www.jcimcr.org

JCIMCR OPEN ACCESS Journal of Clinical Images and Medical Case Reports

ISSN 2766-7820

Short Commentary

Open Access, Volume 6

Voice-related biomarkers and artificial intelligence: Emerging tools in clinical medicine

***Corresponding Author: Pedersen M**
Ear-Nose-Throat Specialist, Head and Neck Surgeon Fellow of the Royal Society of Medicine, UK. Consultant, the Medical Center, Stargate 18, DK 1100 Copenhagen, Denmark.
Website: www.mettepedersen.org

Received: Jun 02, 2025
Accepted: Jun 27, 2025
Published: Jul 04, 2025
Archived: www.jcimcr.org

DOI: www.doi.org/10.52768/2766-7820/3665

Abstract

Objective: Voice-related biomarkers have the attention of clinicians and data scientists because they promise objective, low-cost insights into laryngeal and systemic health. We aimed to combine contemporary evidence, emphasizing methodological quality, regulatory readiness, and the strategic priorities of the Union of European Phoniatricians (UEP) Voice-related biomarker Committee.

Methods: Three sources were used: 1: two MEDLINE/EMBASE searches (2013-2023) for "vocal biomarkers AND artificial intelligence" (n=332) and a focus on the mostly found "voice parameters in Parkinson's disease" (PD, n=98). Study design, sample size, recording protocol, feature set, modelling strategy, and validation approach were reported. 2: The forthcoming multichapter volume book by Springer Publishers: Voice-Related Biomarkers, and 3: Fourteen UEP committee meetings (June 2023 -March 2025)

Results: The proportion of PD articles applying machine-learning rose from 8% (2013-2018) to 39% (2019-2023). Yet only 12% used external validation, three papers adopted prospective designs. Across all conditions, fundamental frequency, jitter, shimmer, Harmonics-to-Noise Ratio, the GRBAS test, Voice Handicap Index, and the Maximal Phonation Time remained the most common metrics. UEP consensus highlights multidimensional assessment and harmonised data stewardship.

Conclusions: Voice-related biomarkers are approaching clinical viability for otolaryngology, neurodegenerative, and genetic disorders, as well as general voice treatment monitoring in speaking and singing. Convergence on robust standards, transparent AI pipelines, and fit-for-purpose trials is required before regulatory clearance can be anticipated.

Introduction

Speech production integrates respiratory airflow, vocal-fold vibration, and vocal-tract resonance under fine neuromuscular control. Perturbations at any level-pulmonary, laryngeal, articulatory, or cortical-modulate the acoustic output in systematic ways. Consequently, voice can act as a window into diverse pathologies ranging from vocal-fold lesions to neurodegenerative and genetic disorders, e.g. Parkinson's Disease (PD) and even heart failure [1-3]. We aimed to combine contemporary evidence, emphasizing methodological quality, regulatory readiness, and the strategic priorities of the Union of European Phoniatricians (UEP) Biomarker Committee. The Biomarkers Committee defines biomarkers as "characteristics that are objectively measured and evaluated as indicators of normal biological processes, pathogenic processes or responses to therapeutic intervention" [1]. Digital biomarkers extend this concept to data from ubiquitous consumer devices, lowering the barrier for frequent, longitudinal monitoring. Interest in Voice-related biomarkers accelerated during the COVID-19 pandemic, when social-distancing mandates disrupted traditional clinic-based assessments and stimulated remote-care modalities [4]. Despite burgeoning research, no voice-based Artificial Intelligence (AI) application has yet received US FDA de novo clearance or Class II CE certification. Major hurdles include dataset shift, cross-linguistic generalizability, privacy concerns, and rapidly evolving AI regulations such as the EU AI Act [5]. To foster quality and reproducibility, the UEP convened a Biomarker Committee in 2023 with representation of otorhinolaryngologists, phoniatricians, and scientists. The committee prioritized, as an example, that included AI PD for its high AI prevalence of voice measured, well-characterized natural history, and objective treatment response to levodopa and deep-brain stimulation [6-8]. This article summarizes the committee's proposals as a result of 14 meetings.

Material and methods

To evaluate the literature, a phoniatrician and librarian designed a sensitive strategy combining MeSH terms and free-text synonyms for "voice measure", "biomarker", "acoustic parameter", "machine learning", and target disorders. MEDLINE and EMBASE records (January 2013–December 2023) were downloaded, deduplicated in EndNote, and screened in Covidence. Eligible studies measured at least one objective voice parameter in human subjects and reported diagnostic, prognostic, or monitoring performance. Exclusion criteria were animal models, non-English abstracts without full translation, and purely theoretical signal-processing papers without data. A piloted spreadsheet captured study design (cross-sectional, cohort, controlled trials), sample characteristics (diagnosis, age, language), recording environment, acoustic and aerodynamic as well as perceptive parameters, machine-learning algorithm, feature-selection method, validation type, and primary outcome metric. Risk of bias was assessed with PROBAST for diagnostic/prognostic models. Quantitative AI pooling was infeasible owing to heterogeneity; therefore, data were grouped under domains of acoustic, aerodynamic, and perceptual, and subgroup analyses for examination.

Results

Acoustic parameters of the Fundamental frequency were quantified in 85% of the papers, usually as mean and SD. Perturbation measures, jitter, and shimmer were reported in 61 % and 49%, respectively, reflecting their historical role in hoarseness evaluation [6]. The Voice Handicap Index (VHI) and GRBAS test, capturing perceived disability, appeared in 53% of papers, demonstrating increasing integration of patient-reported outcome measures. Airflow measurement was carried out with the Maximum Phonation Time measures (MPT). Harmonics to Noise Ratio (HNR)was measured, but only 18% of studies applied cepstral-spectral analysis despite its superior noise robustness. Machine-learning adoption: Of the 332 AI papers, support-vector machines were most common (41%), followed by random forests (22%), convolutional neural networks (17%) and transformer architectures (4%). Feature-set cardinality spanned two orders of magnitude (median 32), indicating lack of consensus on optimal descriptors. Six studies incorporated transfer learning across languages, achieving only marginal degradation in AUC when source and target languages differed by family. Validation practice included internal k-fold cross-validation, dominated (68%); external validation against independent cohorts occurred in only 12% of studies, frequently revealing AUC attrition of 0.07–0.15. Prospective data collection was rare (three papers); none were randomized [9-12]. The UEP voice-related committee proposes a multidimensional framework that is a layered model: Acoustic layer, F0, Jitter, Shimmer, and HNR [9]. Aerodynamic layer, air flow measuring with MPT [9]. Perceptual layer, GRBAS test, and VHI-30 or CAPE-V eventually anchored by rater-calibration videos [9]. A biomechanical layer of High-speed video endoscopy and Optical Coherence Tomography metrics, eventually integrated with inverse filtering, can characterize tissue viscoelasticity [13,14]. As for the aspect of future foundation models of voice-related biomarkers, standardized datasets and setups will have to be made. Alignment with e.g., CONSORT-AI and the forthcoming DECIDE-AI reporting guideline will be necessary. Till now, only 15 % of AI studies disclosed hyperparameter search space or code repositories [15,16].

Discussion and future perspectives

Voice represents a scalable signal whose production depends on the tightly coordinated performance of laryngeal, respiratory, and perceptive systems. The UEP committee indicates that Voice-related biomarkers have progressed from research curiosities to clinical candidates for everyday otolaryngology, neurodegenerative and genetic disease, and many other kinds of routine treatment monitoring in speech and singing [4,8]. Nevertheless, the translation gap remains wide. To bridge it, multilingual, demographically balanced groups of raw audio, metadata, and reference transcriptions released under FAIR licenses [12] are an aspect. This is also the case for open-source code and version-controlled data-processing notebooks, enabling regulators and clinicians to scrutinise each analytical step [17,18]. Randomized or pragmatic trials must be made that interrogate how Voice-related biomarkers feedback changes clinician behavior and improve patient-centered outcomes in communication participation. Requirements mandated by the EU AI Act should be taken into account [5], as well as ethical governance with federated-learning infrastructures to address bias, privacy, and user trust [19].

Conclusion

Based on the work of the UEP committee work on Voice-related biomarkers, voice analytics may soon stand alongside imaging, electrophysiology, and genetics as a routine diagnostic modality in otorhinolaryngology and many other fields [20].

References

1. Biomarkers Definition Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther. 2001; 69: 89-95.
2. Strimbu K, Tavel JA. What are biomarkers? Curr Opin HIV AIDS. 2010; 5: 463-6.
3. Fagherazzi G, et al. Voice as a biomarker of health: algorithms and applications. Digit Biomark. 2021; 5: 78-88.
4. Strohl M et al Implementation of Telemedicine in a Laryngology Practice during the COVIC-19 Pandemic. J Voice 2022; 36(3): 396-402.

5. European Parliament and Council. Regulation laying down harmonised rules on Artificial Intelligence (EU AI Act). Official Journal of the European Union. 2024.

6. Oguz H, et al. Acoustic analysis of voice in Parkinson's disease. Biomed Res Int. 2015; 2015: 156014.

7. Sara JDS, et al. Deep-brain stimulation effects on speech in Parkinson's disease. Mayo Clin Proc. 2023; 98: 1353-75.

8. Evangelista EG, et al. Wearable sensors and AI for motor symptom monitoring in Parkinson's disease. Otolaryngol Head Neck Surg. 2024; 171: 340-52.

9. Lechien JR, et al. Multidimensional voice assessment: current practice in Europe. Eur Arch Otorhinolaryngol. 2023; 280: 5459-73.

10. Califf RM. Biomarker translation: a regulatory science perspective. Exp Biol Med. 2018; 243: 213-21.

11. WHO IPCS. Biomarkers in risk assessment: validity and validation. Geneva: World Health Organization. 2001.

12. Gisladottir RS, et al. Cross-lingual transfer learning for voice biomarker discovery. Sci Adv. 2023; 9: 2969.

13. Pedersen M, Ghirelli V. Overview of voice parameters in Parkinson's disease. In: Voice-Related Biomarkers. Springer Publishers, Berlin, Heidelberg. In print. 2025.

14. Moerman M, Calmecacs V. The UEP Biomarker Committee: mission and objectives. In: Voice-Related Biomarkers. Springer Publishers; Berlin, Heidelberg. 2025.

15. Lee CS, Lee AY. Clinical applications of artificial intelligence in otorhinolaryngology. Transl Vis Sci Technol. 2020; 9: 9.

16. Vasey B, et al. Reporting guideline for early-stage clinical evaluation of decision support systems (DECIDE-AI). Nat Med. 2022; 28: 924-33.

17. Idrisoglu A, et al. Federated learning for cross-institutional voice-disorder detection. Med Internet Res. 2023; 25: 46105.

18. 1Zicari RV, et al. Ethical AI: a practitioner's guide for medical developers. Front Hum Dyn. 2021; 3: 673104.