers.

Digital technology is opening a world of possibilities for the utilization of speech in the screening, diagnosis, prevention and monitoring of a broad range of conditions including:

- Neuromotor and neurodegenerative disorders such as Parkinson's disease and Alzheimer's disease
- Respiratory illnesses such as COVID-19 and chronic obstructive pulmonary disease (COPD)
- Mental health disorders such as depression and schizophrenia
- Cardiovascular diseases such as congestive heart failure and coronary artery disease

The human voice produces multiple frequencies, of which the human ear is capable of perceiving only a narrow sequence. But digital technologies can register frequencies across the entire spectrum and, with the aid of special algorithms, monitor for early sound-based signals of disease. Given the widespread availability of connected personal devices on which such frequencies can be captured, voice-based symptom detection can therefore not only benefit patients through timely detection of potential symptoms but also expand access to screening for underserved populations typically excluded from clinical research.

Voice-based biomarkers also hold the promise of enabling continuous, non-invasive disease monitoring at a significantly lower cost than conventional episodic follow-ups (which often occur months apart) and clinic appointments. Speech tools can track disease progression frequently and over extended periods of time. This consistent monitoring can be leveraged to modify treatment plans and medications as needed while cutting down on patient visits to a doctor's office or a hospital. These technologies are also compatible with hybrid and fully remote decentralized clinical trials, where their potential to capture surrogate endpoints—both established and newly defined—can add valuable insight to assessing the efficacy and effectiveness of investigational therapies.

Based on voice-based biomarkers' anticipated utility, their rapid growth within digital medicine and the intensifying investment and innovation in the space, Coherent Market Insights expects these technologies' global market share to exceed $4 billion by 2027.[1] Nevertheless, implementing voice-based biomarkers is far from straightforward in practice; as of today, there are none approved for clinical use in the United States.

All references appear in the endnotes on page 19.

Digital biomarkers

Let's take a step back to define biomarkers and digital biomarkers in particular. A brief explanation can help those new to the concept fully grasp the implications of developing and embedding voice-based digital biomarkers into clinical research and practice.

Traditional biomarkers are measures of blood, tissues and human functions. The readings are typically classified as normal or abnormal, and they are meaningful to the extent that we understand their relationship to health-related outcomes associated with a condition or disease. Traditional biomarkers are measured in a lab or with common medical devices, such as a thermometer or blood pressure cuff monitor.

Digital biomarkers mimic their conventional counterparts but employ a different collection method. Readings are captured via sensors, wearable trackers or other computational tools, then analyzed using software and hardware applications. Just like traditional biomarkers, digital biomarkers aim to derive standardized, objective measures relevant to individual health or clinical outcomes, including diagnosis, disease progression or response to treatment; self-reported outcomes, genetic information and data collected via durable medical devices, such as pacemakers and defibrillators, are therefore not included among digital biomarkers. Similarly, digital biomarker data are meaningful only in the context of assessing a patient's susceptibility to or risk of developing a disease or in the course of ongoing disease monitoring.[2] Absent this use, the data are not intrinsically meaningful.

Thanks to the broad adoption of digital devices such as smartphones, smartwatches and fitness trackers, the collection of digital biomarkers is becoming increasingly common. According to the digital health venture fund Rock Health, digital biomarkers are decreasing the costs of collecting population-level health data over long periods of time, while also making possible the collection of longitudinal data for individual consumers. In addition, since they don't rely on human-mediated measurement, digital biomarkers are believed to be highly accurate and objective. They are therefore tremendously valuable in "subjective" medical domains such as psychiatry and neurology. And because their interpretations are based on algorithms trained to detect patterns across both time and values being measured, these biomarkers reduce biases in physician judgment that may result from a selective or incomplete reading of such values.[3]

Digital biomarkers can even play a crucial role in digital phenotyping, the multidisciplinary field that was first described in a seminal 2016 paper as "the moment-by-moment quantification of the individual-level phenotype in situ using data from personal digital devices."[4]

All references appear in the endnotes on page 19.

The advent of voice-based digital biomarkers

Voice-based biomarkers, a relative newcomer to the field, screen for abnormalities and early symptoms of disease through detection of acoustic or linguistic changes in speech. This capability may come as a surprise to some, since laypeople don't typically think of voice as an indicator of heart disease or respiratory illness. Yet the production of human speech and the use of voice are complex phenomena that require close coordination between brain and lung functions, intertwining numerous neurological and physiological processes. Changes in patterns of speech or in the ability to produce sound may therefore indicate a health impairment or dysfunction. Through targeted research studies, linguists and speech pathologists in recent years have been able to classify various acoustic, prosodic, emotional and lexical voice-based features associated with conditions such as Alzheimer's disease, depression, schizophrenia, autism spectrum disorder, Parkinson's disease and asthma.[5]

"Voice-based digital biomarkers are very well placed in terms of potential compared to other digital biomarkers because they can cover a wide spectrum of clinical applications," says Guy Fagherazzi, Ph.D., who leads the Deep Digital Phenotyping Research Unit and the Department of Population Health at the Luxembourg Institute of Health. "Whereas traditional digital biomarkers using connected trackers or basic medical devices usually target one specific condition—for example, continuous glucose monitoring devices track blood glucose levels in people with diabetes—with vocal biomarkers you can track multiple conditions, multiple symptoms, and cover pretty much all aspects of the health state of an individual. And you also have a wide range of potential applications—for clinical research, for remote monitoring, for telemedicine purposes."

Developing voice-based biomarkers to identify particular diseases or health conditions therefore requires determining which areas of speech should be assessed and how to capture those assessments. For example, assessing a motor speech impairment may be best accomplished by examining acoustic features in the context of a structured speech task, such as sustained phonation. Patients may be asked to hold an "ahhh" sound for several seconds. In contrast, assessing irregularities in the organization and content of speech that are often associated with cognitive impairment may require analyzing linguistic features derived from a less structured speech task. Patients may be asked to describe a picture or simply to extemporize.

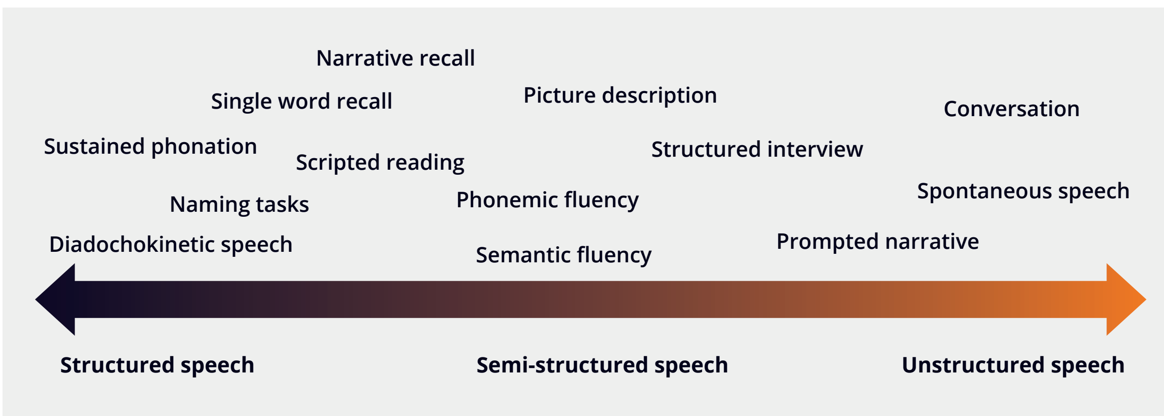
All references appear in the endnotes on page 19.

As Figure 1 demonstrates, speech tasks and assessments can range from highly structured to unstructured, depending on the elements of speech being analyzed:

- Basic vocal tone and quality can be determined through exercises such as sustained phonation (saying "ahhh"), a structured speech task.
- Fluency of phonetics, speed of recall or fluency of concept generation are typically measured through semi-structured tasks, such as repeating set phrases.
- Multifactorial assessment is commonly gauged through unstructured spontaneous speech or free-form conversation.

FIGURE 1:

Example of speech tasks ranging from structured speech to unstructured conversation



Note. From "Evaluation of Speech-Based Digital Biomarkers: Review and Recommendations," by J. Robin, J. Harrison, L. Kaufman, F. Rudzicz, W. Simpson and M. Yancheva, 2020, Digital Biomark,4, pp. 99-108.

Source: https://doi.org/10.1159/000510820

To equip these assessments with maximum precision as a basis for developing fit-for-purpose voice processing techniques and analysis systems, scientists have quantified typical variations in the voice's acoustic features. To date, some of the most studied are:

- Frequency: jitter (local, absolute, ppq5, ddp)
- Amplitude: shimmer (local, db, dda, apq11, apq5, apq3)
- Harmonic: noise-to-harmonic, harmonic-to-noise

Aural Analytics—which develops neuro-biomarkers for detecting cognitive impairment after a stroke, bulbar deterioration in amyotrophic lateral sclerosis (ALS) and respiratory conditions such as COPD—serves as a powerful example of how versatile voice technology can be in the clinical setting. Julie Liss, chief clinical officer at Aural Analytics, reports that the company's stroke-related cognitive detection algorithm focuses on ascertaining changes in a patient's vocabulary rather than the tone or frequency of their voice. However, when monitoring for changes in respiratory health, important features for the algorithm include changes in speaking and pause rate and changes in speaking volume. In a study published last year in the peer-reviewed journal Nature, Aural Analytics also demonstrated a proof of concept for remotely detecting early speech changes that portend a likely deterioration in bulbar function, which affects ALS patients' ability to communicate verbally.[6]

Voice-based technology is like a thermometer—the thermometer doesn't diagnose...but it does give you a cue about how serious a temperature deviation from normal is and in this way educates and informs you about what the appropriate next step would be. That's the right way to think about the kind of information voice-based biomarkers provide.

– Jim Harper, Founder and COO, Sonde Health

Sonde Health is another company developing speech-based technologies, with a focus on helping clinicians to better understand mental health, respiratory illness, sleepiness and hypertension. Its founder and chief operating officer, Jim Harper, Ph.D., a biochemist by training, says voice-based technology is essentially "like a thermometer—the thermometer doesn't diagnose, it doesn't tell you whether you have COVID-19 or the flu or a post-operative infection. But it does give you a cue about how serious a temperature deviation from normal is and in this way educates and informs you about what the appropriate next step would be. That's the right way to think about the kind of information voice-based biomarkers provide."

All references appear in the endnotes on page 19.

Sonde's latest product is an algorithm that helps to detect symptoms of depression and anxiety through changes in voice acoustics. Unlike most other companies innovating in the speech space, Sonde aims to empower consumers directly with this tool. "This product will be a bit like [a] mental Fitbit—it will enable people to look at six-to-ten features derived objectively from their voice that are known to change when one is experiencing depression and anxiety. The readings won't reveal a specific level and they are not meant to diagnose, but if there is a trend, it will tell the user in which direction that trend is increasing risk that their mental fitness or resilience might be decreasing," says Harper. "The goal there is to encourage healthy behaviors, positive self-care, reducing stress and resting more. If engaging in something that is a step removed from therapy is likely to activate more people compared to some employer-sponsored wellness programs, we would have accomplished that goal—not excluding the possibility to also partner with such assistance programs to augment participation even further."

Voice technology has been a long time coming

Although the U.S. Food and Drug Administration (FDA) has yet to approve a voice-based digital biomarker, the concept has a long history. Some of the earliest efforts, which concerned the development of cough-counting technology, appeared in the 1950s; automated cough frequency monitoring arrived in the late 1980s. In both cases, the equipment was bulky and the methodology was exceedingly restrictive to patients. In the 1990s and early 2000s, advances in technology such as MP3 audio, digital storage, miniaturization of microphones and battery power made possible the advent of ambulatory devices that could longitudinally capture and record high-quality data. Yet processing the recordings remained a challenge, because final cough counts relied on manual assessment. Full automation of cough counting was therefore not achieved.[7]

By and large, these early efforts faltered because the technologies available at the time were insufficient to accomplish the task. They pale in comparison, for instance, to the powerful and sophisticated voice technologies used today by Apple, Amazon and Google, which collect millions of hours of voice records through their digital voice devices, including Siri, Alexa and Google Assistant. But the twentieth century researchers were on to something. As Webb writes, they were futurists listening for signals, scanning the data and testing patterns they thought could evolve into more robust trends. And some trends did emerge.

All references appear in the endnotes on page 19.

Around 2007, work on voice-based signal processing began in earnest. A team of Columbia University and IBM researchers recruited a group of 34 adolescents who were deemed to be at high risk of developing psychosis. They then collected voice samples from the subjects over two and a half years and analyzed the voice clips using technology focused on semantic and syntactic speech features known to be predictive of early psychosis. Five of the teens developed psychosis over the course of the study, and researchers reported in 2015 that their machine learning algorithm could predict the onset of psychosis with 100% accuracy using only voice samples.[8]

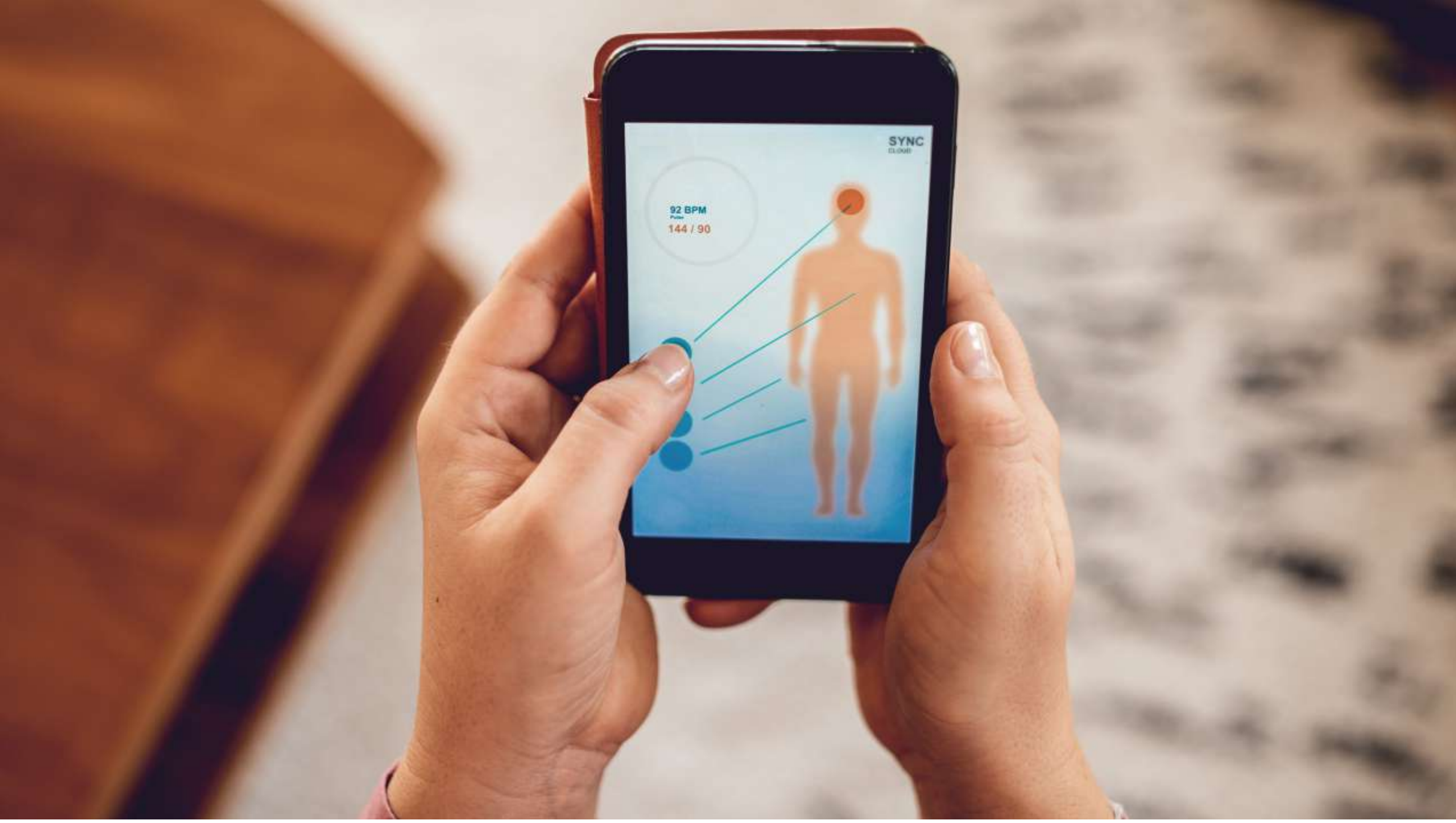
With advances in mobile devices, voice recognition, artificial intelligence (AI) and natural language processing, the pace of R&D in voice-based biomarkers is increasing rapidly. Faster data processing speeds are enabling the analysis of ever-larger volumes of voice data, and the ability to construct predictive and diagnostic meaning from changes in unstructured and natural speech patterns is taking shape in this very moment.

Physicians want to move away from pen-and-paper tests... Once we figure out a way that´s practical for them to do that, I think they will embrace this new frontier.

– Liam Kaufman, Co-founder and CEO, Winterlight Labs

"Psychiatrists and neurologists want more objective tools to assess their patients who are still subjectively diagnosed through pen-and-paper tests that date to the 1970s and 1980s and they want to be able to better characterize their symptoms. Voice-based digital biomarker technology isn't making a diagnosis, but it's objectively quantifying existing symptoms. Physicians can use those objective markers to form their own diagnosis based on their clinical skill and ability, and that's an important distinction," says Liam Kaufman, co-founder and CEO of Winterlight Labs, a Toronto startup that operates a speech-based AI assessment tool to detect the onset of dementia and other psychiatric conditions. "Physicians want to move away from these pen-and-paper tests, if possible. Once we figure out a way that's practical for them to do that, I think they will embrace this new frontier."

All references appear in the endnotes on page 19.



Challenges

Although the future looks bright for voice-based digital biomarkers, many challenges remain, from technical, ethical, regulatory and commercial considerations to patient concerns. Below we provide a brief overview of the main speed bumps vocal biomarker developers need to navigate on the road to successful commercialization or adoption of their technology.

Technical aspects

Any voice-based technology starts with the collection of quality speech data (such as WAV files), which requires the absence or near-absence of noise artifacts. Unwanted acoustics can obstruct audio pre-processing and feature extraction.[9] Other technical challenges revolve around the audio features being extracted. For example, structured speech is far less complex to extract and analyze but may not be as predictive as semi-structured or unstructured speech, particularly where the aim is to analyze the speech's linguistic rather than acoustic content. Even when analyzing an acoustic speech feature, including changes in certain vocal characteristics, such as an increasing tendency to slur certain words, extracting those artifacts and translating them into algorithms may be difficult.

All references appear in the endnotes on page 19.

Dearth of data samples

An algorithm's ability to discriminate with high precision between vocal signals is based on the volume and quality of the samples the algorithm has been trained, tested and validated against. However, there are as of today very few high-quality databanks of speech samples linked with clinical data or disease status, due to the relative infancy of voice-based digital biomarkers as a field of study.[10,11]

Privacy concerns

As researchers look for ways to build more voice datasets, consumers or patients may be wary of providing such samples. Article 4.1 of the European Union's General Data Protection Regulation (GDPR) implicitly considers voice to be part of a person's physical or physiological identity, and therefore non-anonymous data; similarly, in the U.S., the Biometric Information Privacy Act of Illinois (BIPA) includes the concept of "voiceprint" (using voice to reveal an individual's identity) in its definition of "biometric identifiers."[12,13] Biomarkers based on sustained phonation and collected through structured speech tasks may form an exception to these legal privacy protections.

Comparison to baseline

Where a speech biomarker is being used as a screening tool, its underlying algorithm works by extracting and analyzing a feature from a vocal sample and comparing that sample against a large population dataset. Screening tools therefore do not require the algorithms to look at individual changes over time. When biomarkers are being used as monitoring tools, however, the goal is to detect potential changes in speech quality or content over time compared to an individual baseline. But because the field of voice-based biomarkers is relatively young, and since building longitudinal individual-level databases is even more demanding than collecting data for population datasets, there may in many cases be no baseline available to furnish the algorithm. And even where a baseline exists, algorithms may misinterpret individual situational characteristics affecting voice quality, such as poor sleep quality and mood swings, as manifestations of disease symptoms, resulting in false positives.

Confounding variables

Many variables can confound speech analysis and therefore may need to be anticipated into the algorithm. For example, a biomarker being developed for Parkinson's disease would require a strategy for dealing with the potentially confounding variables of the patient's age, dialect or common comorbidities, such as depression.

All references appear in the endnotes on page 19.

Clinical relevance

Good biomarkers are understandable and meaningful to both patients and clinicians and are relevant to a therapeutic benefit. Vocal biomarkers, however, often lack gold standards or established methods to correlate speech aspects and features with specific disease processes. Standards for in-clinic voice evaluation exist for some diseases, such as Parkinson's (where measurements include jitter, shimmer, F0 and harmonics-to-noise ratio), but these are the rare exception.

Due to the dearth of real-world standards, predictive AI models that serve as the foundation for voice-based digital biomarkers may not be validated against established norms. As a result, they may be difficult or impossible to interpret or replicate—a phenomenon known as "AI's black-box problem," whereby algorithms that provide predictions cannot be deconstructed to reveal the bases on which predictions are made or to help researchers understand the limits of their predictive abilities. Take for example an AI algorithm that uses the ninth mel frequency cepstral coefficient (MFCC) as an input variable. The MFCC is related to the neurobiology of speech and to language production and can therefore reflect changes related to disease processes. Such an algorithm may be able to indicate early stages of Parkinson's disease, signs of depression and lung disease. It would do so by relying on what is known as a Fourier analysis of the logarithmic amplitude spectrum of the signals; frequency bands represented in this format closely represent the human auditory system and are used because of their efficiency and prevalence. But a nonspecialist would likely have more trouble grasping how the MFCC works—let alone the predictive AI model—compared to something as concrete and meaningful as a count of coughs over a 24-hour period.

Promising clinical areas

Notwithstanding these challenges, research linking voice-based features to symptom detection or disease progression has shown promise for the development of digital biomarkers in a number of clinical areas. Some of this research is already being applied in the form of non-clinical screening tools, as is the case with screenings for COVID-19, while other initiatives are at the testing and evaluation phase. Other applications, including the use of voice in conjunction with telemedicine for the purpose of diagnosis and monitoring, may come after sufficient time for development, approval and acceptance.

The field of voice-based digital biomarkers is moving very fast, so I think within five years we'll begin to see the first uses in clinical practice. I am pretty sure they will come in the field of neurodegenerative disorders.

– Guy Fagherazzi, Group Lead,
Deep Digital Phenotyping Research Unit, Luxembourg Institute of Health

"This field is moving very fast, so I think within five years we'll begin to see the first uses in clinical practice. I am pretty sure they will come in the field of neurodegenerative disorders, such as Parkinson's disease and aging in general, because these are the first fields that started to experiment with vocal biomarkers, so they have a bit of an advantage," says Guy Fagherazzi. "But soon, within five to ten years, we will see a lot of other applications as well because we will have had sufficient time to validate the technology and [to] go through all the necessary steps."

We review below the three main therapeutic areas where research into vocal biomarkers is most intense or has been ongoing the longest.

Neuromotor and neurodegenerative diseases/central nervous system diseases

- **Parkinson's disease:** In 2015, researchers at Sage Bionetworks launched a large observational iPhone-based study with the aim of monitoring day-to-day variability of Parkinson's disease symptoms and their sensitivity to medication. Participants in the "mPower study" performed four distinct tasks related to memory, motion and voice. The voice task required them merely to record themselves saying "ahhh" for no longer than ten seconds. After processing the audio and applying machine-learning models, researchers at the University of Washington DigiPsych Lab were able to distinguish participants in the Parkinson's disease group from those in the control group 85% of the time, exceeding the average accuracy of clinical diagnosis for both non-specialist doctors and movement specialists.[14,15,16]

All references appear in the endnotes on page 19.

- **Alzheimer's disease, ALS:** Research is underway to use voice to facilitate early detection of frontotemporal degeneration and to distinguish if changes are indicative of Alzheimer's, ALS or Parkinson's. Recent studies have also evaluated linguistic analyses' capacity to predict mild cognitive impairment, a precursor to Alzheimer's disease.[17]
- **Autism:** Researchers are exploring the correlation between voice pitch and the presence of autism.
- **Schizophrenia:** Work on voice-based approaches to early schizophrenia diagnoses and the management of treatment response is in emerging stages, with recent advances including the Columbia University and IBM research team's psychosis prediction algorithm developed in their work with adolescents,[18] among others.[19]
- **Bipolar disease:** One of the first clinical studies exploring voice biomarkers was the PRIORI study to predict bipolar mood swings.

Respiratory and pulmonary conditions

- **COVID-19:** Vocalis Health and Sonde Health are developing voice-based COVID-19 screening tools that can be used by employers to determine whether an employee might be ready to return to work after a COVID-19 diagnosis.
- **Chronic obstructive pulmonary disease:** Clinical tests are underway for using voice-based approaches to detect exacerbations in COPD, allowing physicians to intervene before a patient visits the emergency room or is hospitalized.
- **Pulmonary hypertension (PH):** PH is a rare disease that is difficult to diagnose, requires expensive, invasive diagnostics and is often not detected until later stages. Early studies have shown that specific vocal features inaudible to the human ear are correlated with the condition and could be used as screening or diagnostic tools.[20,21]

Cardiometabolic conditions

- **Congestive heart failure:** Exploratory research has demonstrated a correlation between voice signals and adverse outcomes (including hospitalization and mortality) in CHF patients, suggesting a possible role for voice along with telemedicine in CHF patient care.[22]
- **Coronary artery disease:** Researchers have reported the initial development of algorithms successful in detecting the presence of coronary artery disease using voice.[23]

All references appear in the endnotes on page 19.

Strategic approaches

Clinical development and validation are just two of the prerequisites to voice-based digital biomarkers becoming widely adopted and commercially successful. Organizations interested in bringing such products to market, or that are already designing and developing prototypes and proofs of concept, also need to consider how they intend their intellectual property (IP) to be deployed. Depending upon which vocal features a biomarker measures and the conditions those features might indicate, IP may be used in clinical research, clinical care management, prevention and wellness programs or as part of remote monitoring approaches.

Digital health companies developing voice-based technologies need to consider go-to-market planning early in the process as biomarkers are being tested and validated. The optimal evidence-generation strategy and roadmap depend on decisions taken at the outset. Depending upon the biomarker's intended use and the clinical area being addressed, the developer will need to demonstrate different levels of robustness (for instance, randomized clinical trials, n-of-1 trials, observational studies or case-control studies) to garner acceptance and approval from regulators, the research community and clinicians.[24]

The regulatory front presents another layer of considerations. The FDA's approach to evaluating digital health tools and technologies has evolved over the past few years, and the agency continues to modernize its framework for using real-world data. The agency is now more open to being engaged early in the development process for voice-based biomarkers, including in designing digital biomarker studies, proving clinical and analytical validity and providing evidence and documentation of data integrity with regard to novel endpoints. Nonetheless the FDA has yet to clear a voice-based digital biomarker, and moving from marketing approval to application to acceptance from healthcare professionals can be a lengthy process even in a relatively receptive regulatory environment. For innovators to be successful in this emerging space, it's essential that they not only "tick the boxes" that get their product over the regulatory speed bumps but also position the product in a way that resonates with their would-be end users.

"Like anything in medicine, you have to show that you are saving someone either time or money, or [show] that you are getting much better patient outcomes—or both," says Kaufman. "So if we can show that we can help users measure response to therapy in half the time as conventional pen-and-paper tests, then that's an important insight for pharma R&D. Same for clinicians, too: if they are prescribing an antidepressant to their patients, they're going to want to have better ways to know whether that is actually working in a shorter period of time."

All references appear in the endnotes on page 19.

In responding to these needs, it is important for digital health companies to ensure optimal integration of their voice-based digital biomarker technology within existing infrastructure and clinical workflow processes, even as those processes vary across implementation sites. The following considerations should therefore be kept top of mind at the strategic planning level:

- Operating system and sensor data collection
- Integration and interoperability with clinical information systems
- User interface and user experience for engagement purposes
- Data processing
- Security and privacy

Why digital health expertise is needed

Voice-based digital biomarkers cannot operate well if planned and developed in isolation; multiple factors affect the data collection, evidence generation, physician buy-in and commercialization necessary for their success. As with any novel technology intended to provide diagnostic or therapeutic benefit, their evolution needs to be organized from the outset and managed from research project to proof of concept to pilot to full implementation—and through the many stages in between.

And because voice-based digital biomarker technology is based on software-driven processing and interpretation, it's reasonable to expect the FDA to classify and regulate it as software as a medical device. This additional layer of complexity will require digital health companies to pay special attention preparing for and addressing the many compliance aspects that condition eventual deployment.

A final key to health outcomes is how voice biomarkers are integrated with other sources of data, with clinical workflow and with the patient care journey. Both clinical knowledge and data engineering expertise (particularly in the area of interoperability) are essential determinants for success.

When we think about voice-based digital biomarkers, we need speech language pathologists to understand the specifics about voice behavior, we need medical doctors to create the clinical protocols and we need engineers to make the technology happen. Nobody has all the knowledge to do it by themselves, so you need someone who can see the whole picture and can understand how to make this information actionable for the system.

– Nathalia Nunes, MPH, speech language pathologist and author

All references appear in the endnotes on page 19.

The sooner an organization can find an expert partner to oversee the alignment of these prerequisite functions, the more likely it will be to successfully map and plan the myriad issues of clinical evidence, regulatory approval and compliance and security, and the sooner it can bring voice-based digital biomarkers to market.

FIGURE 2:

Companies working actively to develop speech biomarkers

Company	Conditions with developed and pipeline products
Vocalis Health	Developed: COVID-19, shortness of breath (CE marking) Pipeline: PH, depression, COPD, sleep apnea
Healthymize	Pipeline: asthma, pneumonia, and COPD monitoring
Sonde Health	Developed: respiratory health monitoring (asthma, COPD, COVID-19) Pipeline: depression, anxiety, hypertension detection, sleepiness
Winterlight Labs	Developed: cognitive impairment associated with dementia and mental illness
CompanionMx	Developed: mood and anxiety disorders
Aural Analytics	Developed: ALS, neurological and respiratory health

Endnotes

[1] "Vocal Biomarkers Market to Surpass US$ 4.2 Billion by 2027," Coherent Market Insights, November 27, 2019, https://www.coherentmarketinsights.com/press-release/vocal-biomarkers-market-to-reach-us-16-billion-by-2028-17.

[2] Andrea Coravos, Sean Khozin and Kenneth D. Mandl, "Developing and adopting safe and effective digital biomarkers to improve patient outcomes," npj Digital Medicine 2, no. 14 (2019), https://doi.org/10.1038/s41746-019-0090-4.

[3] Teresa Wang, Tej Azad and Ritu Rajan, "The Emerging Influence of Digital Biomarkers on Healthcare," Rock Health, accessed on June 11, 2021, https://rockhealth.com/reports/the-emerging-influence-of-digital-biomarkers-on-healthcare/.

[4] John Torous et al., "New Tools for New Research in Psychiatry: A Scalable and Customizable Platform to Empower Data Driven Smartphone Research," JMIR Mental Health 3, no. 2 (2016): e16, https://doi.org/10.2196/mental.5165.

[5] Jessica Robin et al. "Evaluation of Speech-Based Digital Biomarkers: Review and Recommendations," Digital Biomarkers 4, no. 3 (2020): 99-108, https://doi.org/10.1159/000510820.

[6] Gabriela M. Stegmann et al., "Early detection and tracking of bulbar changes in ALS via frequent and remote speech analysis," npj Digital Medicine 3, no. 132 (2020), https://doi.org/10.1038/s41746-020-00335-x.

[7] Jocelin Isabel Hall et al., "The present and future of cough counting tools," Journal of Thoracic Disease 12, no. 9 (2020), https://doi.org/10.21037/jtd-2020-icc-003.

[8] Gillinder Bedi et al., "Automated analysis of free speech predicts psychosis onset in high-risk youths," npj Schizophrenia 1, no. 15030 (2015), https://doi.org/10.1038/npjschz.2015.30.

[9] Guy Fagherazzi et al., "Voice for Health: The Use of Vocal Biomarkers from Research to Clinical Practice," Digital Biomarkers 5, no. 1 (2021), https://doi.org/10.1159/000515346.

[10] Sonde Health has built what is currently one of the largest voice sample biobanks, containing more than a million samples from over 80,000 individuals labelled with clinically valid descriptors across a range of conditions. The biobank includes samples from 50,000 volunteers in India (speaking eight different languages) and is collaborating with Michigan State University to recruit another 4,000 volunteers across four African countries. The company has prioritized international data collection beyond western countries in order to minimize potential sources of bias in feature selection and modelling that are associated with language, culture and ethnicity.

[11] The Luxembourg Institute of Health is also building a multilingual audio databank with the goal of obtaining a sufficiently large number of voice samples to be able to address heterogeneity in terms of language, accent, age, gender and cultural background when developing vocal biomarker algorithms. "We are trying to see if we can identify some common features in the voice that hold true across the entire spectrum of different languages and accents. This is a very important challenge because the generalizability of research in vocal biomarkers is definitely an issue," says Fagherazzi.

[12] Regulation (EU) 2016/679 (General Data Protection Regulation) of the European Parliament and of the Council of 27 April 2016 in the current version of the OJ L 119, 4.5.2016, https://gdprinfo.eu/en-article-4.

[13] CIVIL LIABILITIES (740 ILCS 14/) Biometric Information Privacy Act, https://www.ilga.gov/legislation/ilcs/ilcs3.asp?ActID=3004&ChapterID=57.

[14] Brian M. Bot et al. "The mPower study, Parkinson disease mobile data collected using ResearchKit," Scientific Data 3, no. 160011 (2016), https://doi.org/10.1038/sdata.2016.11.

[15] Genevieve Wanucha, “Talk About a Revolution: The Future of Voice Biomarkers in the Neurology Clinic,” Dimensions Magazine, (Spring/Summer 2019): 6-8, https://indd.adobe.com/view/80228d6e-dfd7-4fd1-bfe7-57316f4e2a6b.

[16] Vojtech Illner, Pavel Sovka and Jan Rusz, “Validation of freely-available pitch detection algorithms across various noise levels in assessing speech captured by smartphone in Parkinson’s disease,” Biomedical Signal Processing and Control 58, no. 101831 (2020), https://doi.org/10.1016/j.bspc.2019.101831.

[17] Meysam Asgari, Jeffrey Kaye, Hiroko Dodge “Predicting mild cognitive impairment from spontaneous spoken utterances,” Alzheimer’s & Dementia 3, no. 2 (Feb. 2017): 219-228, https://doi.org/10.1016/j.trci.2017.01.006.

[18] Bedi, “Automated analysis of free speech.”

[19] Natália B. Mota, Mauro Copelli and Sidarta Ribeiro, “Thought disorder measured as random speech structure classifies negative symptoms and schizophrenia diagnosis 6 months in advance,” npj Schizophrenia 3, no. 18 (Apr. 2017), https://doi.org/10.1038/s41537-017-0019-3.

[20] Jaskanwal Deep Singh Sara et al., “Non-invasive vocal biomarker is associated with pulmonary hypertension,” PLoS ONE 15, no. 4 (2020): e0231441, https://doi.org/10.1371/journal.pone.0231441.

[21] Vocalis Health, “Mayo Clinic Collaborates with Vocalis Health for Clinical Development of Vocal Biomarkers,” PR Newswire, Oct 29, 2020, https://www.prnewswire.com/il/news-releases/mayo-clinic-collaborates-with-vocalis-health-for-clinical-development-of-vocal-biomarkers-301163290.html.

[22] Elad Maor et al., “Vocal Biomarker Is Associated With Hospitalization and Mortality Among Heart Failure Patients,” Journal of the American Heart Association 9, no. 7, (Apr. 2020): e013359 https://doi.org/10.1161/JAHA.119.013359.

[23] Heather Mack, “Mayo Clinic study shows voice-analyzing app may be useful in heart disease diagnosis,” MobiHealthNews, Nov 14, 2016, https://www.mobihealthnews.com/content/mayo-clinic-study-shows-voice-analyzing-app-may-be-useful-heart-disease-diagnosis.

[24] OCEBM Levels of Evidence Working Group, “The Oxford Levels of Evidence 2,” Oxford Centre for Evidence-Based Medicine, https://www.cebm.ox.ac.uk/resources/levels-of-evidence/ocebm-levels-of-evidence.

About the authors

Todd Greenwood is a senior director in ZS's Princeton, NJ office in its digital connected health practice. He now leads the evidence generation and outcomes research space, focusing on generating clinical grade evidence for the effectiveness of digital health solutions. Todd has more than two decades of experience building impact and outcome strategies for patient-centric digitally-enabled solutions and programs across dozens of different therapeutic areas. He holds a Ph.D. from Indiana University, a master's in public health focusing on health behavior and health education from the University of Michigan and a B.I.D. in industrial design from Pratt Institute.

Nathalia Nunes is a global market researcher at ZS. She holds a B.S degree in speech-language pathology from the University of Sao Paulo, an MBA in health economics from the Federal University of Sao Paulo and an executive education certificate from the Harvard T.H. Chan School of Public Health. Nathalia is also an advisor and consultant and works with companies, investors, providers and payers to solve challenges with market mapping and positioning, innovation and product launch and development. Nathalia was the head of product at Informa Markets and part of the advisory board for Berrini Ventures, the first healthcare acceleration program in Brazil. She teaches health tech, innovation and the healthcare market at Brazilian universities, such as University of Sao Paulo and Butantan Institute. In 2020, Nathalia was elected among 39 health leaders under 39 to participate in an intensive executive program by the Brazilian College of Healthcare Executives, designed to prepare the new leadership for the healthcare industry in Brazil.

We are grateful for the efforts of **Gergana Koleva** in helping with the writing and editing of this white paper.

About ZS

ZS is a professional services firm that works side by side with companies to help develop and deliver products that drive customer value and company results. We leverage our deep industry expertise, leading-edge analytics, technology and strategy to create solutions that work in the real world. With more than 35 years of experience and 10,000-plus ZSers in more than 25 offices worldwide, we are passionately committed to helping companies and their customers thrive.

Learn more: www.zs.com/digital-health