Bratisl Lek Listy 2011; 112 (3)
120 – 124

CLINICAL STUDY

Distinction of cough from other sounds produced by daily activities in the upper airways

Martinek J[1], Zatko T[1], Tatar M[1], Javorka M[2]

Department of Pathophysiology, Comenius University, Jessenius Faculty of Medicine, Martin, Slovakia.
martinek@jfmed.uniba.sk

Abstract: *Objectives:* The aim of this study was to validate the successfulness of our developed system for distinction between cough and other sounds which are present in daily human activities from the upper airways.
Background: To date, methods used for monitoring of cough sound were primarily subjective. A reliable measure of cough is needed so that the severity of cough in various patients and the effectiveness of treatment can be assessed.
Methods: Sounds of induced cough and sneezing, voluntary throat and nasopharynx clearing, forced ventilation and laughing, snoring, eructation, loud swallowing, and nasal blowing were studied. Characteristics of the sound events in 20 volunteers were calculated using the time-domain, spectral and non-linear analysis. The classification tree was constructed for classification between cough and non-cough sounds. We have validated the usefulness of our developed algorithm against subjective cough counts, which were performed by two trained observers.
Results: The value of sensitivity for distinction between cough and other sounds was 86 % and the value of specificity was 91 %. The value of sensitivity for distinction between voluntary and induced cough sounds was 96% and specificity was 43 %. The value of sensitivity between cough sounds and voluntary throat clearing was 96 % and specificity was 85 %. The value of sensitivity between cough sounds and induced sneezing was 95 % and specificity was 93 %.
Conclusion: We have developed an algorithm for distinction between cough and other sounds with a relatively high degree of accuracy *(Tab. 1, Fig. 5, Ref. 15)*. Full Text in free PDF *www.bmj.sk.*
Key words: cough sound, spectral analysis, non-linear analysis, classification tree.

Cough is the commonest symptom of many respiratory diseases forcing the patients to seek a medical advice. To date, the methods used to assess the cough have been primarily subjective (1). Cough monitoring systems have been proposed recently, based either on sound recordings alone (2) or on simultaneous sound and electromyography recordings (3–4). In order to make cough monitoring applicable to clinical practice, it is necessary to develop accurate automatic system for recording, detection and counting of coughs. With the availability of digital recording devices and the advances in digital storage media, battery powered mp3 recorder can be used to make a high quality ambulatory sound recordings. Data can be transferred to personal computer and the recordings can be used to develop algorithms for cough sounds identification (5).

Several systems for automatic cough recognition and monitoring based on sound recordings have been described recently (6–8). They are based on automatic cough detection algorithms that operate reliably in the ambulatory settings. Different analyses were applied to resolve the problem of recognition of cough sounds, while rejecting other sounds with similar characteristics.

Because these mathematical algorithms are not commercially available, we have recently prepared an algorithm based on the time-domain, spectral and non-linear analysis for distinction between voluntary cough sound and speech in healthy volunteers (9). Because other sounds occur by human daily activities too, the aim of this study was to confirm the effectiveness of previously developed algorithm for distinction cough sounds (induced, voluntary and spontaneous cough sounds) from induced sneezing, voluntary throat and nasopharynx clearing, voluntary forced ventilation and laughing, voluntary snoring, eructation, loud swallowing, and nasal blowing.

Materials and methods

Subjects

Our study group included 20 healthy volunteers (8 females – median age 28.5 yrs, range 20 – 36 yrs; 12 males – median age 33 yrs, range 26 – 60 yrs). All subjects were without respiratory infection at the time of the study. Forced expiratory volume during one second (FEV_1) was higher than 80 % of predicted value.

[1]Department of Pathophysiology, Comenius University, Jessenius Faculty of Medicine, Martin, Slovakiam, and [2]Department of Physiology, Comenius University, Jessenius Faculty of Medicine, Martin, Slovakia

Address for correspondence: J. Martinek, MD, Dept of Pathophysiology, Comenius University, Jessenius Faculty of Medicine, Sklabinska 26, SK-037 53 Martin, Slovakia.
Phone: +421.43.4238213

Acknowledgement: The paper is dedicated to the memory of Professor Juraj Korpas.

Fig. 1. Portable digital MP3 recorder Sony ICD-MX 20 (Sony Corporation, China).

Fig. 2. Miniature omnidirectional condenser lavalier microphone ATR35s (Audio-Technica U.S.,Inc., Philippines).

Recording system

The sound recording system consisted of a portable digital voice recorder (Sony ICD-MX20, Sony Corporation, China) with the sampling frequency of 8 kHz (Fig. 1) and a miniature omnidirectional condenser lavalier microphone (ATR35s, Audio-Technica U.S., Inc., Philippines) with a flat frequency response from 50 to 18 000 Hz (Fig. 2). The microphone was attached to the subject's chest and was covered by plastic foam membrane to suppress sounds coming from the outer environment. The sounds were stored during recording time to the MP3 recorder memory card, and after recording time the obtained sound files were transferred to the PC and converted to the WAV file format.

Protocol

The induced cough sounds were evoked by single-breath inhalation of capsaicin aerosol (250 µmol/l; bolus 1.2s) using KoKo Spirometer & KoKo Digidoser system (Ferraris Respiratory, Pulmonary Data Services, Inc. Louisville, Co, USA). Sneezing

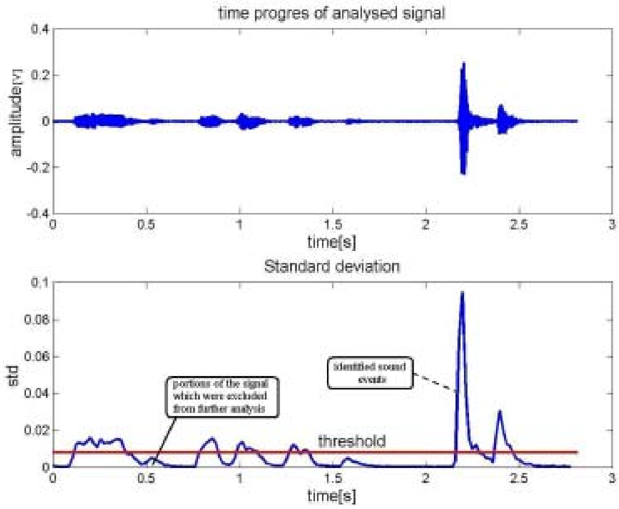

Fig. 3. The time progress of 3 sec section from the obtained sound record and the time progress of its standard deviation. For each position of the moving window (0.018s) the value of standard deviation was calculated, and compared to empirically determined threshold value. The portions of the signal containing no sound events reached only low value of standard deviation and they were excluded from further analysis (an example is indicated). The portions of the signal containing sound events reached relatively high degree of standard deviation, which was above the given threshold value, and they underwent further analysis (an example is indicated).

was induced by intranasal administration of histamine solution (8 mg/ml; 25 ml). Volunteers were requested to perform voluntary coughs, voluntary throat and nasopharyngeal clearing, voluntary forced ventilation and snoring, laughing, eructation, loud swallowing, and nasal blowing.

Determination of the sound events and sound events analysis

The selection of the sound events from the obtained sound recordings was the first step in the sound events analysis. The algorithm for determination of the sound events from the raw records and sound events analysis was described in our previous study (9). Briefly, we determined the sound events using the moving window, which moved over whole audio signal without overlap. For each position of the moving window, the value of standard deviation (SD) was calculated and compared to our empirically determined threshold value. The portions of the signal containing no sound events showed only small SDs related to the inherent noise present in the signal. These portions were below the threshold value, and they were excluded from further analysis. The portions of recordings containing sound events reached relatively high values of SDs. Identified sound events were stored as separate files and underwent further analysis (Fig. 3).

The characteristics of the identified sound events were calculated using time-domain, spectral and non-linear analysis. For each identified sound event, its duration was quantified (parameter *length*) and the time progress of the power spectral density (PSD) was determined. Moving averaging further smoothed PSD.

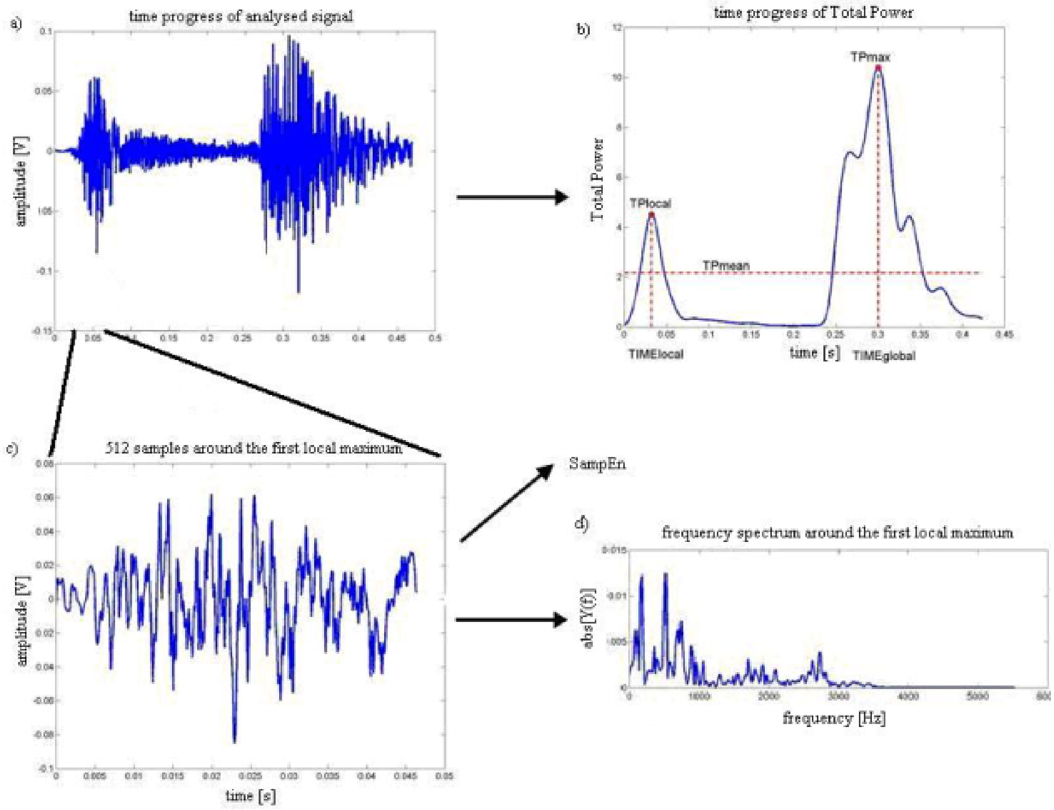

Fig. 4. Example of the characteristic parameters calculation from the determined sound event (a). These parameters were calculated using time-domain, spectral and non-linear analysis. From the sound event, the time progress of Total Power was determined (b). From the Total Power curve, the value of first local maximum (*TPlocal*) and the time of its occurrence (*TIMElocal*), the value of global maximum (*TPmax*) and the time of its occurrence (*TIMEglobal*) and the mean value from all Total Power values (*TPmean*) were determined. From the 512 samples corresponding to the local maximum (c,) the value of Sample Entropy was determined. From the frequency spectrum (d) calculated from these 512 samples, the values of skewness and kurtosis were determined.

Total power (TP) corresponding to the area under the PSD curve was computed as a measure of the sound intensity. From the TP curve parameters *TPmax* (maximal total power), *TPmean* (mean total power) and the time of the first local and global maximum occurrence – $time_{local}$ *and* $time_{global}$ were computed. The ratio of the sum of TPs of all the local maxima divided by the sum of TPs of all local minima in given sound events gives the parameter *ratio*. The value of the first local maximum of TP was divided by the time of its occurrence (from the start of the sound event) – this parameter is named *slope*. Non-linear time series analysis included calculation of sample entropy (SampEn) values for 512 samples corresponding to the local and global maxima. SampEn is a measure of irregularity and unpredictability of the signal. It is higher for noisy and noise-like and complex signals compared to periodic oscillations. SampEn (m, r, N) is a negative natural logarithm of the conditional probability that two sequences similar for *m* points remain similar at the next point. Algorithm for SampEn computation was published elsewhere (10).

SampEn was calculated for two values of input parameter *r* (tolerance; $r = 0.1$ and $r = 0.2$ times SD of the window). The length of compared sequence ($m = 2$) and the length of analysed window ($N = 512$ samples) was fixed. SampEn for local maximum were denoted as $SampEn_{local}$ *(0.1)* and $SampEn_{local}$ *(0.2)*, SampEn values corresponding to the global maximum were denoted as $SampEn_{global}$ *(0.1)* and $SampEn_{global}$ *(0.2)*. From the spectrum corresponding to the first local maximum and global maximum (512 samples), we computed parameters skewness ($skewness_{local}$, $skewness_{global}$) and kurtosis ($kurtosis_{local}$, $kurtosis_{global}$) of the PSD values distribution (in the frequency band 0–1000 Hz) (Fig. 4). Skewness is a measure of distribution symmetry. A distribution, of a data set is symmetric, if it looks the same to the left and right of the centre point. Kurtosis is a measure of whether the data are peaked or flat relative to a normal distribution. Data sets with higher kurtosis tend to have a distinct peak near the mean, decline rather rapidly, and have heavy tails. Data sets with lower kurtosis tend to have a flat top, near the mean rather than a sharp peak (11).

Classification of the identified sound events

Based on the determined parameters, the identified sound events were classified into cough and non-cough sounds using the classification tree. Trees are directed graphs beginning with

Tab. 1. Sensitivity, specificity, true positive, true negative, false positive and false negative in particular classification tasks, for which classification trees were constructed.

Used classification tree	Sensitivity[%]	Specificity[%]	True positive(n)	True negative(n)	False positive(n)	False negative(n)
cough vs sneezing	95	93	144	28	2	8
cough vs throat clearing	96	85	146	53	9	6
voluntary vs induced cough	96	43	50	40	52	2
cough vs other sounds	86	91	131	1425	129	21

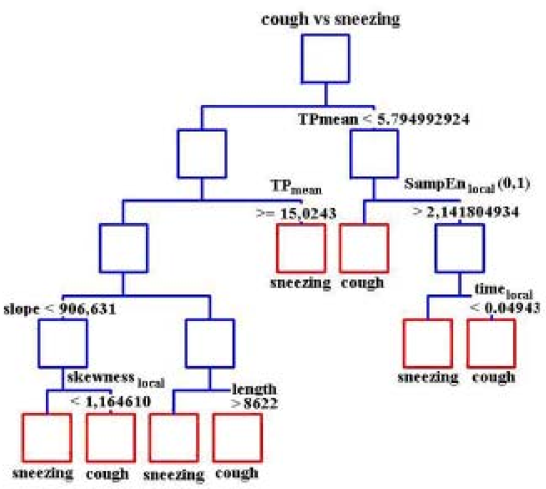

Fig. 5. The classification tree for distinction between cough sounds and sneezing. The top node contains the entire sounds. Each remaining node contains a subset of sounds in the node directly above it. Furthermore, each node contains the sum of the samples in the nodes connected to and directly below it. Each node can be thought of as a cluster of sounds. The sounds are divided to nodes according to the values of evaluated parameters. The most useful variables for distinction between cough sounds and sneezing, which were included in this classification tree are:
- ***length*: the entire length of the entire analyzed sound event,**
- ***slope*: the value of the first local maximum divided by the time of its occurrence,**
- ***TPmean*: the arithmetic mean of the Total Power for the whole sound event,**
- **$time_{local}$: the time of the first local maximum occurrence,**
- **$SampEn_{local}$ (0.1): sample entropy for 512 samples corresponding to the first local maximum for r = 0.1,**
- **$Skewness_{local}$: the value of skewness computed from the frequency spectrum, corresponding to the first local maximum.**

one node and branching into many. Classification trees are used for prediction. They have become popular as alternatives to regression, discriminant analysis, and other procedures based on algebraic model (12). We have constructed four classification trees. The first tree was constructed for distinction between cough and other sounds, the second tree was constructed for distinction between voluntary and induced cough sounds, the third tree was constructed for distinction between cough sounds and voluntary throat clearing and the fourth tree was constructed for distinction between cough sounds and induced sneezing (Fig. 5). The input parameters for tree construction included all assessed variables and the output parameter of the tree was the classification of given sound into cough or non-cough sound events. From the all assessed variables the algorithm for tree construction selected the parameters, which were the most useful for distinction between cough and non-cough sounds.

We have compared the effectiveness of our developed mathematical algorithm to the subjectively classified sound files by two trained observers (visual observation) into particular categories. The system performance was evaluated by calculating the sensitivity, specificity, number of true positives, true negatives, false positives and false negatives.

Statistics

The differences in the sound parameters between cough and non-cough sounds were evaluated using the nonparametric Mann-Whitney U-test. For classification of sound events, the classification tree was constructed (10). Statistical analysis was performed using the statistical package Systat 10, SPSS Inc.

Results

All cough sounds measures were different ($p<0.001$) between cough and sneezing, cough and throat clearing, cough and other sounds. From 20 recordings, we have obtained 1706 sound events. The cough sounds included voluntary (n=52), spontaneous (n=8) and induced coughs (n=92). The non-cough events included voluntary throat (62) and nasopharynx clearing (23), voluntary forced ventilation (19), voluntary snoring (26) induced sneezing (30), laughing (125), swallowing (0), eructation (18), nasal blowing (85), speech (1073) and other sounds (93). Sensitivity, specificity, number of true positives, true negatives, false positives and false negatives for cough vs sneezing, cough vs throat clearing, voluntary vs induced cough, and cough vs other sounds are summarized in the Table 1.

Because the intensity of sounds corresponding to swallowing was very low, these sounds were not detected as sound events.

Discussion

The aim of this study was to verify the effectiveness of our recently developed algorithm for the distinction of the cough sounds from other sounds, which occur in daily human activities. The algorithm was tested against manual cough counts, which is the gold standard for objective evaluation of cough frequency (5). We have found that our algorithm is valuable for the distinction of cough sounds from voluntary throat and nasopharynx clearing, voluntary forced ventilation, voluntary snoring, induced sneezing, laughing, eructation, nasal blowing and speech. On the other hand, many sounds of induced cough were determined as sounds of voluntary cough. It is indicated by low value of specificity. Because the main goal of our study was focused to discriminate the cough sounds from the other sounds produced from the upper airways activities, this negative result is not very important. Our algorithm was able to count the number of coughs, nevertheless it was a single cough or cough epoch.

Until now, published algorithms for distinction between cough and non-cough sound events use methods of time domain and spectral analysis (6–8, 13). Novel non-linear methods (e.g. SampEn) were recently applied to physiological time series analysis (14, 15). We have shown that these measures are also useful for cough sound analysis – all regression trees included at least one of the SampEn measures.

The classification trees are easy to understand and to program specific algorithms. For distinction between cough and non-cough sounds other authors use probabilistic neural network (7), or hidden Markov model (8). Our reached values of sensitivity and specificity are comparable with recently published studies.

We conclude that the classification trees based on the cough sound analysis (including nonlinear measures) are useful for the discrimination of the cough sound events from the other sounds generated from the upper airways in healthy volunteers.

References

1. French CT, Irwin RS, Fletcher KE, Adams TM. Evaluation of a cough-specific quality-of-life questionnaire. Chest 2002; 121: 1123–1131.

2. Subburaj S, Parvez L, Rajagopalan TG. Methods of recording and analysing cough sounds. Pulm Pharmacol 1996; 9: 269–279.

3. Munyard P, Busst C, Logan-Sinclair R, Bush A. A new device for ambulatory cough recording. Pediatr Pulmonol 1994; 18: 178–86.

4. Chang AB, Newman RG, Phelan PD, Robertson CF. A new use for an old Holter monitor: An ambulatory cough meter. Eur Respir J 1997; 10: 1637–1639.

5. Smith JA, Earis JE, Woodcock AA. Establishing a gold standard for manual cough counting: video versus digital audio recordings. Cough 2006; 2:6 doi:10.1186/1745-9974-2-6, http://www.coughjournal.com/content/2/1/6.

6. Coyle MA, Keenan DB, Henderson LS, et al. Evaluation of an ambulatory system for the quantification of cough frequency in patients with chronic obstructive pulmonary disease. Cough 2005; 1:3 doi:10.1186/1745-9974-1-3, http://www.coughjournal.com/content/1/1/3.

7. Barry SJ, Dane AD, Morice AH, Walmsley AD. The automatic recognition and counting of cough. Cough 2006; 2:8 doi:10.1186/1745-9974-2-8, http://www.coughjournal.com/content/ 2/1/8.

8. Matos S, Birring SS, Pavord ID, Evans DH. An automated system for 24-h monitoring of cough frequency: the Leicester cough monitor. IEEE Trans Biomed Eng 2007; 54: 1472–1479.

9. Martinek J, Tatar M, Javorka M. Distinction between voluntary cough sound and speech in volunteers by spectral and complexity analysis. J Physiol Pharmacol 2008; 59.

10. Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. Am J Physiol 2000; 278: H2039-H2049.

11. NIST/SEMATECH e-Handbook of Statistical Methods, http://www.itl.nist.gov/div898/handbook/eda/section3/eda35b.htm.

12. Systat 10 Statistics, USA 2000.

13. Murata A, Ohota N, Shibuya A, Ono H, Kudoh S. New Non-invasive Automatic Cough Counting Program Based on 6 Types of Classified Cough Sounds. Intern Med 2006; 45 (6): 391–397.

14. http://physionet.incor.usp.br/citations.shtml.

15. Aboofazeli M, Moussavi ZK. Comparison of recurrence plot features of swallowing1 and breath sounds. Chaos, Solitons & Fractal (2006), doi: 10.1016/j.chaos.2006.09.026.

Received June 26, 2009.
Accepted October 27, 2010.