Electroglottography in Medical Diagnostics of Vocal Tract Pathologies: A Systematic Review.

Electroglottography in Medical Diagnostics of Vocal Tract Pathologies: A Systematic Review

[1]Julia Zofia Tomaszewska, and Apostolos Georgakis, *London, UK*

Summary: Electroglottography (EGG) is a technology developed for measuring the vocal fold contact area during human voice production. Although considered subjective and unreliable as a sole diagnostic method, with the correct application of relevant computational methods, it can constitute a most promising non-invasive voice disorder diagnostic tools in a form of a digital vocal tract pathology classifier. The aim of the following study is to gather and evaluate currently existing digital voice quality assessment systems and vocal tract abnormality classification systems that rely on the use of electroglottographic bio-impedance signals. To fully comprehend the findings of this review, first the subject of EGG is introduced. For that, we summarise most relevant existing research on EGG with a particular focus on its application in diagnostics. Then, we move on to the focal point of this work, which is describing and comparing the existing EGG-based digital voice pathology classification systems. With the application of PRISMA model, 13 articles were chosen and analysed in detail. Direct comparison between chosen studies brought us to pivotal conclusions, which have been described in Section 5 of this report. Meanwhile, certain limitations arising from the literature were identified, such as questionable understanding of the nature of EGG bio-impedance signals. The appropriate recommendations for future work were made, including the application of different methods for EGG feature extraction, as well as the need for continuous EGG datasets development containing signals gathered in various conditions and with different equipments.

Key Words: Electroglottography–Bio-impedance–Voice pathology classification–Voice pathology detection–Deep learning–Statistical classifier–EGG signal classification–Closed quotient–CQ.

INTRODUCTION

Electroglottography (EGG) is a non-invasive and cost-effective technology for the assessment of human vocal fold vibrations generated during the phonation process. It was first introduced by Philippe Fabre in 1940 as a method proposed for registration of arterial pulse frequencies.[1] In 1957, referring to it as "high-frequency glottography", Fabre suggested the previously reported method could be applied in studies of human phonation and the function of vocal folds—in literature also referred to as vocal cords.[2] Returning as "electroglottography" and experiencing a significant surge in scientific interest during the late 1980s and early 1990s,[3] EGG emerged as a promising diagnostic tool for multitude of voice disorders.[4] Various leading electroglottography researchers argued that its use may be the crucial step toward the development of a non-invasive preliminary diagnostic tool, particularly for laryngeal dysfunction and speech pathology.[4–6] Nonetheless, there are several factors that can easily compromise the quality of the EGG signal,[7,8] which contributed to the electroglottograph never becoming a sole diagnostic tool for the medical industry. Those include abrupt corruption of recorded signal due to misplacement of electrodes,[8] delivering incorrect information on the motion of the glottis,[9] as well as susceptibility to external interference, such as equipment or ambient noise and movement artifacts.[7] The depiction and illustration of the EGG signal's waveform have also been majorly debated, contributing towards confusion regarding EGG measurements and their reliability.[9,10]

Although influenced by several factors easily degrading its reliability, electroglottography is the closest currently existing non-invasive alternative to endoscopic laryngeal imaging and glottal airflow evaluation. Providing current technological development, with a particular focus on computational advantages, the correct evaluation and interpretation of electroglottographic measurements could be the crucial step in developing a novel non-invasive glottal level assessment and diagnostic tool.

The following work investigates relevant literature on electroglottography, its function, and—most crucially—its application in glottal-level pathologies diagnostics, which is the focal point of this review. For study selection, we implemented the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) model.[11,12] In this review, we hope to elucidate the purpose of electroglottography and clear some confusion regarding its reliability in diagnostics. First, we introduce the concept of electroglottography, along its brief history, with particular focus on its use in voice pathology diagnostics. Subsequently, we describe the methods applied in this work for efficient literature review and fulfilment of this work. Finally, we investigate currently available literature on the implementation of electroglottography in novel glottal-level pathology classifiers, with elements of digital voice quality assessment systems.

Accepted for publication December 4, 2023.
From the School of Computing and Engineering, University of West London, London, UK
[1] The author conducted the research as part of a PhD at the University of West London under the Vice Chancellor Scholarship Scheme..
Address correspondence and reprint requests to Julia Zofia Tomaszewska, School of Computing and Engineering, University of West London, St Mary's Road Ealing, London W55RF, UK. E-mail: julia.tomaszewska@uwl.ac.uk
Journal of Voice, Vol xx, No xx, pp. xxx–xxx
0892-1997
https://doi.org/10.1016/j.jvoice.2023.12.004

BACKGROUND KNOWLEDGE

EGG is a method for monitoring vocal fold vibrations produced during the human phonation process. Put simply, the signal generated by an electroglottograph shows the changes in voltage or current flow as the vocal folds come into contact and separate during phonation.[5]

The signal produced by an electroglottograph is the amplitude modulation of a weak alternating high-frequency current.[10] Accordingly, the EGG signal waveform represents the variations in impedance of vocal tract and neck tissues produced in response to that current.[13] The EGG signal can also be interpreted as a representation of conductivity between vocal folds. The changes in such conductivity of the tissues in the neck area are related to the proximity of the vocal folds, thus corresponding to the opening and closing of the vocal folds; as the vocal folds come into contact, the impedance of the tissues decreases. Those measurements in turn provide insights into vocal fold behaviour, human phonation, and various aspects of voice production.

The procedure of electroglottographic evaluation involves placing two electrodes on each side of a patient's thyroid cartilage (Figure 1). One electrode serves as the source of the current (passes the voltage), while the other electrode collects the resulting electrical signals. With the application of high-frequency low-amperage electrical current through the electrodes, EGG depicts alternations happening within the vocal fold contact area (VFCA) in a form of a time-varying signal.[3] The decrease of the impedance can be observed in the closed phase of phonation (closure of the glottis – vocal folds in contact), while the increase of the impedance measurement takes place in the open phase (glottis open – no contact between the folds).[5]

For the purposes of this research, the EGG waveform will be represented in the now conventional way, where Y-axis corresponds to VFCA, meaning the rise of EGG signal's amplitude indicates the decrease of bio-impedance. The ideal stereotypical waveform of such signal can be observed in Figure 2, with all stages of glottal opening and closing described according to Childers et al,[14,15] Rothenberg,[16] and Baken.[10]

The idealised waveform shows the stages of the glottal cycle and can be interpreted as follows:[3]

a. initial contact of the lower vocal fold margins (initiation of closing phase),
b. initial contact of the upper vocal fold margins,
c. maximum contact (yet not necessarily implying actual complete contact) of the vocal folds (end of closing phase),
d. initial separation of lower vocal fold margins (initiation of opening phase),
e. initial separation of upper vocal fold margins,
f. glottis fully open with minimal contact area between vocal folds.

Interpretation of EGG Signal's Waveform

Throughout the related literature, the visual representation of the electroglottography signal varies significantly, which causes vast amount of confusion in the understanding of produced EGG waveform.[10] The inconsistency in interpretation of an electroglottographic waveform is a result of misunderstanding of the flow of the current in the electrical circuit. There are two most prevalent approaches in constructing an electrical circuit for the application of EGG; one where increasing amplitude of the produced signal is equivalent to increasing impedance, and another where an increase in the amplitude of the signal relates to an increase of the VFCA. However, since the exact construction of a circuit is closely related to electronics, it is beyond the scope of this work.

The inconsistency of EGG waveform illustration relies on the difference in Y-axis implication, thus the significance of the signal's amplitude increases. Two practiced representations of the EGG waveform include:

1) **Increase of Y-axis parameter corresponds to the increase of the impedance – Childers's representation** (Figure 3). This approach has been taken by researchers such as Childers et al,[5,6,17,18,19] Colton and Conture,[7] as well as Rothenberg in his work related to multichannel electroglottograph.[16,20] In this representation, the increasing amplitude of the EGG signal corresponds to the increasing impedance measurement.

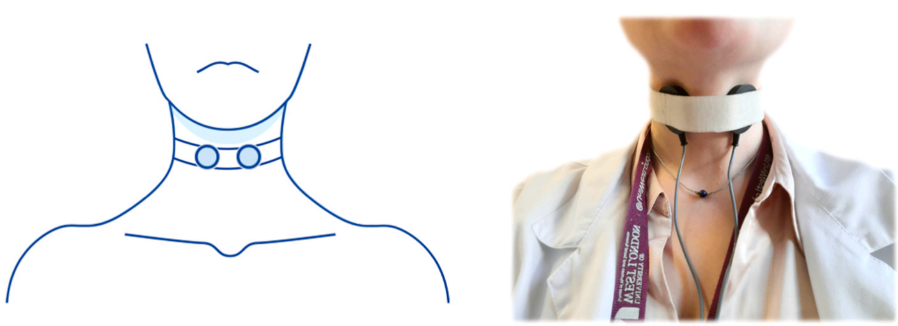

FIGURE 1. Electrodes placement in electroglottography (EGG).

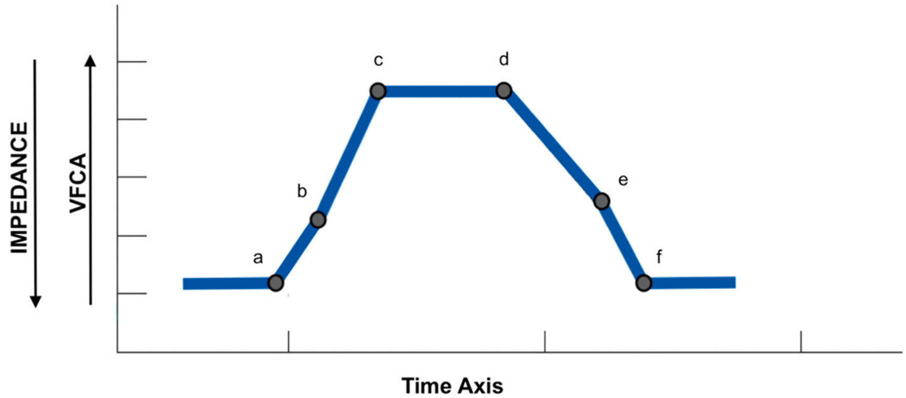

FIGURE 2. Relationship between stereotypical EGG signal's waveform, vocal fold contact area, impedance, and phases of the glottal cycle.

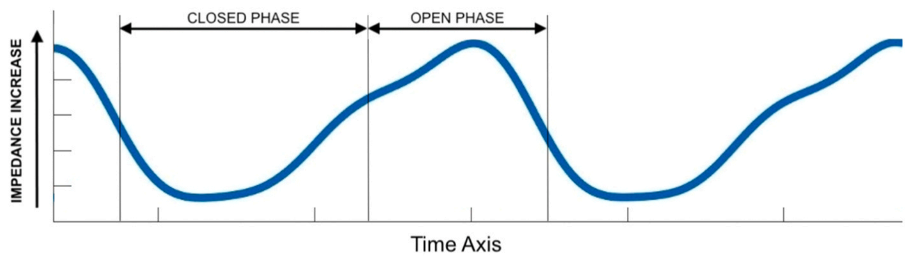

FIGURE 3. Y-axis corresponds to bio-impedance; the rise of EGG signal's amplitude indicates the increase of bio-impedance and decrease in vocal fold contact area.

2) **Y-axis depicting the vocal fold contact area (VFCA) – Fourcin's representation** (Figure 4). This representation of the EGG waveform was adopted by the inventor of laryngograph himself, Fourcin.[21–23] It is the most prevalent depiction of EGG signal in the recent literature,[3,10,24] and it is also a form of the signal produced by commonly used electroglottographs nowadays, such as Kay Pentax and Kay 6103,[25,26] as well as the Speech Studio laryngograph and software.[27] In this representation, the increase of the signal's amplitude corresponds to decreasing value of the impedance.

As shown above, most of the recent literature, as well as commonly applied electroglottographs show the increase in signal's amplitude that is parallel to the decrease in that signal's impedance. The impedance measures the opposition to alternating current presented by the combined effect of resistance and reactance in a circuit. The resistance represents a measure of the opposition to current flow in an electrical circuit, making two concepts alike. Generally, the concept of the resistance is used for direct current (DC), while the concept of the impedance can be interpreted as its equivalent for alternating currents (AC). According to Ohm's Law, the current (I) is proportional to the voltage (V), divided by the resistance (R).

$$I = \frac{V}{R}$$

Respectively, the current (I) will be proportional to the voltage (V) divided by the impedance (Z) for a circuit with alternating current, such as the EGG.[10]

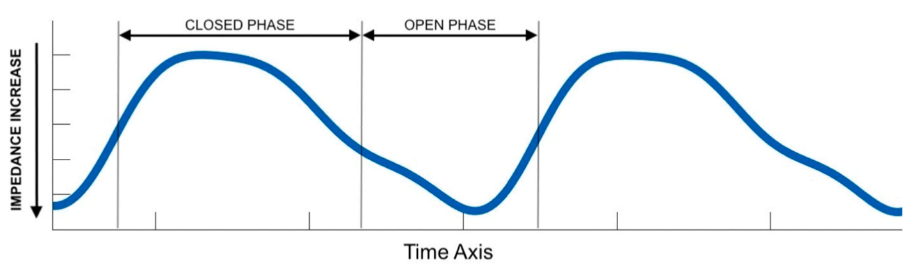

FIGURE 4. Y-axis corresponds to VFCA; the rise of signal's amplitude indicates the decrease of bio-impedance and increase in vocal fold contact area.

$$I = \frac{V}{Z}$$

Following this law, we can state that impedance is proportional to the voltage divided by the current. Provided that the measurements generated by electroglottography show the reciprocal of the impedance, it is fair to state that the EGG signal is proportional to the current divided by the voltage.

$$G = \frac{I}{V}$$

Where G stands for the measurement of conductance. The conductance is a reciprocal of resistance, thus measures the ease with which the current passes through a circuit.

Summarising the above information, although regularly interpreted as a varying impedance or simply the VFCA, the signal generated by an electroglottograph can be interpreted as a measurement of conductance.

In literature, a signal from EGG is very commonly referred to as Lx waveform.[9,10,21] This is a reference to an electroglottographic signal after required basic pre-processing. The EGG measures subtle impedance changes within the larynx; however, the signal flows from one electrode to another through a large amount of neck tissue. Therefore, the glottal impedance changes account for only about 1%–2% of the total neck impedance.[10] To extract the signal related to the activity of the vocal folds, the original signal is subjected to a high-pass filter—a filter that passes only the frequencies higher than a set threshold and attenuates all frequencies below that set threshold.[28] For that reason, the unfiltered, original output of the signal flowing through all the neck tissues is often referred to as Gx, while the intended signal acquired normally by the application of the high-pass filter (and thus, related only to the behaviour of the vocal folds) is referred to as Lx.

Nonetheless, while interpreting electroglottography measurements, it is important to note:

1. **Impedance increases during the open phase of phonation:** The EGG signal typically shows higher impedance when the vocal folds are apart or not in contact. This is because air is present between the vocal folds, creating a lower conductivity path for the electrical current.
2. **Impedance decreases during the closed phase of phonation:** When the vocal folds come into contact during phonation, the EGG signal usually shows a lower impedance. This is due to the increased conductivity caused by the presence of tissue contact, which provides a better path for the electrical current to flow.

Overview of EGG's History and Parameters

The history of electroglottography dates back to 1940s. Although invented by Fabre as a potential method for registration of arterial pulse frequencies, electroglottography quickly became a significant point of interest for many medical and vocal tract pathology researchers.

Already in his second work, Fabre admitted his then called "high-frequency glottograph" is most appropriate for studying human phonation processes.[2] His assumptions were supported between 1967 and 1970 by van Michel,[29,30] who identified and presented EGG waveform patterns of various voice pathologies, validated the EGG signals with simultaneously captured high-speed films, and developed the first electroglottograph named "Mark 4 EGG".[30]

Soon thereafter Fourcin and Abberton built and described the laryngograph, which became the first electroglottograph available on the market.[21] Fourcin and his co-workers enhanced previous findings on voice quality evaluation using EGG by publishing various waveforms for normal, breathy, creaky voice, as well as a voice affected by unilateral paralysis, laryngitis, and hoarseness.[22] Until now, Fourcin is considered one of the leading electroglottography researchers of his time, advocating the use of EGG in the rehabilitation and its monitoring in range of speech disorders.[23]

Between 1975 and 1980, the studies of electroglottography being an appropriate tool for quality and pathology voice assessment continued[9,31,32] with the emphasis on correct identification of vocal fold vibration phases (opening and closing). In 1981, Smith argued the acoustic vibrations of the larynx are too prominent to consider electroglottography a reliable diagnostic and medical assessment tool.[33] Nonetheless, in 1983 Hanson et al validated the EGG generated signals using photoglottograms and performed the calculations of open and speed quotients.[34] The researchers argued that glottographic parameters can help with diagnostic procedures, particularly in case of patients affected by voice pathologies associated with neuromuscular disorders.

Soon after work completed by Hanson, the "golden era of electroglottography" (as it is referred to in literature[3]) began. Between late 1980 and 1990, three landmark review papers were published in which all relevant methods, findings, and pitfalls of electroglottography were summarised. Those articles were written by Childers and Krishnamurthy,[19] Colton and Conture,[7] and Baken.[10] It is important to note that the "landmark papers of electroglottography" have later been expanded by Herbst, who in 2019 summarised most crucial work of 20th century and complemented it by adding most relevant developments of the past 25 years.[3]

Childers' research began early 1980 and since then provided highly accurate results in distinguishing pathological larynges from those unaffected by vocal tract abnormalities. In his work, Childers also included the EGG derivative, which supported his findings on opening and closing instants. Based on simultaneous high-speed videos, the researcher found EGG provides accurate measurements of opening and closing of the glottis. His modelling[14,15] combined with Rothenberg's observations[16] allowed for detail analysis of vocal fold instants and explicit description of glottal cycle phases.[10] Nonetheless, Childers[14,19] believed electroglottography is a representative measure of the glottal area, which was soon proven to be inadequate.[10]

In 1992, Childers undertook a development of a quantitative measurement system intended for laryngeal function assessment using EGG and speech signals.[6] This work is described further in the "Results" section of this article as a first attempt in building a digital vocal tract pathology recognition system.

In 1990, the history of electroglottography was systematised by Colton and Conture.[7] The researchers also focused on the drawback of EGG that majorly related to the correct placement of the electrodes, the signal-to-noise ratio, as well as variability within the subject groups—for instance, difficulties of obtaining undisturbed EGG signal from female participants or children. Nonetheless, the authors confirmed glottographic parameters can contribute towards the correct diagnosis of illness, such as oedema, nodules, or tumours. Furthermore, they admitted that EGG performs better than acoustic signals in fundamental frequency acquisition, as well as representation of duty cycle (the time of periodic process divided by its total period).

Baken, the author of the third landmark electroglottography review, was one of the first researchers to prove the ease with which fundamental frequency can be extracted from EGG waveform.[35,36] He suggested that electroglottography provides a mean of extracting certain features of phonatory function that are unobtainable by any other means.[10] Baken disagreed with the hypothesis put forward by Smith in 1981. Smith argued that the bio-impedance changes observed in EGG signal are not related to the vocal fold contact area, instead, they are primarily due to compression of laryngeal tissue caused by acoustic vibrations.[33] According to Baken, such hypothesis would implicate that the EGG signal is a result of a microphonic effect—the electrical change caused solely by mechanical vibrations of the system, in this case the larynx. Baken compared simultaneous recordings of EGG and an accelerometer, proving the absence of a significant microphonic component in the EGG signal. In 1992, in his landmark review, Baken concludes that while it does not provide information on the exact area of glottal opening (or the glottal space involved in opening and closing instants), applied with other appropriate tools of laryngeal observation electroglottography it can contribute significantly to clinical and therapeutics assessment.[10]

The same year, Rothenberg published his work on a tracking multichannel electroglottograph, then also referred to as TMEGG.[20] Previous EGG devices were entirely lacking the spatial resolution, but this new multichannel EGG-enabled vertical tracking of the larynx movements during voice production. Hence, as anticipated, it also enabled the verification of the correct electrode placement.[20] These finding were first questioned by Laukkanen et al in 1999, who suggested the multichannel EGG is a valid tool for analysis of the larynx's vertical movement only for sustain vowels in well-defined laboratory settings.[37] Nonetheless, the evolving research of EGG in early 2000s sought to prove spatial information can be extracted from various implementations of EGG with good results.[38,39] Subsequently, Rothenberg continued his research on electroglottography, investigating how choice of high-pass filter cut-off frequency can distort the EGG waveform, as well as methods for adequate phase correction.[40]

Following the significant improvements of computational methods of early 2000s, the subsequent investigation of EGG spread from phonation mechanisms and voice physiology to speech processing, phonetics, singing, and various medical research fields involving psychology, hearing, as well as swallowing, where EGG became a crucial non-invasive alternative for videofluorographic imaging.[41] Due to the non-invasive and cost-effective nature of EGG, its application in diagnostics research has also significantly increased. Thus, most importantly for this work, the application of EGG regenerated much interest within medical assessment of the larynx, including various developments within voice pathology classification systems.[42-44]

Although numerous papers have been published through the past three decades, their majority pertain to the physical analysis of the EGG signal itself and its directly related parameters, such as glottal closure instants (GCIs), glottal opening instants (GOIs), EGG contact quotient (later also referred to as open and closed quotients), as well as the EGG derivative (dEGG).[45] All those terms have become crucial within the research of EGG applications and can be explained as follows:

1. GCIs – temporal location of sudden vocal fold excitation that occurs during voiced phonation process. The start of closed phase.
2. GOIs – temporal location where vocal folds begin to reopen due to muscle tension and air pressure. The start of open phase.
3. Fundamental period – it is the duration between two consecutive glottal closing instants.[45]
4. Open phase – when vocal folds come apart (vocal folds in abducted position) – it is the duration between the glottal opening instant and the consecutive glottal closing instant.[45]
5. Closed phase – when vocal folds come together (vocal folds in adducted position) – it is the duration between the glottal closing instant and the consecutive glottal opening instant.
6. Open quotient – it is the ratio between open phase of vocal folds and the fundamental period of vocal fold oscillation.[45]
7. Closed quotient – it is the ratio between closed phase and the fundamental period of fold oscillation.
8. dEGG – it is the (normally first) derivative of the EGG signal, also referred to as differentiated EGG. The dEGG signal shows strong positive peaks at the beginning of EGG near-maximum slope, often associated with glottal closure, and the negative peak at the EGG signal decrease, often associated with glottal closure.[46]

Nevertheless, it is important to note that each of the described parameters is generally considered hypothetical, often changing its position along the actual EGG signal waveform depending on the researchers' approach. Their interpretation should be done with severe caution, particularly while performing statistical analysis based on any of those parameters. Their investigation is ongoing, however, as this work is designed to focus on the application of EGG in medical field, with a particular focus on voice pathology diagnostics, the introduced parameters shall not be investigated here further.

Due to the significantly increasing number of EGG-related papers being published throughout the past three decades, only some of them were chosen to be discussed further in this work. The selection criteria focused majorly on subjects related to the vocal tract physiology and EGG in diagnostics of voice pathologies. Nonetheless, some of the most crucial work related to non-pathological glottal activity, as well as EGG signal analysis topics were also investigated. The brief overview of the most important findings of EGG-related work with a particular focus on its application in diagnostics can be found in the appendix (Table 2).

Overview of EGG in Diagnostics

Electroglottography is a well-established method in research, and its application is widely spread across multiple scientific branches. Its medical utilisation, however, is still strongly debated and currently considered as non-reliable while administered on its own.

Through years of research, the potential of EGG has been investigated in various clinical subjects, including reflux,[47,48] chronic cough,[49-51] multiple sclerosis,[52] and Parkinson's disease.[53] Nonetheless, the application of EGG in diagnostics and evaluation of these pathologies most often implies its use along another well-established method, such as audio analysis[49,48,54] or stroboscopy.[4,13,32] One of most successful fields of clinical application of EGG has been detection and evaluation of various dysphonia type, including muscle tension dysphonia,[55] spasmodic dysphonia,[56] vocal fold paralysis,[57,58] and others.[6,42] Large majority of those studies unanimously showcased that EGG can be a very effective and reliable method for dysphonia detection, as well as progress monitoring throughout its treatment and recovery.

The clinical research of electroglottography also includes other cases related strictly to vocal fold physiology. Given that the EGG-derived parameters (such as GOI and GCI) mostly pertain to the contact between vocal folds and their movement patterns, the studies of EGG in relation to vocal fold physiology hold the most potential and are the main focus of this work. Those topics include, among others, vocal fold nodules and polyps,[26] Reinke's oedema, as well as laryngitis.[54]

Considering the vast development of computational methods and their application in diagnostics, this work focuses primarily on the use of EGG in novel digital diagnostic systems of glottal-level pathologies, and EGG-related voice quality assessment systems, which are further discussed in the "Results" section.

METHODOLOGY

The electroglottography, also referred to as laryngography, is generally regarded as a well-established scientific method. Particularly in the past 30 years, the number of publications on EGG and variety of its applications has increased significantly. Regardless, its application within medical field remains controversial. Considering the main focus of this review is the application of EGG within diagnostics of glottal-level pathologies, majority of newer publications unrelated to this topic were rejected. The scope of this review covers the investigation of electroglottography and its application in voice quality assessment systems and vocal tract abnormality classification systems related to upper respiratory tract disorders.

This review is divided into two parts: first investigates the beginnings of EGG and its brief history, and the second one—being the focal point of this report—focuses on the use of electroglottography in novel digital diagnostic and voice quality assessment systems that implement EGG derived parameters. For investigation of the latter one, we implemented the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) model.[11] In this report, we first introduce EGG in section titled "Background Knowledge", then, in the "Results" section, we investigate literature on the use of EGG in novel diagnostic systems of vocal tract pathologies, and EGG-related voice quality assessment systems.

Search Strategy

This review comprises two subject matters—the overview of EGG and its application in digital diagnostic systems. The literature regarding both topics was explored through PubMed, IEEE Xplore, and Scopus databases.

For the overview of EGG, the key words of "electroglottography", "glottography", "laryngography", as well as "electroglottograph", "glottograph", and "laryngograph" were used. All searches resulted in the sum of 1828 papers, from which majority were duplicates. After the removal of duplicates, the papers were reviewed manually according to their relevance to development of EGG and its role in diagnostics.

For the focal point of this review—the digital diagnostic systems of vocal tract pathologies and voice quality assessment systems, the three databases were searched with key words of "electroglottography", "electroglottograph", "laryngography", "laryngograph", "glottography", "glottograph", "vocal tract", "vocal folds", "vocal cords", "voice", "pathology", "pathologies", "disorder", "disorders", "diagnostics", "diagnosis", "detection", "classification" in various configurations using Boolean operators "OR" and "AND". These resulted in 56 publications on IEEE Xplore, 226 publications on Scopus, and 192 publications on PubMed—474 in total. There were 12 additional papers related to the research topic that were identified through citations in other full-text reviewed articles. All papers were assessed, eliminating duplicates, concluding at the count of 431. The resulting papers were then subjected to the inclusion and exclusion criteria pertaining to the research objectives.

Inclusion and Exclusion Criteria

In the following review, we consider all literature related to EGG as a main parameter in vocal tract pathology diagnostics, including findings in favour of EGG being an accurate diagnostic tool, as well as findings postulating against it.

This review is to investigate currently existing EGG-based digital vocal tract assessment systems in order to evaluate their reliability and the potential of EGG signals in digital diagnostic systems. We established three inclusion criteria and three adequate exclusion criteria to narrow down gathered records to most relevant articles.[59] The inclusion criteria were:

(a) EGG signals are primary data implemented in the study,
(b) the study concerns vocal tract and upper respiratory tract disorders or voice quality assessment,
(c) conclusions on the application of EGG as a tool for glottal-level pathology diagnostics are drawn based on digital processing of the data.

Adequately, the exclusion criteria were:

(a) EGG not being main data medium,
(b) not enough focus on vocal tract pathology or voice quality assessment,
(c) no assessment of EGG's potential in glottal-level pathologies diagnostics. This criterium of exclusion applies when a considered paper does not assess the potential and performance of a given EGG-based approach to vocal tract pathology diagnostics.

Based on the inclusion criteria, 62 full-text articles were selected and assessed, from which 13 most relevant papers were chosen to be further described in this work (Figure 5).

Information Extraction

To assess and systematise the full-text articles selected for this review, the following categories were established:

1) Population: This criterion describes the population of a specific study regarding present pathologies. It defines whether participants were affected by pathologies of vocal tract or not.

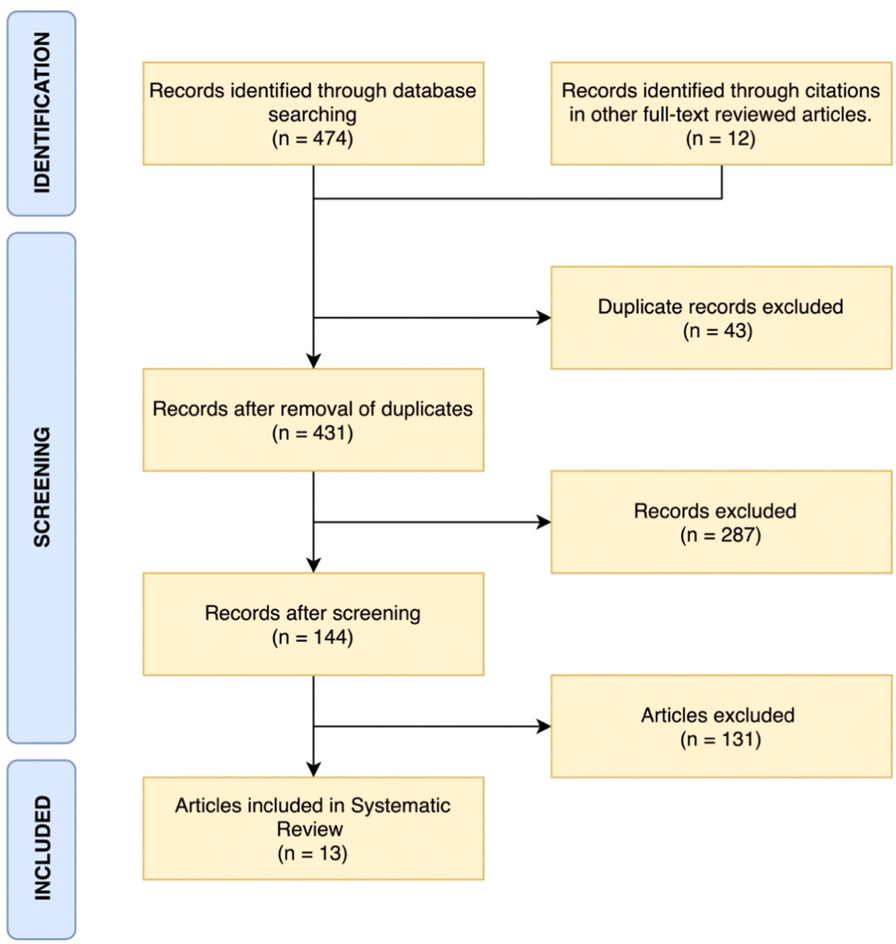

FIGURE 5. PRISMA flowchart showing article selection process for this systematic review.

2) Sample size: This parameter shows the size of population and—if relevant—lists number of participants in each study group, including the pathologies, if relevant.
3) Gender: This parameter clarifies gender of investigated study population.
4) Study objective: This category describes the hypothesis of a study. It clarifies the objective in relation to results and findings.
5) Methods: Here, methods applied within a study are described. Focus is paid to applied digital signal processing, classification methods, and the equipment used for the EGG data collection. If used in a study, pre-existing EGG datasets are also listed.
6) Findings: Here, we describe all results of a study, including statistical data, as well as drawn conclusions.

RESULTS

Overview of EGG in Diagnostics

For a thorough and reliable investigation of EGG's potential in digital diagnostic systems of vocal tract disorders, the choice of dataset is crucial. In this section, we describe the two databases that are most commonly used while investigating EGG and its application in pathology detection or classification systems.

Saarbruecken Voice Database

The Saarbruecken Voice Database was first developed and published by Manfred Pützer and Jacques Koreman in 1997 in collaboration with the Department of Phoniatrics and ENT at the Caritas Clinic St. Theresia in Saarbrücken.[60] The database is still developing and is currently managed by Manfred Pützer and William J. Barry.[61] It consists of recordings of 71 various vocal tract pathologies, collected from over 2000 subjects in total, each performing three simple tasks of sustained vowels "i", "a", "u" produced at normal, high, and low pitch, the sustained vowels "i", "a", "u" of rising-falling pitch and a recording of words "Guten Morgen, wie geht es Ihnen?" ("Good morning, how are you?"). All recordings are sampled at 50 kHz. With regards to digital systems of voice pathology detection and classification, the Saarbruecken Voice Database is applied most commonly.

Massachusetts Eye and Ear Infirmary KayPENTAX Voice Disorders Database

Another noteworthy dataset of EGG and audio samples related to various voice pathologies is the Massachusetts Eye and Ear Infirmary (MEEI).[62] It was developed by MEEI Voice and Speech Lab and the KayPENTAX Corp and released in 1994. It combines the recordings of sustained phonation of the vowel "ah" from 53 healthy participants and 657 pathological, as well as the first sentence of the "rainbow passage" gathered from 53 healthy participants and 662 pathological.[62] All recordings from MEEI are sampled at over 25 kHz, excluding 17 recordings of the "rainbow passage". This dataset is considered as landmark as it is believed to capture all the relevant American accent phenomes.

EGG-Based Digital Systems for Vocal Tract Pathology Diagnostics

The main purpose of this article was to review the currently existing literature on the use of electroglottography in digital systems for assessing voice quality and classification of vocal tract pathologies. First, most relevant work was chosen following PRISMA model, subsequently, it was systematised in a form of a table including the applied methods and study findings (Table 1). The following section describes the results of this work with regards to implemented methods.

Statistical Approach

Since the understanding of EGG signals relies on the evaluation of its quotients and other mathematically derived EGG-related parameters, first digital systems for EGG signal classification were heavily based on statistical methods. Those included evaluation of pitch period and amplitude perturbation,[6,63] calculation of speed quotient, open quotient, and closed quotient,[43] analysis of fundamental frequency and harmonic content of the EGG signal,[44] calculation of glottal instants and EGG derivative,[24] as well as various types of statistical classifiers, such as Mann-Whitney *U* test,[26] random forest classifier and Gaussian mixture model.[64]

One of the pioneering digital methods for detection of pathological EGG signal was the system developed in by Childers and Bae.[6] The authors compared the use of speech and electroglottography-derived parameters in pathological voice detection, where speech signals were analysed using Linear Predictive coding and Vector Quantisation, and EGG signals were investigated using perturbation analysis of pitch period and amplitude. To extract the EGG features, the signals were analysed in time domain visually, following by derivation of eight parameters related to signal's amplitude, the intervals between cycles, as well as open and closed phase of the signal. As the parameters were dependent on pitch period of a subject, the appropriate ratios were established. The achieved accuracy was 75.9% and 69% for speech and EGG, respectively. According to authors, the slight decrease in the system's accuracy in case of EGG signals versus the speech is due to signal's fragment selection—the selected EGG signal's portions were considered stable, while the authors believe speech disorders are best manifested in stable signal interspersed with unstable cycles.[6]

Another system taking advantage of perturbation parameters of audio and EGG signals for detection of voice pathology was the one described by Hosokawa et al.[63] The authors suggested the perturbation parameters derived

TABLE 1.
Brief Overview of Currently Existing EGG-Based Digital Voice Pathology Classification Systems

Authors	Population	Sample Size	Gender	Objective	Methods	Findings
Childers and Bae, 1992.[6]	Affected by vocal tract pathologies.	52 healthy, 29 pathological	Male and female	Quantitative assessment and detection of laryngeal function. Binary classification (healthy vs pathological).	Synchrovoice EGG Laryngeal function assessment using Linear Prediction and Vector Quantisation for speech signals, and Perturbation Analysis for EGG signals (pitch period and amplitude perturbation). Classification with two methods: closed-threshold test and discriminant analysis	Found the results for threshold method (with false alarm probability of 9.6%): For speech: 75.9% correct detection of pathological voice for pitch asynchronous LPC analysis method, 44.8% correct detection of pathological voice pitch synchronous analysis method. For EGG: 69.0% correct detection of pathological voice based on manual counting of features whose values exceeded the threshold.
Jiang et al, 1998.[43]	Affected by vocal tract pathologies.	7 healthy, 10 recurrent paralysis, 9 superior paralysis	Male and female	Integrated classification of laryngeal function based on EGG and PGG (photoglottogram). Binary classification (healthy vs paralysis).	Synchrovoice EGG Calculation of EGG and PGG derivatives (speed quotient, open quotient, closed quotient…) Digital classification system measuring match probability between new patient's signal and knowledge database by comparing area of overlap between distribution of derivative in both signals.	Found the designed integrated EGG and PGG classifier achieved 43% accuracy for healthy patients, 73% for patients with recurrent paralysis, and 57% for patients with superior paralysis. Overall, 64% accuracy for all testing dataset samples. Found the system requires training with much larger dataset, but provides potential for a future classification tool.
Ritchings et al., 2001.[44]	Affected by vocal tract pathologies.	77 pathological (at different stages of cancer treatment)	Male	Intelligent digital classification, based on EGG signals, of voice at different stages of larynx cancer treatment and recovery. Multiclass classification (stages of recovery).	Electrolaryngograph PCLX. EGG derived long-term features (fundamental frequency and its standard deviation, percentage of voiced signal), and short- term features (related to the structure of the first few harmonics, and the glottal noise). 2-layer 7-output Multi-layer Perceptron (MLP) neural networks. Validation against 7-point ranking made by speech and language therapist.	Achieved 92% accuracy. Concluded appropriate use of EGG parameters in intelligent digital diagnostic systems allows for use of EGG as a clinical tool. MLP structure: 20-nodes input layer, 25-nodes hidden layer, and 7-nodes output layer. Application of backpropagation training algorithm, softmax activation function, and cross-entropy error function. Input parameters: mean fundamental frequency, its standard deviation, voiced signal percentage, harmonic linearity measure, glottal noise, Gaussian distribution of position of first five harmonics.

TABLE 1. *(Continued)*

Authors	Population	Sample Size	Gender	Objective	Methods	Findings
Hosokawa et al., 2014.[63]	Affected by vocal tract pathologies.	39 muscle tension dysphonia, 48 vocal fold lesions, 40 healthy	Male and female	Assessment of EGG perturbation parameters (superior to acoustic) in detection of mild vocal roughness and other abnormalities of vocal tract. Multiclass assessment.	KAY Model 6103 electroglottograph. Calculation of period perturbation quotients and amplitude perturbation quotients of both audio and EGG signals (here referred to as Ac-PPQ/-APQ and EGG-PPQ/-APQ). Receiver operating characteristic (ROC) analyses for evaluation of discriminative capabilities.	Achieved the following specificity for discrimination of healthy signal and Ac-PPQ, Ac-APQ, EGG-PPQ, EGG-APQ, respectively: 80.0%, 92.5%, 90.0%, 75.0% for regular muscle tension dysphonia. 87.5%, 97.5%, 75.0%, 70.0% for regular organic dysphonia. 57.5%, 80.0%, 90.0%, 75.0% for mild muscle tension dysphonia. 72.5%, 80.0%, 75.0%, 70.0% for mild organic dysphonia. Demonstrated that both the period and the amplitude perturbation parameters of the EGG signals showed higher diagnostic accuracies than those of acoustic signals, especially for the detection of mild vocal roughness.
Deshpande and Manikandan, 2017.[24]	Unaffected by vocal tract pathologies.	3	Male and female	Automatic glottal instant detection (GCIs, GOIs, and EGG parameters) for assessment of voice pathologies.	Application of pre-existing dataset (CMU Arctic). Adaptive Variational Mode Decomposition (aVMD) for extraction of glottal waveform and suppression of low-frequency and high-frequency artifacts. Determination of glottal instants using the autocorrelation features and detecting positive and negative zero crossings of the signal.	Achieved overall accuracy of over 94% for 30–15 dB signal-to-noise ratio. Noise-free environments: identification rate of 96.49%, overall accuracy of 94.38%. Noisy environments: identification rate of 95.34%, overall accuracy of 95.06%.
Borsky et al., 2017.[64]	Unaffected by vocal tract pathologies.	28	Male and female	Assessment of voice types (modal, breathy, strained, and rough) using MFCC feature extraction on audio signals, glottal inverse filtered signals, and EGG signals. Multiclass classification (four voice types).	Media: audio, glottal inverse filtered signal, and EGG signal. Feature extraction: COVAREP feature set (only for audio), MFCCs and first-order dynamic MFCCs. Classification: Gaussian mixture model (GMM), support vector machine (SVM), random forest classifier (RF), and deep neural network (DNN).	Achieved accuracy of 56% for voice type classification based solely on EGG signals, with MFCC feature extraction. This suggested MFCC method applied on EGG gives average results. For conjoined acoustic, glottal inverse filtered and EGG signals, with application of static and dynamic MFCC coefficients achieved accuracy of 78.38% for SVM, 79.53% for RF, 56.26% for GMM, and 76.7% for DNN.

TABLE 1. (*Continued*)

Authors	Population	Sample Size	Gender	Objective	Methods	Findings
Nacci et al., 2020.[26]	Affected by vocal tract pathologies.	36 healthy, 24 functional dysphonia, 21 bilateral vocal polyps, 23 unilateral polyps, 21 unilateral cysts	Male and female	Assessment and differentiation between functional dysphonia and organic pathologies (polyps, cyst, nodules) of the vocal tract. Multiclass assessment.	KAY Model 6103 electroglottograph. Investigation of EGG signal using amplitude-speed combined analysis expressed as variability index (VI). VI compared using Kruskal-Wallis test. Mann-Whitney *U* test corrected with Bonferroni for individual pathologies (functional dysphonia, polyps, nodules, cysts)	Achieved specificity of 66.7% for VI-tot and 77.8% for VI-Q2%, proving VI-tot and VI-Q2% as highly significant. Found that EGG-derived variability index (combined analysis of amplitude and speed of vibration) is significantly greater in pathological subjects than healthy participants. Proved healthy and pathological voice can be classified correctly based on the variability of the EGG. Mann-Whitney *U* test corrected with Bonferroni showed *P* values for comparisons between functional dysphonia, polyps, nodules, cysts all below 0.021.
Muhammad and Alhussein, 2021.[65]	Affected by vocal tract pathologies.	281 healthy, 791 pathological	Male and female	Detection of voice pathology with EGG signals, voice audio signals, and bimodal approach combining EGG and audio. Comparison of modalities. Binary classification (healthy vs pathological).	Saarbruecken Voice Database. Feature extraction from spectrograms and Mel-spectrograms using CNN – harmonic distortion reduced before obtaining spectrograms, then spectrograms fed into the pre-trained Convolutional Neural Network (ResNet50, Xception, and MobileNet models tested). Classification performed with Long Short-term Memory Network.	For following modalities, with the application of pre-trained CNN as feature extractor, the system achieved accuracy of: 93.94% for audio, 93.71% for EGG, and 95.65% for fusion (EGG and audio together). Found Xception performed best for voice pathology detection from among three pre-trained CNN models (ResNet50, Xception, and MobileNet).

TABLE 1. (*Continued*)

Authors	Population	Sample Size	Gender	Objective	Methods	Findings
Miliaresi et al., 2022.[66]	Affected by vocal tract pathologies.	687 healthy, 140 laryngitis, 204 dysphonia, 210 paralysis	Male and female	Classification of voice pathologies using multimodal data (audio, medical data, and EGG features) into modular deep learning algorithms, with emphasis on EGG signals application. Multiclass classification (healthy vs dysphonia vs laryngitis vs paralysis).	Saarbruecken Voice Database. Oversampling algorithm applied to overcome dataset imbalance. For neural network, Adam gradients descent and cross-entropy. Audio (transposed into short-term feature vectors) processed to acquire MFCCs (13 features), their derivatives (13 features), and Mel filter banks (26 features) for each recording. Features transposed into image of $N \times 52$ dimensions, where N is the sequence length. Fed into 4-layer convolutional sub-network, max pooling and batch-normalisation. Medical records and jitter, fundamental frequency, and harmonic-to-noise ratio extracted from audio using autocorrelation-based algorithm. Transposed into 5×1 feature vectors, fed into two-layer feed-forward sub-network. EGG signals processed as three channel wavegrams,[67] (also attempted spectrograms and closed quotient) with dimensions of $300 \times N \times 3$, then fed as vectors in three-layer of convolutional sub-network and max pooling. All modalities fed into two final dense layers with 128 nodes, then outputted by softmax layer as one of four classes: healthy, dysphonia, laryngitis, or paralysis.	Achieved accuracy of 89.30% for multimodal network of conjoined acoustic signals, medical records, and EGG wavegrams. Unimodal network tested solely on EGG signals delivered accuracy of: 58.30% for closed quotient only, 59.40% for EGG wavegrams only, 26.50% for EGG spectrograms only. Two modal network achieved accuracy of: 81.10% for acoustic signal and closed quotient, 82.60% for acoustic signal and EGG wavegrams, 79.20% for medical records and EGG wavegrams. The results suggest EGG wavegram performs better than closed quotient and spectrograms in pathology classification tasks. Also, multimodality, with implementation of acoustic signals is highly beneficial for voice pathology classification.

TABLE 1. (*Continued*)

Authors	Population	Sample Size	Gender	Objective	Methods	Findings
Islam et al., 2022.[68]	Affected by vocal tract pathologies.	150 healthy, 30 dysphonia, 25 laryngitis, 10 vocal fold nodules.	Male and female	Detection (binary CNN classifier) and classification (multiclass CNN classifier) of voice pathologies (dysphonia, laryngitis, vocal fold nodules) based on EGG or speech signals. Binary classification (healthy vs pathological) AND multiclass classification (laryngitis vs polyp vs dysphonia).	Saarbruecken Voice Database. Feature extraction using Convolutional Neural Network. Dual cascaded Convolutional Neural Networks—first for binary classification (healthy vs pathological), second for pathologies classification (multiclass).	For binary classification, healthy vs pathological, the system achieved the accuracy of: 72.10% for EGG signals, and 80.30% for speech signals (audio). This suggests audio is a better mean for discrimination of pathological voices from healthy ones. For multiclass classification of dysphonia, laryngitis, and vocal fold nodules, the system achieved the accuracy of: 88.67% for EGG signals, and 76.48% for speech signals (audio). This suggests that EGG is a better mean for recognition of specific pathologies and their classification.
Islam et al., 2022.[69]	Affected by vocal tract pathologies.	25 healthy, 25 dysphonia	Male and female	Detection of voice pathologies using MFCCs based on speech signals or EGG signals. Binary classification (healthy vs pathological).	Saarbruecken Voice Database. Feature extraction using MFCCs for both audio and EGG. Convolutional Neural Networks classifier (binary classifier).	For binary classification between healthy and pathological voice, the system achieved the accuracy of: 50.41% for EGG signals, and 74.28% for speech signals (audio). Authors suggested audio is a better mean for discrimination of pathological voices from healthy ones.
Geng et al., 2022.[54]	Affected by vocal tract pathologies.	613 healthy, 566 pathological (leucoplakia, laryngitis, Reinke's Oedema, paralysis, vocal nodules and polyps)	Male and female	Detection and classification of voice pathologies (leucoplakia, laryngitis, Reinke's Oedema, paralysis, vocal nodules and polyps) based on multimodal approach of audio and EGG. Binary classification (healthy vs pathological) AND multiclass classification (leucoplakia, laryngitis, Reinke's Oedema, paralysis, vocal nodules and polyps).	Saarbruecken Voice Database. Mel-spectrograms derived from audio, mel-spectrograms derived from EGG. Convolutional Neural Network (pre-trained ResNet18 model) with multimodal transfer module.	Achieved the accuracy of 100% for detection of pathology/healthy, and for classification of pathologies the accuracy, recall, specificity, and F1 score were 98.02%, 98.23%, 97.82%, and 97.95%, respectively. Accuracy calculated as percentage of correctly detected pathologies (true positive) and correctly identified healthy voices (true negative).

TABLE 1. *(Continued)*

Authors	Population	Sample Size	Gender	Objective	Methods	Findings
Kumar et al., 2023.[70]	Affected by vocal tract pathologies.	303 healthy, 303 affected (Reinke's oedema, vocal fold polyp, leukoplakia, and dysphonia.)	Male and female	Binary classification system of EGG signals—pathological and non-pathological EGG. Binary classification (healthy vs pathological).	Saarbruecken Voice Database (EGG signals of vowels at normal pitch). Feature extraction: various methods assessed with minimum redundancy maximum relevance algorithm for selection of feature set, including Equivalent Rectangular Bandwidth (ERB) Spectrum and Gammatone cepstral coefficients that achieve highest feature scores. Classifiers: support vector machine (SVM), k-nearest neighbour (KNN), ensemble learner and neural networks (NN).	Achieved 93.15% accuracy for binary classification of pathological EGG signals using ensemble learner algorithm, with maximum precision of 96.70%, recall of 90.29%, and F1-score of 93.38%. Worst performing classifier was KNN, with accuracy of 86.60%. The combination of acoustic and EGG signals achieved lower accuracy of 79.97%. Results suggest EGG-based classification performs best with implementation of ERB Spectrum and Gammatone cepstral coefficients as feature extraction, and ensemble learner as classifier.

from audio are less reliable than those derived from the EGG signals, which were consistently higher in value for the subjects in dysphonic groups. Furthermore, the EGG-derived perturbation parameters exhibited greater differences between healthy participants and those affected by dysphonia, than the parameters derived from audio. The results were also assessed using the receiver operating characteristic, which showcased the EGG-derived parameters are far more accurate than audio in classification of signals affected by mild dysphonia. Nonetheless, the specificity calculated for audio can be interpreted as equally as high as for EGG-derived parameters.

A particular case of statistical approach implemented in EGG-based voice pathology detection system is the method described by Jiang et al.[43] The authors described a digital system of laryngeal pathology detection based on integration of EGG and the signals obtained from photoglottogram (PGG). To distinguish between signals gathered from healthy participants and those affected by vocal fold paralysis, the authors focused on similarities and differences in signals between two groups. Those were justified based on signals' amplitude, with focus on the last highest amplitude point and the lowest amplitude point, as well as tracking of the transition periods. The system achieved the accuracy of 43%, 73%, and 57% for the detection of healthy, recurrent paralysis, and superior paralysis, respectively,[43] whereas overall, the system performed with 64% accuracy in distinguishing between healthy and affected subjects. Although the authors argued that EGG signals can be unreliable, they also suggest the system has a great potential if provided with a large dataset and more effective control methods of data acquisition.

Considering the importance of accurate determination of EGG quotients and glottal instants in voice pathology assessment, Deshpande and Manikandan proposed an automated system for detection and extraction of these parameters, specifically for the assessment of voice pathology.[24] This subject, however, had previously been explored by Thomas and Naylor who in 2009 proposed SIGMA algorithm.[71] With the application of stationary wavelet transform, group delay function, and the detection of true and false with Gaussian mixture modelling, SIGMA detects the GCIs with the accuracy of 99.47%, while detection of GOIs, with additional processing based on detected GCIs, reaches the accuracy of 99.35%.[71] To this day, SIGMA remains one of most accurate state-of-the-art algorithms used for EGG-based glottal opening and closing detection.

On the other hand, the system proposed by Deshpande and Manikandan works in five consecutive stages: removal of high- and low-frequency artifacts, extraction of the EGG signal, detection of the glottal instants with positive and negative zero crossing detection, removal of nonglottal instants, and extraction of all EGG parameters, including fundamental frequency, closed quotient, open quotient, and speed quotient. This method achieved the identification rate of 96.49% and overall accuracy of 94.38% for signals gathered in noise-free environments, while signals collected in noisy environments yield the identification rate of 95.34% and the overall accuracy of 95.06%. Although the accuracy seems to be lower than one reported by Thomas and Naylor, Deshpande and Manikandan suggested the accuracy of SIGMA drops significantly in noisy environments, furthermore, the detection of GOIs was also reported lower than expected.

Although neither method—SIGMA or system developed by Deshpande and Manikandan—substitutes a voice pathology detection systems on its own, both deliver accurate results for EGG parameters extraction, which is considered crucial for understanding and thus classification of EGG signals.

To assess the reliability of EGG signal as a carrier of features related to dysphonia, as well as to evaluate possible differences between organic and functional dysphonia that could aid in appropriate diagnosis, Nacci et al analysed EGG data gathered from 125 subjects—36 healthy participants, 24 experiencing functional dysphonia, and 65 with various vocal fold polyps.[26] The method combined the analysis of amplitude and speed, of which the variation was expressed in a new Variability Index parameter (VI). The VI was calculated for the entire signal, and separately for each stage of the glottal cycle. According to the glottal cycle, the authors distinguished VI-Q1 parameter corresponding to the initial contact between vocal folds, VI-Q2 at the full vocal folds contact, VI-Q3 at the first phase of vocal folds disconnecting, and VI-Q4 at the last phase of glottal cycle.[26] Once calculated, the median of each parameter was derived and compared using Kruskal-Wallis test. Furthermore, the authors applied Mann-Whitney U test to evaluate the median values of VIs derived from pathological signals against the VIs obtained from healthy voices. Finally, Mann-Whitney U test corrected with Bonferroni was applied for comparison between individual pathologies (functional dysphonia, polyps, nodules, cysts). The authors found that VI calculated for the entire signal, as well as VI-Q2 were significantly higher in case of pathological signals. Kruskal-Wallis test showed statistically relevant difference for all illnesses in VI for entire signal, as well as VI-Q4. Given the achieved specificity of 66.7% for VI for entire signal and 77.8% for VI-Q2, the authors suggested these stages of the EGG signal can be highly significant for differentiation between healthy and pathological voices.

Lastly, in their work, Borsky et al compared various forms of signals and classification models to find the most effective method for classification of voice based on its breathy, strained, and rough qualities.[64,72] The authors investigated EGG, audio, and glottal inverse filtered waveform, with the application of various feature types and three types of statistical classifiers: random forest classifier (RF), support vector machines (SVM), and Gaussian mixture model (GMM). For comparison, the authors have also investigated the classification capabilities of a simple deep neural network classifier (DNN) of feed-forward architecture, with one hidden layer of 100 neurons and sigmoid activation function. Nonetheless, the authors did not investigate further the DNN classifier, as it delivered slightly lower accuracy to some of the other classifiers evaluated in this work. The authors found that COVAREP feature set (including glottal source features and harmonic model features) performed best, achieving 79.97%, 79.79%, 76.98%, and 68.12% accuracy for SVM, RF, DNN, and GMM, respectively. However, these results were only obtained from COVAREP features derived from audio signals alone—no COVAREP were tested on EGG signals. The application of MFCC features in classification of breathy, modal, and strained voice also delivered accuracy between 74% and 79%, nonetheless, it was far more successful for audio and glottal inverse filtered waveform than for the EGG signals—the classification of voice qualities based solely on EGG signals gave average accuracy of 55%–57% while classified with random forest. Furthermore, the authors argued that the classification of EGG signals performs at the average level, and even while combined with other signals no improvement of accuracy was observed.[64]

Deep Learning Approach

Most of the currently existing EGG-based digital systems of voice pathology classification that result in highest accuracy benefit from deep learning methods, involving implementation of artificial neural networks. One of the earliest systems of this type was one described by Ritchings et al.[44,73] The system was intended to give an objective assessment of voice quality in patients at different stages of cancer treatment and recovery, based solely on classification of the EGG signals. It was intended as an extension of the authors' previous work where it was found that EGG signal and its derivative parameters can be applied in Multi-layer Perceptron neural network training, resulting in accuracy of 80% in detection of pathological voice.[73] In 2001, the number of features fed into the network was expanded by adding parameters referred to by authors as "long-term"—those included mean fundamental frequency, its standard deviation, and voiced signal percentage. Previously present features (referred to as "short-term") included parameters related to the glottal noise and first five harmonics, for instance Gaussian distribution calculated with the harmonic's position, width, and amplitude. All input parameters of the proposed system included mean fundamental frequency, its standard deviation, voiced signal percentage, harmonic linearity measure, glottal noise (as a parameter derived from the fundamental-harmonic normalised spectrum, but based on the normalised noise energy), and the Gaussian distribution of position of first five harmonics. Enhanced system was intended to deliver 7-grade classification, alike one used by Speech and Language Therapists.[44] An interesting method applied by the authors to ensure the least inter-patient variability was the application of derivation of fundamental-harmonic normalised spectral representation by using an estimate of fundamental frequency for each frame of the signal.

A two-layer 7-output MLP was trained using the back-propagation algorithm, softmax activation function, and cross-entropy error function, achieving 92% accuracy of classification. The authors put great emphasis on the importance of all parameters in the process of classification; the system increased its performance accuracy from 26.5% with just one parameter of the first harmonic's Gaussian distribution, to 67.7% with five of the harmonics, to 92% with the application of all short- and long-term features derived.

Majority of most accurate voice pathology and quality digital classification systems emerged recently, as large EGG datasets became more prominent and accessible. In 2021, with the application of Saarbruecken Voice Database, the EGG-based system with 95.65% accuracy was proposed.[65] The system relied on deep neural network for the feature extraction, as well as the classification of the signals. The EGG signals, in a form of spectrograms, were to be fed into one of pre-trained convolutional neural networks—ResNet50, Xception, and MobileNet. Once generated, the features were to be fused and fed subsequently into a bidirectional long short-term memory network for their classification. The high accuracy of 95.65% of the system proposed in this work can be the result of two advantages; the large dataset of consistently gathered and pre-processed EGG signals, as well as the pre-trained neural networks utilised for the extraction of features.

A similar approach to classification of vocal pathologies and voice qualities based on the EGG signals, also utilising the Saarbruecken Voice Database, was pursued by Islam et al.[68] In this research, two types of signals were compared for their abilities to retain features relevant for voice pathology classification: the speech and the EGG. The chosen pathologies included dysphonia, laryngitis, and vocal fold polyps. The authors proposed a two-stage CNN classifier, where pathological voices were first discriminated from healthy ones in binary classification process (CNN-1), then, they were subsequently classified according to predicted pathology (CNN-2). The system was able to extract features from raw temporal signals in a form of 100 by 100 matrix, requiring no prior feature extraction. Overall, the authors reported the average accuracy of 73.33% for EGG signals in binary classification between healthy and pathological signals, while for speech signals (audio) the accuracy was 82.34% on average. According to confusion matrix, the multi-class pathology classification of EGG signals reached the accuracy of 77%, 78.67%, and 80% for laryngitis, vocal fold polyps, and dysphonia, respectively. In case of audio signals, the accuracy was of 78.83%, 63%, and 82.33% for laryngitis, vocal fold polyps, and dysphonia, respectively. The results brought authors to a conclusion that whilst audio outperforms EGG in discrimination between healthy and pathological signals, the EGG performs better in classification of those pathologies.

In a separate work, Islam et al proposed another CNN voice pathology detection system for EGG and speech signals, this time, however, the features in a form of MFCCs were to be extracted before being fed into the network.[69] The proposed system reached the accuracy of 50.41% for EGG signals (58.33% for healthy, 42.50% for detection of pathological voices), and 74.28% for speech (73.33% for healthy, 75.00% for detection of pathological signals). These results suggest EGG-derived MFCCs decrease the ability of those signals to be classified correctly according to related pathology.

An interesting matter related to voice pathology classification is multimodality. Allowing for simultaneous processing of various types of data, multimodality has become increasingly important in the application of EGG in classification systems. An example of voice pathology classification system benefiting from multimodality is the one proposed by Miliaresi et al.[66] The system also utilised Saarbruecken Voice Database and focused on classification of signals gathered from healthy participants, and those affected by dysphonia, laryngitis, and vocal fold paralysis. The system fused three modalities: audio signals, EGG signals, and medical records including demographic data on participants. First modality consisted of audio-derived MFCCs, their derivatives, and Mel filter bank outputs, that were to be processed by CNN branch built of "convolutional-max pooling-batch normalisation" layer units with rectified linear unit activation function. Second modality consisted of feed forward neural network branch processing demographic data and perturbation features, such as fundamental frequency and harmonic-to-noise ratio. In the third modality, "wavegrams" derived from EGG signals following the method described by Herbst et al[67] were processed by CNN branch with rectified linear unit activation function and global max pooling layer. All modalities were then to be concatenated into a fully connected branch with four possible outputs corresponding to three pathologies and healthy class. With the application of all three modalities, the system achieved 89.30% accuracy. While tested on EGG signals alone, the system yielded 59.40% accuracy for EGG wavegrams and 26.50% for EGG spectrograms, suggesting wavegrams retain more features relevant for signals' classification. While tested on two modalities, the system reached 82.60% accuracy for acoustic signal and EGG wavegrams, and 79.20% for medical records and EGG wavegrams.

The work of Miliaresi et al can be directly compared to that completed by Geng et al,[54] which also relies on multimodality and the application of Saarbruecken Voice Database. From among the multiple described here studies, this work implemented the widest spectrum of vocal tract pathologies, including, but not limited to leucoplakia, laryngitis, Reinke's Oedema, paralysis, vocal nodules and polyps. The two modalities applied in Geng's work were also audio and EGG, and the study also heavily relied on CNN model. However, unlike Miliaresi et al, the authors in this work benefitted from the application of the multimodal transfer module. Furthermore, both audio and EGG signals were processed in a form of Mel-spectrograms rather than MFCCs and wavegrams. The proposed system achieved the accuracy of 100% for binary classification between pathological and healthy signals, and for classification of pathologies, the accuracy reached 98.02%. Nevertheless, the authors utilised the pre-trained ResNet18 model of CNN, which can be an advantage related to higher accuracy of the system.

Most recent work on voice pathology detection system based on EGG signals is the one proposed by Kumar et al.[70] The article investigated the application of 25 various feature extraction algorithms and their influence classification of EGG signals using four classifiers: support vector machine (SVM), k-nearest neighbour (KNN), ensemble learner, and neural network. To efficiently compare the performance of features extracted using different methods, the minimum redundancy maximum relevance algorithm was applied. The score calculated with the algorithm was significantly higher for ERB Spectrum features and Gammatone cepstral coefficients than any other feature set, hence suggesting these retain most significant information for accurate EGG signal classification. The following accuracy was achieved: 93.15%, 91.15%, 90.50%, and 86.60% for ensemble learner, neural network, KNN, and SVM, respectively.

DISCUSSION AND CONCLUSIONS

This work introduced the concept of electroglottography and its brief history, with particular focus on the understanding of electroglottographic signal and its application in medical diagnostics. The focal aim of this report, however, was to gather and review most relevant work completed on digital classification systems based on the application of EGG signals, aimed at classification of voice pathologies and various vocal qualities. For that, we utilised PRISMA model, resulting in 13 different systems, that were analysed and described in Table 1 and "Results" section.

Overall, we analysed five pathology detection systems—those applying binary classification to differentiate between healthy and pathological signals; seven classification systems—those looking at multiple outputs; and one system built for EGG parameter extraction for voice pathology recognition.[24] From the five systems of binary classification, four delivered information on EGG's performance as a sole source of signals for classification.[6,65,69,70] The remaining study provided results for the system of integrated EGG and PGG.[43] Out of five binary classification systems, two applied statistical approach,[6,43] while remaining three relied on deep learning approaches. The binary classification system of the highest accuracy was one utilising pre-trained convolutional neural network and achieving over 93% accuracy in sole EGG signal application.[65]

From the described classification systems with multiple outputs, four relied on deep learning,[44,54,66,68] two on statistical methods,[26,63] and one tested both approaches.[64] Two of the deep learning classifiers were multimodal.[54,66] The classification system achieving the highest accuracy—that of 98.02%—was one proposed by Geng et al,[54] which also implemented the pre-trained CNN model (ResNet18).

Summarising all systems described, the assumption could be made that deep learning methods achieve higher accuracy in pathology detection and classification from EGG signals, than statistical methods. The accuracy of EGG-based voice pathology classification with the application of statistical classifiers was 69.0% using perturbation analysis,[6] 64% using match probability between new patient's signal and knowledge database,[43] 56% using random forest classifier,[64] as well as specificity between 66.7% and 77.8% using amplitude-speed combined analysis and Mann-Whitney *U* test corrected with Bonferroni.[26] Another instance of statistical approach was presented by Hosokawa et al[63] who focused on period perturbation quotients and amplitude perturbation quotients, achieving between 70% and 90% accuracy depending on an illness detected.[63] Most deep learning approaches to EGG-based voice pathology classification reached close to 90% accuracy, as follows: 92% using Multi-layer Perceptron fed with 10 different parameters[44]; 93.71% using spectrograms and DNN for the feature extraction and pre-trained DNN, such as ResNet50, Xception, and MobileNet; 88.67% using CNN for both feature extraction and classification[68]; 93.15% using ensemble learner algorithm.[70] Special cases of deep learning algorithms for EGG signals classification are those proposed by Miliaresi et al[66] and Geng et al[54]—both systems relied on multimodality of audio and EGG signals, achieving the accuracy of 89.30% and 98.02%, respectively. The first system processed EGG signals in a form of wavegrams and utilised CNN model, while the second employed EGG-derived Mel-spectrograms and multimodal transfer module. Furthermore, the system proposed by Geng et al was built with the application of a pre-trained ResNet18 model, which allowed the system to reach higher accuracy quicker than a newly created network.

We noticed many of the systems we investigated also compared the performance of EGG signals in voice pathology classification against another type of data, such as audio.[64,66,68,69] In these cases, the application of audio or speech signals often increased the accuracy of the system, or simply, the application of audio outperformed the sole use of EGG signals. Nevertheless, those studies applied similar or identical feature extraction methods to both audio and EGG, which could have been more suitable for audio than bio-impedance measurements.

Most of the systems described in this work, although also investigated the use of EGG signals, heavily relied on the application of audio or speech signals. This leads us to conclusion there are very few reliable systems that can accurately classify the EGG signals alone. Furthermore, many of these systems attempt to compare EGG's classification capabilities with those of audio, often noting that classification of the features derived from audio achieves greater accuracy.[64,66,68,69] Nevertheless, this could be due to a number of reasons, including the use of feature extraction methods that are likely not suited for bioimpedance signals, for example, the MFCCs.

An important element of signal classification, especially while supplemented with a deep learning model, is the extraction of features. We had noticed many authors pursuing EGG-related research follow approaches normally applied in speech processing, such as MFCCs. However, while comparing the work of Islam et al,[68,69] the conclusion can be drawn that these feature extraction methods seem to decrease the effectiveness of EGG-based classification systems. Another research suggests Mel-spectrograms derived from EGG, in a right classification setting, can reach the accuracy as high as 100%.[54] Nevertheless, although indeed produced by human phonatory system, the EGG signal is a bio-impedance or conductance measurement. We therefore believe other methods of feature extraction, those not necessarily associated with digital processing of speech, could perform better considering the nature of EGG. Examples of such methods include Equivalent Rectangular Bandwidth (ERB) Spectrum and Gammatone cepstral coefficients, which have been proven to perform better on EGG signals by Kumar et al.[70] Other examples of such feature extraction worth investigating could be Linear Predictive Coding, Linear Predictive Cepstral Coefficients, or Gammatone Frequency Cepstral Coefficients.

Another crucial parameter in classification systems is the dataset itself. Many of the systems described in this report, and almost all those utilising deep learning approach, focus on exploring one existing dataset—Saarbruecken Voice Database. Although this dataset is considered highly reliable, we believe the development of other databases, possibly utilising different equipment, could be of high relevance. Additionally, the change of recording environment to one that best reflects usual hospital settings should be investigated further. We also believe that selecting an existing dataset often implies the data have already been pre-processed, which may limit the development of the study or be misleadingly advantageous given that the data could have been pre-processed to obtain better results. Furthermore, the use of new data allows for examining and documenting the impact of data pre-processing methods on the final result of the system.

Lastly, we noticed certain limitations arising from the investigated literature. One, described already, relates to limited datasets. We believe the currently existing voice pathology classification systems could benefit from being trained and tested on new datasets, possibly of signals recorded with various pieces of equipment and in different environments. Another limitation is an arguable misinterpretation of electroglottographic signals, often treated as speech. We believe this research area could benefit from more investigation into the nature of electroglottography overall, as well as the appropriate methodology pertaining to EGG feature extraction. Other challenges include the sensitivity of EGG measurements to electrode placement and signal artifacts due to movement or poor contact. Additionally, while EGG provides valuable information about vocal fold contact, it may not directly reflect vocal fold dynamics or subtle changes in pathology. Despite its usefulness, the current literature highlights the need for further validation studies, standardised protocols, and advancements in signal processing techniques to address these limitations and enhance the clinical applicability of EGG in diagnosing vocal tract pathologies.

Considering future developments within the area of electroglottography for vocal tract pathology diagnostics, it is imperative to delve deeper into the nature of electroglottographic signals, exploring their nuanced characteristics for enhanced understanding and interpretation. A possible enhancement of validation method for the positioning of the electrodes could be of immense importance for accurate and appropriate measurement taking. Additionally, focusing on refining feature extraction methods tailored specifically for bio-impedance signals could significantly elevate the accuracy and reliability of EGG-based diagnostics. Rigorous validation of emerging EGG-driven classification systems across diverse datasets is essential, ensuring their robustness and generalisability in real-world clinical scenarios.

Recommended future research directions in the area of digital analysis of the EGG signals are (a) further investigation into the interpretation of electroglottographic signals as well as the validation of electrodes' placement, (b) improving EGG device design for better signal accuracy, (c) the exploration, development, and application of feature extraction methods that could be suited better for bio-impedance signals, as well as (d) rigorous validation of new digital EGG-based voice pathology classification systems, possibly including training and testing on various datasets.

Appendix

Table 2.
Brief History of Electroglottography and Its Medical Application in Voice Pathology Diagnostic

Authors	Population	Sample Size	Gender	Findings
Fabre, 1940.[1]	Unaffected by vocal tract pathologies.	N/A	N/A	Bio-impedance measurements were collected from tracheal level using electrodes as a method proposed for registration of arterial pulse frequencies.
Fabre, 1957.[2]	Unaffected by vocal tract pathologies.	N/A	N/A	

Table 2. (*Continued*)

Authors	Population	Sample Size	Gender	Findings
				"High-frequency glottography" was used in studies of human phonation and vocal fold function.
Chevrie-Muller, 1964.[74]	Affected by vocal tract pathologies.	N/A	N/A	Investigated the use of electroglottography in diagnosis of specific disorders, such as stuttering.
Fant et al., 1966.[75]	Affected by vocal tract pathologies.	Several	Male	Validated the EGG waveform against voice inverse filtering method.
van Michel, 1967.[29]	Affected by vocal tract pathologies.	6	Male and female	Presented waveform patterns of a hypokinetic voice disorder, hyperkinetic voice disorders including recurrent paralysis, nodule, abduction hypotonicity, and ventricular phonation, as well as subjects unaffected by voice pathologies.
Frokjaer-Jensen and Thorvaldsen, 1968.[76]	N/A	N/A	N/A	Presented the electrical circuit of an electroglottograph based on Fabre's design.
van Michel and Raskin, 1969.[77]	Unaffected by vocal tract pathologies.	N/A	N/A	Developed an electroglottograph "Mark 4 EGG".
van Michel et al., 1970.[30]	Unaffected by vocal tract pathologies.	1	Male	Concluded EGG signal can be correlated with opening and closing of vocal fold based on comparison between EGG signals and simultaneously captured high-speed films.
Fourcin and Abberton, 1971.[21]	Affected by vocal tract pathologies.	Several	Male and female	Showed EGG signals (here referred to as Lx waveform) can be used for voice quality evaluation—proved that EGG signal varies depending on voice qualities; presented different waveforms for normal, breathy, and creaky voice. First application of a laryngograph—based on a pair of double electrodes, with ground reference.
Abberton and Fourcin, 1972.[22]	Affected by vocal tract pathologies.	Several	Male and female	Enhanced previous work showing EGG signal waveforms of unilateral paralysis, laryngitis, hoarse voice, and a deaf speaker. Explained how EGG can be used to extract human voice fundamental frequency, suggesting for the purposes of fundamental frequency extraction the EGG signal is simpler than acoustic signal.
Lecluse et al, 1975.[9]	Unaffected by vocal tract pathologies.	1	N/A	Appropriately represented the opened/closed phase, where highest amplitude of EGG signal corresponds to the lowest impedance measurement, meaning closed phase (**Y-axis corresponding to value of vocal fold contact area**). One of EGG model used on excised larynx exhibited responses to acoustic vibrations and demonstrated variations in waveforms across different vowels. The remaining instruments provided indications of vocal fold contact.
Wechsler, 1976.[31]	Affected by vocal tract pathologies.	20	Male and female	Noted differences in frequency distribution in patients experiencing vocal tract pathologies before and after voice therapy. Argues that EGG can detect anomalous laryngeal function even when voice appears normal.
Pederson, 1977.[32]	Unaffected by vocal tract pathologies.	20	Male and female	Confirmed EGG signal can be correlated with opening and closing of vocal folds based on comparison between EGG

Table 2. (*Continued*)

Authors	Population	Sample Size	Gender	Findings
				signals and stroboscope. Defined sequential stages of the opening and closing of vocal folds. Concluded electroglottography may be useful in medical applications.
Rothenberg, 1981.[16]	Unaffected by vocal tract pathologies.	3	Male and female	Gave insight into two parameters recordable using EGG: air flow at the glottis and the vocal fold contact area (VFCA). Representation of EGG signal with **Y-axis corresponding to value of impedance**.
Smith, 1981.[33]	Unaffected by vocal tract pathologies.	N/A	N/A	Argued EGG is unreliable as a medical tool due to signal being influenced by acoustic vibrations of the larynx.
Hanson et al., 1983.[34]	Affected by vocal tract pathologies.	4	Male and female	Validated EGG with photoglottograms, computed open and speed quotients. Concluded open quotient differs between a patient unaffected by voice pathologies, Parkinson's disease, spastic dysphonia, and arsenic poisoning.
Smith and Childers, 1983.[18]	Affected by vocal tract pathologies.	24	Male and female	Provided results that EGG signals with the application of discriminant analysis can distinguish speakers with pathological larynges from those with larynges unaffected by pathologies with 75% accuracy. Concluded EGG may be useful in medical applications. EGG signal represented in its inverted form, where highest amplitude of EGG signal corresponds to the lowest contact area between vocal folds, meaning highest impedance measurement (**Y-axis corresponding to value of impedance**).
Childers and Larar, 1984.[5] Childers and Krishnamurthy, 1985.[19]	Affected by vocal tract pathologies. Unaffected by vocal tract pathologies.	Several	Male and female	Argued instants of glottal closure and opening can be identified from EGG by using EGG and simultaneous high-speed films. Suggested the use of EGG derivative as a meaningful parameter for medical assessment of vocal fold physiology. Explored the concepts of closed and open quotients thoroughly explaining their characteristics. Suggested EGG can assist with discrimination of pathological larynx from larynx unaffected by pathologies due to its abnormal vibratory patters. Represented the EGG signal in its inverted form, where highest amplitude of EGG signal corresponds to the lowest contact area between vocal folds, meaning highest impedance measurement (**Y-axis corresponding to value of impedance**).
Rothenberg and Mahshie, 1988.[78]	Unaffected by vocal tract pathologies.	5	Male and female	Described a method for estimating the degree of vocal fold abduction from EGG signal (**Y-axis corresponding to value of impedance**), based on a threshold method—chosen level line based on percentage of the amplitude between its minimum and maximum within one glottal cycle (50% for a normal to pressed voice and 35% for a relaxed voice). Found the method in robust, but by its nature it is imprecise and should be interpreted with care.
Titze, 1990.[28]		8		

Table 2. (*Continued*)

Authors	Population	Sample Size	Gender	Findings
	Unaffected by vocal tract pathologies.		Male and female	Proved crucial influence of electrode size and orientation on the signal-to-noise ratio and linearity of the EGG signal. Suggested better results are obtained in small inter-electrode distance and electrode angle.
Childers et al., 1990.[79]	Affected by vocal tract pathologies.	12	Male and female	Formulated mathematical equation representing a mathematical model of an EGG waveform. Suggested the use of EGG derivative as a meaningful parameter for medical assessment of vocal fold physiology. Concluded certain EGG features can be associated with vibratory characteristics of both pathological and not-pathological larynges. Represented the EGG signal in its inverted form, where highest amplitude of EGG signal corresponds to the lowest contact area between vocal folds, meaning highest impedance measurement (**Y-axis corresponding to value of impedance**).
Colton and Conture, 1990.[7]	SYSTEMATIC REVIEW	N/A	N/A	Identified and organised the pitfalls of electroglottography, including easily distorted nature of the measurements, difficulty in electrode placement, electrode-to-skin ratio influencing the measurements, as well as differences between recordings obtained from children, male, and female subjects. Showed that the presence of mucus affects the EGG signal. Confirmed the advantages of EGG, listing the accurate duty cycle, fundamental frequency acquisition remaining more accurate than its extraction from acoustic signals, as well as accurate closing time representation (**Y-axis corresponding to value of impedance**). Concluded the identification of longer closing time in EGG can constitute to accurate disclosure of illnesses such as oedema, nodules, or tumours.
Kitzing, 1990.[13]	SYSTEMATIC REVIEW with consultations with 17 specialists in the field.	N/A	N/A	Argued the use of EGG as a sole diagnostic tool in unreliable, but in conjunction with other methods as photoglottography or stroboscopy it provides valuable additional medical information unobtainable with other methods. Concluded EGG is the best method for measurement of glottal vibratory period, as well as the quotients.
Childers and Lee, 1991.[80]	Unaffected by vocal tract pathologies.	52 healthy, 23 pathological	Male and female	Used EGG to differentiate four voice types (modal, vocal fry, falsetto, and breathy) through pulse width, pulse skewness, the abruptness of glottal closure, and turbulent noise. Suggested the results of voice investigation with the application of EGG can be used for healthy vs pathological voice modelling.
Rothenberg, 1992.[20]	N/A	N/A	N/A	Developed a multichannel EGG, allowing for more pairs of electrodes to be connected. Developed alternative electrode configuration for EGG, where single electrodes can be connected either in parallel or in series. Representation of

Table 2. (*Continued*)

Authors	Population	Sample Size	Gender	Findings
				EGG signal with **Y-axis corresponding to value of impedance**.
Baken, 1992.[10]	SYSTEMATIC REVIEW	N/A	N/A	Confirmed EGG signal is an ideal mean for fundamental frequency measurement, and that is it free of supraglottal influence or other variables, such as the airflow, thus disagreeing with Smith.[33] Acknowledged two different representations of EGG waveform in relation to Y-axis implication, suggested that most appropriate representation of EGG signal is with **Y-axis corresponding to value of vocal fold contact area**.
Logemann, 1994.[41]	Unaffected by vocal tract pathologies.	N/A	N/A	Suggested EGG is a successful tool in study of swallowing, non-invasive alternative to videofluorographic imaging.
Hillman et al, 1997.[4]	N/A	N/A	N/A	Suggested EGG is a reliable clinical tool for medical diagnostic while used along videostroboscopic assessment.
Laukkanen et al, 1999.[37]	Unaffected by vocal tract pathologies.	2	Male and female	Compared Rothenberg's dual-channel EGG[20] with videofluoroscopy, confirming similar trends in larynx's vertical movements, but disagreements in the amount of these movements depending on shifts in the larynx's initial position and changes in the position of cartilages. Suggested multichannel EGG is valid in clinical application, but its applicability for studying laryngeal biomechanics is limited.
Carding et al, 1999.[42]	Affected by vocal tract pathologies.	45	Male and female	Found the EGG signal can be assessed qualitatively by clinicians to establish the process of minor non-organic laryngeal pathologies treatment. Suggested EGG signal is a suitable method for medical assessment of those illnesses and larynx function.
Rothenberg, 2002.[40]	N/A	N/A	N/A	Discussed how choice of high-pass filter cut-off frequency can distort the EGG waveform. Proposed hardware and software methods for adequate phase correction.
Zagolski and Carlson, 2002.[57]	Affected by vocal tract pathology.	16 healthy, 22 pathological	Female	Found that EGG is a reliable method for vocal fold paralysis diagnosis. Concluded EGG is a suitable tool for measuring progress during therapy of vocal fold paralysis.
Henrich et al, 2004.[45]	Unaffected by vocal tract pathology.	18	Male and female	Thoroughly investigated and discussed EGG derivative and glottal instants derived using dEGG signal. Applied a correlation-based algorithm (DECOM – DEgg Correlation-based Open quotient Measurement) to automatically calculate fundamental frequency and open quotient from dEGG. Suggested dEGG peaks are related to instants of glottal opening and closing, however, only for a healthy male voice.
Kob and Frauenrath, 2009.[38]	Unaffected by vocal tract pathology.	N/A	Male and female	Suggested multichannel EGG with 12 electrodes (36 channel measurements, time-multiplex algorithm) is a reliable tool in clinical application, the diagnosis of voice, speech, and swallowing disorders.

Table 2. (*Continued*)

Authors	Population	Sample Size	Gender	Findings
Vertigan et al, 2008.[49]	Affected by vocal tract pathology.	56 chronic cough, 8 paradoxical vocal fold movement (PVFM), 55 combined CC-PVFM, 25 muscle tension dysphonia, 27 healthy	Male and female	Found that EGG with simultaneous application of audio analysis is a suitable and effective method for chronic cough and paradoxical vocal fold movement assessment. Study based on statistical approach and manual comparison of parameters: mean fundamental frequency, standard deviation of fundamental frequency, jitter, and harmonic-to-noise ratio.
Sarvaiya et al, 2009.[81]	N/A	N/A	N/A	Published details related to EGG circuit design.
Gibson and Vertigan, 2009.[50]	Affected by vocal tract pathology.	50 untreated, 47 treated with chronic cough.	Male and female	Found that EGG-derived fundamental frequency distribution and the duration of the closed phase show no significant changes between participants before and after speech pathology treatment.
Qin et al, 2009.[82]	Unaffected by vocal tract pathology.	1	Female	Used EGG and HSV (high-speed video) integrated system for investigation of vocal fold vibration inverse parameters. Focused on glottal instants based on dEGG signal. Concluded the integrated system was more accurate than usual methods for inverse parameters of vocal fold vibration.
Thomas and Naylor, 2009.[71]	Unaffected by vocal tract pathology.	N/A	N/A	Proposed SIGMA algorithm for accurate detection of glottal opening and closing instants, with the application of multiscale analysis for singularity detection, group delay function for spike detection, and Gaussian mixture modelling for removal of detections with unlikely features. Achieved accuracy of 99.47% for GCI detection, and 99.35% for GOI detection—most accurate results for glottal instants detection up to date.
Herbst et al, 2010.[67]	Unaffected by vocal tract pathologies.	N/A	N/A	Developed "wavegrams"—a highly successful method for analysing and displaying EGG signals and their first derivatives. Wavegram image represents variations in vocal fold contact as a sequence of events changing with pitch, loudness, and voice type. It provides insight into individual glottal cycles, time-varying fundamental frequency of EGG signal, and changes of vocal fold contact phase.
Hosokawa et al, 2012.[55]	Affected by vocal tract pathologies.	19 healthy, 19 dysphonic, 19 affected by muscle tension dysphonia	Male and female	Found that EGG parameters pertaining to regularity of vocal fold vibration are a valid diagnostic tool for muscle tension dysphonia.
Ayazi et al, 2012.[47]	Affected by vocal tract pathologies.	55 healthy, 32 pathological	Male and female	Found that Gastroesophageal Reflux patients had significantly higher irregularity in both voice frequency and amplitude based on EGG measurements.
Yamout et al, 2013.[52]	Affected by vocal tract pathologies.	15 healthy, 24 pathological	Male and female	Found that EGG-derived mean closed quotient for sustained vowels [a] and [e] in multiple sclerosis and healthy participants are comparable, except in patients with dysphonia. Suggested EGG is a reliable tool for dysphonia diagnostics; however, it is not for multiple sclerosis recognition.
	N/A	N/A	N/A	

Table 2. *(Continued)*

Authors	Population	Sample Size	Gender	Findings
Herbst et al, 2014.[83]				Suggested that positive and negative dEGG peaks do not necessarily precisely coincide with events of glottal closure and initial opening. Research based on excised canine larynx, time-synchronized EGG, and ultra-HSV.
Tang et al, 2015.[84]	Unaffected by vocal tract pathologies.	N/A	Male and female	Proposed that the utilisation of EGG electrodes positioned at an angle, to be employed concurrently with an ultra-sound measurement probe placed directly on the larynx, could yield sufficient EGG waveforms encompassing significant reference points indicative of the utmost augmentation and reduction in vocal fold contact area (VFCA).
Barona-Lleo and Fernandez, 2016.[85]	Unaffected by vocal tract pathology.	44 children with ADHD, 35 non-affected children	Male and female	Showed that children with ADHD suffer significantly more often from dysphonia or hyperfunctional vocal behaviour as compared to children unaffected by ADHD. With the application of audio, EGG and endoscope, found that over 78% of ADHD-affected children suffer from vocal nodules.
Somanath and Mau, 2016.[56]	Affected by vocal tract pathologies.	12 healthy, 12 pathological	Male and female	Built a digital spasmodic dysphonia detection system based on EGG signals. Concluded EGG is unable to differentiate signals gathered from affected and unaffected by an illness participants.
Hampala et al, 2016.[86]	N/A	N/A	N/A	Investigated relation between EGG and actual vocal fold contact area, concluded EGG deviates slightly from VFCA, and although can be a reasonable first approximation, but its results must be interpreted with caution. Research based on deer larynges.
Borsky et al, 2016.[72]	Unaffected by vocal tract pathologies.	11	Male and female	Classified modal, breathy, rough, pressed, and soft voice types based on EGG signal, using MFCCs as feature extraction method, and cepstral-based features and multivariate Gaussian mixture model for classification. Achieved 83% frame-level accuracy and 91% utterance-level accuracy. Argued different voice types can be classified using MFCCs due to differences in frequency content.
Syndergaard et al, 2017.[87]	N/A	N/A	N/A	Proposed a method for VFCA vs EGG signal investigation by creating electrically conductive vocal fold replicas.
Ramirez et al, 2017.[48]	Affected by vocal tract pathologies.	17 healthy, 17 pathological	Male and female	Using EGG and audio analysis, established that shimmer, jitter, open quotient, and irregularity are significantly increased in the patients with Laryngopharyngeal Reflux.
Szklanny et al, 2019.[88]	Affected by vocal tract pathologies.	37 healthy children, 57 affected by vocal fold nodules.	Male and female	Found that in evaluation of vocal fold nodules in children the EGG signals are far more accurateHSVchanges in EGG were detected in 95% of children with vocal fold nodules, while acoustic signals only confirmed the 63% of affected children. Investigated EGG through closed quotient, and audio through peak slope—calculations computed and evaluated manually.

References

1. Fabre P. Sphygmographie par simple contact d'electrodes cutanees, introduisant dans l'artere de faibles courants de haute frequence detecteurs de sees variations volu- metriques. *CR Soc Biol (Paris)*. 1940;133:639–641.
2. Fabre P. Un precede electrique percutane d'inscription de l'aceolement glottique au cours de la phonation: glottogra- phie de haute frequence. Premiers resultats. *Bull Acad Med.* 1957;141:66–69.
3. Herbst CT. Electroglottography–an update. *J Voice.* 2019;34:503–526.
4. Hillman RE, Montgomery WW, Zeitels SM. Current diagnostics and office practice: appropriate use of objective measures of vocal function in the multidisciplinary management of voice disorders. *Curr Opin Otolaryngol Head Neck Surg.* 1997;5:172–175.
5. Childers DG, Larar JN. Electroglottography for laryngeal function assessment and speech analysis. *IEEE Trans Biomed Eng.* 1984;12:807–817.
6. Childers DG, Bae KS. Detection of laryngeal function using speech and electroglottographic data. *IEEE Trans Biomed Eng.* 1992;39:19–25.
7. Colton RH, Conture EG. Problems and pitfalls of electroglottography. *J Voice.* 1990;4:10–24.
8. Herbst CT, Dunn JC. Fundamental frequency estimation of low-quality electroglottographic signals. *J Voice.* 2019;33:401–411.
9. Lecluse FLE, Brocaar MP, Verschuure J. The electroglottography and its relation to glottal activity. *Folia Phoniatr.* 1975;27:215–224.
10. Baken RJ. Electroglottography. *J Voice.* 1992;6:98–110.
11. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Internal Med.* 2009;151:264–269.
12. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Int Surg J.* 2021;88:105906.
13. Kitzing P. Clinical applications of electroglottography. *J Voice.* 1990;4:238–249.
14. Childers DG, Hicks DM, Moore GP, et al. A model for vocal fold vibratory motion, contact area, and the electroglottogram. *J Acoust Soc Am.* 1986;80:1309–1320.
15. Childers DG, Alsaka YA, Hicks DM, et al. Vocal fold vibrations: an EGG model. In: Baer T, Sasaki C, Harris KS, eds. *Laryngeal Function in Phonation and Respiration.* Boston, MA: Little, Brown; 1987:181–202.
16. Rothenberg M. Some relations between glottal air flow and vocal fold contact area. *Asha Rep.* 1981;11:88–96.
17. Childers DG, Naik JM, Larar JN, et al. Electroglottography, speech, and ultra-high speed cinematography. *Vocal Fold Physiology and Biophysics of Voice.* Denver: Denver Center of Performing Arts; 1983:202–220.
18. Smith AM, Childers DG. Laryngeal evaluation using features from speech and the electroglottograph. *IEEE Trans Biomed Eng.* 1983;BME-30:755–759.
19. Childers DG, Krishnamurthy AK. A critical review of electroglottography. *Crit Rev Biomed.* 1985;12:131–161.
20. Rothenberg M. A multichannel electroglottograph. *J Voice.* 1992;6:36–43.
21. Fourcin AJ, Abberton E. First applications of a new laryngograph. *Med Biol Illus.* 1971;21:172–182.
22. Abberton E, Fourcin AJ. Laryngographic analysis of intonation. *Br J Disord Commun.* 1972;7:24–29.
23. Fourcin AJ. Laryngograph examination of the vocal fold vibration. *Ventilatory and Phonatory Control Systems: An International Symposium.* Oxford University Press; 1974:315–333.
24. Deshpande PS, Manikandan MS. Effective glottal instant detection and electroglottographic parameter extraction for automated voice pathology assessment. *IEEE J Biomed.* 2017;22:398–408.
25. Nacci A, Romeo SO, Cavaliere MD, Macerata A, Bastiani L, Paludetti G, et al. Comparison of electroglottographic variability index in euphonic and pathological voice. *Acta Otorhinolaryngol Ital.* 2019;39(6):381.
26. Nacci A, Macerata A, Bastiani L, et al. Evaluation of the electroglottographic signal variability in organic and functional dysphonia. *J Voice.* 2020;36:881-e5.
27. Laryngograph. Speech Studio. ⟨http://www.laryngograph.com/pr_studio.htm⟩. Accessed 20 November 2023.
28. Titze I. Interpretation of the electroglottographic signal. *J Voice.* 1990;4:1–9. https://doi.org/10.1016/S0892-1997(05)80076-1.
29. Van Michel CL. Morphologie de la courbe glottographique dans certains troubles fonctionnels du larynx. *Folia Phoniatr.* 1967;19:192–202.
30. Van Michel CL, Pfister KA, Luchsinger R. Electroglottographie et cinematographie laryngee ultra-rapide. *Folia Phoniatr.* 1970;22:81–91.
31. Wechsler E. A laryngographic study of voice disorders. *Br J Disord Commun.* 1976;12:9–22.
32. Pederson MF. Electroglottography compared with synchronized stroboscopy in normal persons. *Folia Phoniatr.* 1977;29:191–199.
33. Smith S. Research on the principle of electroglottography. *Folia Phoniatr.* 1981;33:105–114.
34. Hanson DG, Gerratt BR, Ward PH. Glottographic measurement of vocal dysfunction: a preliminary report. *Ann Otol Rhinol Laryngol.* 1983;92:413–420.
35. Baken RJ, Orlikoff RF. The effect of articulation on fundamental frequency in singers and speakers. *J Voice.* 1987;1:68–76.
36. Baken RJ, Orlikoff RF. Changes in vocal fundamental frequency at the segmental level: control during voiced fricatives. *J Speech Hear Res.* 1988;31:207–211.
37. Laukkanen A-M, Takalo R, Vilkman E, et al. Simultaneous videofluorographic and dual-channel electroglottographic registration of the vertical laryngeal position in various phonatory tasks. *J Voice.* 1999;13:60–71. https://doi.org/10.1016/S0892-1997(99)80062-9.
38. Kob M, Frauenrath T. A system for parallel measurement of glottis opening and larynx position. *Biomed Signal Process Control.* 2009;4:221–228. https://doi.org/10.1016/j.bspc.2009.03.004.
39. Hézard T, Hélie T, Doval B, et al. Non-invasive vocal-folds monitoring using electrical imaging methods. 100 years of electrical imaging-workshop. 2012; 1–4.
40. Rothenberg M. Correcting low-frequency phase distortion in electroglottograph waveforms. *J Voice.* 2002;16:32–36. https://doi.org/10.1016/S0892-1997(02)00069-3.
41. Logemann JA. Non-imaging techniques for the study of swallowing. *Acta Otolaryngol.* 1994;48:139–142.
42. Carding PN, Horsley IA, Docherty GJ. A study of the effectiveness of voice therapy in the treatment of 45 patients with nonorganic dysphonia. *J Voice.* 1999;13:72–104.
43. Jiang J, Tang S, Dalal M, et al. Integrated analyzer and classifier of glottographic signals. *IEEE Trans Rehab Eng.* 1998;6:227–234. https://doi.org/10.1109/86.681189.
44. Ritchings R, McGillion M, Moore C. Intelligent classification of electrolaryngograph signals. Annual International Conference of the IEEE Engineering in Medicine and Biology-Proceedings. 2001;2:1715–1718 doi: 10.1109/IEMBS.2001.1020547.
45. Henrich N, d'Alessandro C, Doval B, et al. On the use of the derivative of electroglottographic signals for characterization of nonpathological phonation. *J Acoust Soc Am.* 2004;115:1321–1332.
46. Henrich N, d'Alessandro C, Doval B, et al. Glottal open quotient in singing: Measurements and correlation with laryngeal mechanisms, vocal intensity, and fundamental frequency. *J Acoust Soc Am.* 2005;117:1417–1430.
47. Ayazi S, Pearson J, Hashemi M. Gastroesophageal reflux and voice changes: objective assessment of voice quality and impact of antireflux therapy. *J Clin Gastroenterol.* 2012;46:119–123. https://doi.org/10.1097/MCG.0b013e31822f386e.
48. Ramirez D, Jimenez V, Lopez X, et al. Acoustic analysis of voice and electroglottography in patients with laryngopharyngeal reflux. *J Voice.* 2017;32:281–284. https://doi.org/10.1016/j.jvoice.2017.05.009.

49. Vertigan AE, Theodoros DG, Winkworth AL, et al. Acoustic and electroglottographic voice characteristics in chronic cough and paradoxical vocal fold movement. *Folia Phoniatr.* 2008;60:210–216.
50. Gibson PG, Vertigan AE. Speech pathology for chronic cough: a new approach. *Pulm Pharmacol Ther.* 2009;22:159–162.
51. Vertigan AE, Theodoros DG, Winkworth AL, et al. A comparison of two approaches to the treatment of chronic cough: perceptual, acoustic, and electroglottographic outcomes. *J Voice.* 2008;22:581–589.
52. Yamout B, Al-Zaghal Z, El-Dahouk I, et al. Mean contact quotient using electroglottography in patients with multiple sclerosis. *J Voice.* 2013;27:506–511.
53. Vojtech JM, Stepp CE. Effects of age and Parkinson's disease on the relationship between vocal fold abductory kinematics and relative fundamental frequency. *J Voice.* 2022. https://doi.org/10.1016/j.jvoice.2022.03.007.In Press.
54. Geng L, Liang Y, Shan H, et al. Pathological voice detection and classification based on multimodal transmission network. *J Voice.* 2022. https://doi.org/10.1016/j.jvoice.2022.11.018.In Press.
55. Hosokawa K, Yoshida M, Yoshii T, et al. Effectiveness of the computed analysis of electroglottographic signals in muscle tension dysphonia. *Folia Phoniatr.* 2012;64:145–150. https://doi.org/10.1159/000342146.
56. Somanath K, Mau T. A measure of the auditory-perceptual quality of strain from electroglottographic analysis of continuous dysphonic speech: application to adductor spasmodic dysphonia. *J Voice.* 2016;30:770.e9–770.e21. https://doi.org/10.1016/j.jvoice.2015.11.005.
57. Zagolski O, Carlson E. Electroglottographic measurements of glottal function in vocal fold paralysis in women. *Clin Otolaryngol All Sci.* 2002;27:246–253. https://doi.org/10.1046/j.1365-2273.2002.00571.x.
58. Zagolski O. Electroglottography in elderly patients with vocal-fold palsy. *J Voice.* 2009;23:567–571.
59. Boland A, Dickson R, Cherry G. Doing a systematic review: a student's guide. Doing a Systematic Review. 2017:1–304.
60. Pützer M, Koreman J. A German database of patterns of pathological vocal fold vibration. *Phonus.* 1997;3:143–153.
61. Saarbruecken Voice Database: Handbook. 〈https://stimmdb.coli.uni-saarland.de/help_en.php4〉. Accessed 5 September 2023.
62. Massachusetts. Eye and Ear Infirmary, Voice disorders database, (Version 1.03 cd-rom), Kay Elemetrics Corp., Lincoln Park, NJ, 1994.
63. Hosokawa K, Ogawa M, Hashimoto M, et al. Statistical analysis of the reliability of acoustic and electroglottographic perturbation parameters for the detection of vocal roughness. *J Voice.* 2014;28:263-e9.
64. Borsky M, Mehta DD, Van Stan JH, et al. Modal and nonmodal voice quality classification using acoustic and electroglottographic features. *IEEE/ACM Trans Audio Speech Lang.* 2017;25:2281–2291.
65. Muhammad G, Alhussein M. Convergence of artificial intelligence and internet of things in smart healthcare: a case study of voice pathology detection. *IEEE Access.* 2021;9:89198–89209.
66. Miliaresi I, Pikrakis A, Poutos K. A Deep Multimodal Voice Pathology Classifier with Electroglottographic Signal Processing Capabilities. *7th IEEE International Conference on Frontiers of Signal Processing ((ICFSP)).* 2022:109–113.
67. Herbst CT, Fitch W, Švec JG. Electroglottographic wavegrams: a technique for visualizing vocal fold dynamics noninvasively. *J Acoust Soc Am.* 2010;128:3070–3078.
68. Islam R, Abdel-Raheem E, Tarique M. Voice pathology detection using convolutional neural networks with electroglottographic (EGG) and speech signals. *Comput Methods Programs Biomed Update.* 2022;2:100074.
69. Islam R, Abdel-Raheem E, Tarique M. Deep learning based pathological voice detection algorithm using speech and electroglottographic (EGG) signals. IEEE International Conference on Electrical and Computing Technologies and Applications (ICECTA). 2022;127–131.
70. Kumar D, Satija U, Kumar P. Analysis and classification of electroglottography signals for the detection of speech disorders. IEEE National Conference on Communications (NCC). 2023;1–6.
71. Thomas MR, Naylor PA. The SIGMA algorithm: a glottal activity detector for electroglottographic signals. *IEEE/ACM Trans Audio Speech Lang.* 2009;17:1557–1566.
72. Borsky M, Mehta DD, Gudjohnsen JP, et al. Classification of voice modality using electroglottogram waveforms. *Proc Interspeech.* 2016:1–5.
73. Ritchings RT, McGillion M, Conroy G, et al. Objective assessment of pathological voice quality. *Proc IEEE SMC99.* 1999;2:340–345.
74. Chevrie-Muller C. Etude de fonctionnement larynge chez les bègues par la méthode glottographique. *Rev Laryngol Otol Rhinol (Bord).* 1964;85:763–774.
75. Fant G, Ondrackova J, Lindqvist-Gauffin J, et al. Electrical glottography. *STL-QPSR.* 1966;7:15–21.
76. Frokjaer-Jensen B, Thorvaldsen P. Construction of a Fabre glottograph. *ARIPUC.* 1968;3:1.
77. Van Michel CL, Raskin L. L'Electroglottomètre Mark 4, son principe, ses possibilités. *Folia Phoniatr.* 1969;21:145–157.
78. Rothenberg M, Mahshie JJ. Monitoring vocal fold abduction through vocal fold contact area. *J Speech Lang Hear Res.* 1988;31:338–351.
79. Childers DG, Hicks DM, Moore GP, et al. Electroglottography and vocal fold physiology. *J Speech Lang Hear Res.* 1990;33:245–254.
80. Childers DG, Lee CK. Vocal quality factors: Analysis, synthesis, and perception. *J Acoust Soc Am.* 1991;90:2394–2410.
81. Sarvaiya J, Pandey P, Pandey V. An impedance detector for glottography. *IETE J Res.* 2009;55:100–105. https://doi.org/10.4103/0377-2063.54892.
82. Qin X, Wang S, Wan M. Improving reliability and accuracy of vibration parameters of vocal folds based on high-speed video and electroglottography. *IEEE Trans Biomed Eng.* 2009;56:1744–1754.
83. Herbst CT, Lohscheller J, Švec JG, et al. Glottal opening and closing events investigated by electroglottography and super-high-speed video recordings. *J Exp Biol.* 2014;217:955–963.
84. Tang S, Zhang C, Wang S. A preliminary study for a slantwise-placed electroglottography. *J Voice.* 2015;29:129-e19.
85. Barona-Lleo L, Fernandez S. Hyperfunctional voice disorder in children with Attention Deficit Hyperactivity Disorder (ADHD). A phenotypic characteristic? *J Voice.* 2016;30:114–119.
86. Hampala V, Garcia M, Svec J, et al. Relationship between the electroglottographic signal and vocal fold contact area. *J Voice.* 2016;30:161–171. https://doi.org/10.1016/j.jvoice.2015.03.018.
87. Syndergaard KL, Dushku S, Thomson SL. Electrically conductive synthetic vocal fold replicas for voice production research. *J Acoust Soc Am.* 2017;142:EL63–EL68.
88. Szklanny K, Gubrynowicz R, Ratyńska J, et al. Electroglottographic and acoustic analysis of voice in children with vocal nodules. *J Pediatr Otorhinolaryngol.* 2019;122:82–88.