Radiotherapy and Oncology 87 (2008) 230–236
www.thegreenjournal.com

Voice function

Prospective, longitudinal electroglottographic study of voice recovery following accelerated hypofractionated radiotherapy for T1/T2 larynx cancer☆

Rehan Kazi[a,b], Ramachandran Venkitaraman[a], Catherine Johnson[a], Vyas Prasad[a], Peter Clarke[a], Kate Newbold[a], Peter Rhys-Evans[a], Christopher Nutting[a], Kevin Harrington[a,b,*]

[a]*Head and Neck Unit, Royal Marsden Hospital, London, UK,* [b]*The Institute of Cancer Research, Cancer Research UK Centre for Cell and Molecular Biology, London, UK*

Abstract

***Background and purpose*: To measure voice outcomes following accelerated hypofractionated radiotherapy for larynx cancer.**

***Materials and methods*: Twenty-five patients with T1/T2 glottic cancer underwent serial electroglottographic and acoustic analysis (sustained vowel/i/ and connected speech) before radiotherapy and 1, 6 and 12 months post-treatment. Twenty-five normal subjects served as a reference control population.**

***Results*: Pre-treatment measures were significantly worse for larynx cancer patients. Median jitter (0.23% vs 0.97%, $p = 0.001$) and shimmer (0.62 dB vs 0.98 dB, $p = 0.05$) and differences in data ranges reflected greater frequency and amplitude perturbation in the larynx cancer patients. Pre-treatment Mean Phonation Time (MPT) was significantly reduced (21 s vs 14.8 s, $p = 0.002$) in larynx cancer patients. There was a trend towards improvement of jitter, shimmer and normalized noise energy at 12 months post-treatment. MPT improved but remained significantly worse than for normal subjects (21 s vs 16.4 s, $p = 0.013$). Average fundamental frequency resembled normal subjects, including improvement of the measured range (91.4–244.6 Hz in controls vs 100–201 Hz in post-treatment larynx cancer patients).**

***Conclusions*: This non-invasive technique effectively measures post-treatment vocal function in larynx cancer patients. This study demonstrated improvement of many key parameters that influence voice function over 12 months after radiotherapy.**

Keywords: Larynx cancer; Electroglottography; Radiotherapy; Voice function

Early stage (T1 or T2) glottic cancers can be successfully treated with either radical radiotherapy or laser surgery [1]. Both modalities have high local control and long-term cure rates and, therefore, the treatment choice often depends on local expertise, patient preference and functional outcome. Surgical resections are often accompanied by functional and anatomical deficits in spite of intensive rehabilitation [1]. On the other hand, radiotherapy is often preferred because it results in very high cure rates and is, perhaps, associated with less impairment of vocal function. Thus far, qualitative studies of post-radiation voice quality have been scarce and inconclusive [2–16]. As a result, there remains disagreement amongst head and neck surgeons, oncologists and speech–language pathologists with respect to the quality of the voice after radical radiotherapy.

In order to further our understanding of the influence of radiotherapy on the voice, a reliable method of voice assessment is required. Such an objective voice assessment might also be useful in aiding speech therapists treating laryngeal cancer patients. However, most published work has been derived from the analysis of the speech signal of sustained vowels. This can be inadequate as the speech in laryngeal cancer patients can be aperiodic and highly variable, resulting in a large proportion of the speech sample being excluded from analysis [17–19]. As an alternative, electroglottography (EGG) is a robust and reliable method of assessment that analyses the glottal waveform using both connected speech and a sustained vowel [17]. This is achieved by directly measuring the electrical impedance

☆ Funding: Dr. R. Kazi was supported by a research grant from the Head and Neck Cancer Research Trust.

0167-8140/$ - see front matter doi:10.1016/j.radonc.2008.01.006

across the neck of the subject. Several of the voice parameters that are now considered to be determinants of quality including ''jitter'' (a measure of the perturbation of fundamental frequency), ''shimmer'' (a perturbation peak amplitude), and normalized noise energy (NNE as a measure of signal to noise ratio) can be accurately determined using EGG. However, this method has not been used previously to analyse the speech of laryngeal cancer patients. We hypothesized that EGG analysis would overcome the problems of low sampling rates and inadequate pitch extraction algorithms that plague regular speech signal analysis methods and would provide a means of studying the voice of larynx cancer patients.

To the best of our knowledge, this is the first study that has sought to examine the effects of radical radiotherapy on the voice of patients with early stage laryngeal cancers. The primary objective was to conduct a prospective, longitudinal study of the influence of radiotherapy on the voice and to compare the data obtained with a reference population of normal subjects with no history of laryngeal disease. The secondary objective was to determine pretreatment and treatment factors predictive of voice outcomes in patients treated with radical radiotherapy for early stage larynx cancer.

Patients and methods

Patients and normal subjects

Twenty-five patients with early stage larynx cancer (T1N0 = 18; T2N0 = 7) were recruited to this study. All patients underwent examination under anaesthesia, biopsy and radiological imaging (CT or MRI) of the larynx and neck to confirm disease stage. In all cases the histology confirmed the diagnosis of invasive squamous cell cancer. Twenty-two patients were males and three were females. The median age at the time of study entry was 64 years (range: 48–79). In addition, 25 normal subjects without a prior history of laryngeal disease provided voice recordings to serve as a reference control population. The median age of the control population was 65 years (range: 33–80) and they included 17 males and 8 females. The study was approved by the local Research Ethics Committee.

Radiotherapy delivery

Patients were immobilised using a custom-made Cabulite (co-polyester) head and shoulder shell with their head in a neutral position. Prior to treatment all patients had isocentric simulator reference images taken using the Acuity digital simulator (Varian Medical Systems, Palo Alto, CA). Patients with T1N0 tumours were treated with 5×5 cm lateral opposed fields centred on the vocal cord (as determined at the time of treatment simulation). Patients with T2N0 tumours were treated with lateral opposed fields with a superior border at the hyoid bone and an inferior border at the inferior margin of the cricoid cartilage. These treatment field borders were customized when necessary to encompass any supra/subglottic extension with an adequate margin. Twenty-one patients whose treatment fields encompassed an area of 36 cm^2 (or lower) received radiotherapy to a dose of 50 Gy in 16 fractions. Four patients whose treatment fields encompassed an area of >36 cm^2 received radiotherapy to a dose of 55 Gy in 20 fractions. All treatments were delivered on a 2100 CD Varian linear accelerator (Varian Medical Systems) using 6 MV photons prescribed to the 100%.

Methods and equipment

EGG and acoustic analysis was performed using the 'Speech Studio' equipment and software provided by Laryngograph Ltd., United Kingdom [19]. All the recordings were obtained in a quiet room with the subject in a comfortable seated position. A pair of round gold-plated surface electrodes (3 cm in size) were attached on either side of the thyroid alae with a small capsule electret microphone (from Sony Ltd., United Kingdom) held in front of the chest at a constant mouth-to-microphone distance of 15 cm at an angle of 45° to receive the sound. The electrodes were supplied with an AC sinusoidal current of high alternating frequency (3 MHz) and the resultant signal was transmitted to the Laryngograph processor. The speech (Sp) and EGG (Lx) signals were acquired at 16 kHz and 16 bits signal resolution. Pitch was extracted using a dedicated hardware circuit and the period was counted by a 12 MHz clock being rounded down to single microsecond to ensure accuracy.

Voice recording protocol

The larynx cancer patients underwent voice recordings before the start of radiotherapy and at 1, 6 and 12 months post-treatment. None of the patients received voice therapy in the follow-up period. The normal subjects provided a single voice recording. In all cases, subjects provided synchronous electroglottographic and acoustic recordings in a single session. The recording protocol consisted of (1) sustained vowel/i/ for at least 5 s duration produced at a comfortable pitch and loudness; (2) maximum phonation time (s): sustained vowel/i/ produced at a comfortable pitch and loudness for as long as they could manage after maximal inspiration; (3) connected speech using the standard text provided ('Arthur the rat') read at a comfortable pitch and loudness. The EGG signal amplitude was adjusted for each individual subject to the optimal gain position prior to the actual recording.

Measures

The EGG waveform (Lx) that was obtained allowed various voice parameters to be determined from the glottic signal [17–19]. The derivation of these parameters is fundamentally different from analysis of the standard acoustic waveform. We derived the following voice parameters: (1) The average fundamental frequency (Fx) which corresponds to the tone/pitch and is calculated from the Lx waveform on a cycle by cycle basis rather than from the Fast Fourier analysis of the acoustic waveform; (2) Jitter which represents short-term frequency perturbations or changes that reflect fine motor control of the voice; (3) Shimmer which represents short-term amplitude perturbation and also reflects fine motor control of the voice; and (4) Normalized Noise Energy (NNE) which is a log ratio of energy contained in the noise to the energy in the signal. NNE is another more select/specific measure of the ''Harmonic