Routine acoustic voice analysis: time to think again?
Meike Brockmann-Bauser and Michael J. Drinnan

Introduction
Acoustic analysis of the speech (voice) waveform is a common assessment technique used in the study of voice pathology [1,2]. In general, instrumental acoustic voice analysis refers to a family of computer-based techniques which measure acoustic signal properties of a spoken (prolonged) vowel or speech. It has the great attraction of potentially providing objective information about the human voice, a phenomenon that is otherwise described by subjective methods (e.g. perceptual rating of voice quality).

What is routine acoustic voice analysis?
Acoustic voice analysis has been cited 265 times in MEDLINE in the past 2 years. These techniques are used to assist differential diagnosis of voice disorders [2–5], to measure the effectiveness of interventions to treat voice disorders [6–8] and to corroborate patients' perceived voice disability [9,10].

A large variety of acoustic voice assessment techniques are available to the clinician. These include frequency and amplitude perturbation measurements (such as jitter and shimmer), estimates of the proportion of aperiodicity (as signal-to-noise ratio), spectral analysis-based techniques including cepstral analysis, and methods based on nonlinear dynamics and chaos analysis [1–4,12].

Underlying this body of literature is the assumption that pathological changes in the vocal fold mass or tension lead to increased measurable irregularity or noise in the human voice [1,13].

In studies involving clinical voice patients, jitter and shimmer remain by far the most commonly applied acoustic voice measures. Our Medline search identified 86 papers using these parameters in the past 2 years. The next most common measures were 'signal-to-noise ratio' and 'nonlinear dynamics' with 16 entries each.

Jitter and shimmer assess the involuntary variation (or perturbation) in fundamental frequency (F0) and voice intensity (sound pressure level, SPL) from one vibratory cycle to the next [14]. Both are almost exclusively measured in steady state vowels where the patients are asked to produce a long and continuous vowel (mostly /a/, /o/ or /i/) at 'comfortable pitch and loudness' [15].

Purpose of review
This paper reviews recent evidence regarding the validity and reliability of acoustic voice analysis in routine clinical assessments. The current role of jitter and shimmer, the most-used indices, and how their clinical application might be improved are evaluated.

Recent findings
Even though the evidence is limited, acoustic analysis is widely used to assist differential diagnosis, documentation and evaluation of treatment for clinical voice disorders. Recent clinical data have not shown that jitter and shimmer are absolute or independent indices of voice pathology or perceptual hoarseness. However, in pretreatment and posttreatment comparisons within patients, acoustic analysis might have value as an outcome measure.

Yet, the true value of clinical acoustic analysis might be masked by the confounding effects due to assessment system, gender, vowel and especially speaking voice intensity.

Summary
The validity of acoustic assessments in clinical applications remains unproven. Measurement reliability is still limited and might be greatly improved with relatively simple changes and consensus in measurement protocols and techniques. For instance, clinical assessment procedures and current normative values would have to be revised considering gender and vowel. Thus, future research might establish the validity and potential of clinical acoustic assessments.

Keywords
acoustic analysis, diagnostics, dysphonia, voice clinics, voice disorder

Current Opinion in Otolaryngology & Head and Neck Surgery 2011, 19:165–170

Link to publisher