

후두 발적에 대한 컴퓨터 평가 시스템의 신뢰도 연구

울산대학교 의과대학 서울아산병원 이비인후과학교실

문병재 · 남순열 · 김상윤 · 최승호

= Abstract =

Reliability of Computerized Measurement of Laryngeal Erythema

Byoung-Jae Moon, MD, Soon Yuhl Nam, MD, Sang Yoon Kim, MD and Seung-Ho Choi, MD

Department of Otolaryngology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea

Background and Objectives : While considerable progress has been made in enhancing the quality of laryngoscopy and image processing, the evaluation of laryngeal erythema is still based on the clinician's judgement. The purpose of this study is to quantitatively measure the degree of erythema and to examine the relationship with clinical grading.

Materials and Methods : Color images of larynx from 100 subjects were captured from video-documented examinations of laryngoscopy. The amount of erythema within the digitized larynx image was quantified using the software developed and was compared with a grading system (0 to 3 scale) based on visual inspection by 4 experienced clinicians. The results were compared by deriving Kappa, Kendall and Spearman statistic.

Results : There was high intra-observer ($R=0.402-0.755$) and inter-observer correlation ($R=0.789$). Among parameters, the red composite value had most remarkable agreement with clinical grading ($R=0.827$).

Conclusion : The results suggest that the computer based analysis of laryngeal erythema can provide quantitative data on degree of erythema and the basis for further development of an expert system.

KEY WORDS : Laryngeal erythema · Color analysis.

서 론

발적은 해당 부위 혈류의 양이 증가했음을 의미하며 후두에서는 급성 및 만성 후두염, 인후두 역류 등의 질환에서 흔히 나타나는 소견이다. 후두의 발적 유무 및 정도의 판정은 굴곡성 후두경, 전자 내시경 등 첨단화된 장비를 사용하는 현재에 이르기까지 전적으로 검사자의 시각에 의존하고 있다. 전문가의 시각에 의한 판단은 객관화, 계량화하기는 어렵다는 단점은 있지만 현재 시점에서 최선의 방법이며, 인후두 역류 환자 대상 연구에서 후두 발적에 관해 검사자 간 혹은 검사자내 재현성이 높음이 보고된 바 있다.¹⁾ 근래에 컴퓨터가 발달하면서 심전도, 혈액세포분석 등 많은 분

야에서 의사의 판독을 보조하여 검사기계에서 자동으로 판독이 이루어지고 있으며 이러한 경향은 점차 확산될 것으로 예상된다. 후두 분야에서도 stroboscopy 영상 분석을 통하여 asymmetric index 등의 척도를 산출하거나,²⁾ 성대의 색상 짜임새를 분석하여 표면의 불규칙성을 계량화하는 등³⁾ 후두 expert system 개발을 위한 연구가 계속 발표되고 있다.

저자들이 알기로 영상 분석을 통하여 후두의 발적을 평가한 것은 Hansen 등의 연구가 유일하며, 그 연구에서는 인후두역류 환자에서 정상인에 비해 발적의 정도가 높으며 역류 치료를 함에 따라 발적이 개선됨을 계량적으로 보여 주었다.⁴⁾

CCD 카메라로부터의 NTSC 출력을 컴퓨터에서 디지털화하면 픽셀 단위로 구성된 컬러 영상을 얻을 수 있다. 24 비트 컬러 영상의 픽셀은 적색(R), 녹색(G), 청색(B)이 각각 8비트 정보량(0-255)을 가지며 이들 삼원색의 조합에 의해 다양한 색상으로 표현된다. 반대로 픽셀을 분석하면 R, G, B 각각의 성분을 얻을 수 있으므로 관심 영역의 색

논문접수일 : 2005년 5월 4일

심사완료일 : 2005년 5월 31일

책임저자 : 최승호, 138-736 서울 송파구 풍납동 388-1

울산대학교 의과대학 서울아산병원 이비인후과학교실

전화 : (02) 3010-3710 · 전송 : (02) 489-2773

E-mail : shchoi@amc.seoul.kr

상에 대한 정보를 계량화 하는 것이 가능하다.

저자들은 첫째, 후두 영상을 보고 전문의가 내리는 판단과 기계적으로 측정된 발적의 척도간에 얼마나 상관관계가 있으며, 둘째 가장 적합한 발적의 척도를 찾아내고, 셋째 후두 발적에 관한 전문의 판단의 재현성을 알아보려고 연구를 시행하였다.

대상 및 방법

1. 대 상

서울아산병원 이비인후과 외래에서 후두경 검사를 시행하여 동영상으로 저장된 기록 중 100예를 선정하였다. 동영상은 Divx 4.12 코덱, 6000kilobit/sec 용량으로 압축된 것이었으며 성대의 노출이 충분하지 않거나 성대 표면이 균일하지 않은 경우는 제외하였다. 동영상으로부터 성대와 피열부가 잘 보이는 시점에 정지영상을 640×480pixel 크기, 24bit 컬러의 비트맵 형식으로 캡처하였다.

2. 컴퓨터에 의한 발적 척도의 계산

연구를 위해 특별히 제작된 프로그램에 의해 Table 1과 같은 척도들을 산출하였다(Fig. 1). 척도 중 Relative redness와 Red composite value는 각각 후두⁴⁾와 결막⁵⁾의 발적에 대해 연구한 문헌에서 참고하였다. 정지영상에서 일측 성대 부위 또는 일측 피열부의 영역을 선택하면 선택 영역의 픽셀 각각에 대하여 RGB 값을 획득하고 RGB 값으

Table 1. Numerical parameters calculated from RGB values for the computerized measurement of the laryngeal erythema

Numerical parameters	Function
Relative redness	$R / (R+G+B)$
Red composite value	$(2R-G-B) / 2(R+G+B)$
Brightness	$0.30R + 0.59G + 0.11B$

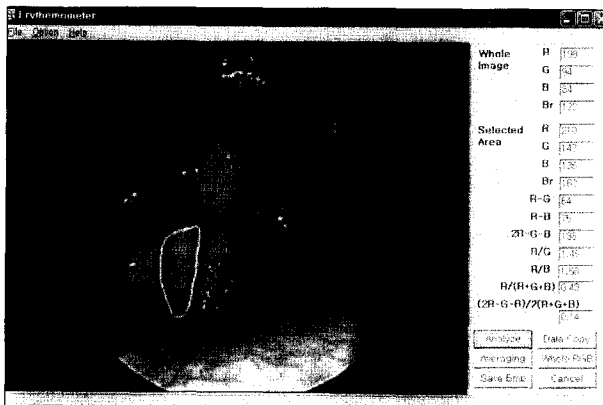


Fig. 1. After the manual selection of the area of interest, numerical parameters are calculated at each pixel and averaged for the selected area.

로부터 척도들을 산출한 후 선택 영역에서의 평균을 구하도록 하였다.

3. 주관적 평가

대상이 된 100예의 후두경 정지영상을 이비인후과 전문의 4명이 1주일 간격으로 2회씩 평가하였다. 검사자에게 대상 환자의 정보는 제공되지 않았으며 두번째 평가 시에는 정지영상의 순서를 바꾸었다. 정지영상은 LCD 모니터 (SyncMaster 175N, Samsung)를 통해 제공되었다. 발적의 정도는 none(0), mild(1), moderate(2), severe(3)의 4단계로 평가하였다.

4. 통계 분석

주관적 검사의 검사자내 및 검사자간 재현성을 Kappa/Kendall 법을 이용하여 분석하였다.정량화된 컴퓨터 측정값과 순위 척도인 주관적 평가 값의 비교를 위하여 주관적 평가값의 평균값을 산출하였고, 상관관계는 Spearman 이변량 상관관계 분석법을 이용하였다. R 값과 주관적 검사치의 상관관계는 B와 G의 영향을 배제하기 위하여 편상관계수법을 이용하여 분석하였다. 모든 통계 처리는 SPSS 10.0.5 프로그램을 사용하였다.

결 과

성대와 피열부 모두 주관적 검사의 검사자 내 그리고 검사자간 재현성은 매우 높아 컴퓨터 측정값과 비교하기에 무리가 없다는 사실을 확인하였다(Table 2). 컴퓨터에서 산출한 여러 척도들 중에 Relative redness와 Red composite value가 주관적 검사와 비교적 높은 상관관계를 나타냈고 명암도(Brightness)와 R 값은 주관적 검사와 음의 상관관계를 보였다(Table 3). R 값과 주관적 검사의 상관관계는 편상관분석법을 이용하여 G, B 값의 요인을 배제하였을 때 성대에서는 상관계수 0.484(p<0.001), 피열부에서는 상관계수 0.478(p<0.001)로 양의 상관 관계를 보

Table 2. Intra-observer and inter-observer reproducibility of the subjective assessment of the laryngeal erythema

	True vocal cord		Arytenoid	
	Coefficient*	p-value	Coefficient*	p-value
Intra-observer				
Observer 1	0.544	<0.05	0.562	<0.05
Observer 2	0.755	<0.05	0.388	<0.05
Observer 3	0.463	<0.05	0.462	<0.05
Observer 4	0.402	<0.05	0.353	<0.05
Inter-observer				
	0.789	<0.01	0.690	<0.01

* : Kappa coefficient for intra-observer analysis and Keddall coefficient for inter-observer analysis

였다. 주관적 검사와 각 척도의 상관성은 피열부에서보다 성대에서 다소 높은 경향을 보였고, 주관적 검사의 변화량에 대비한 Relative redness와 Red composite value의 변화도는 피열부보다 성대에서 크게 나타났다(Fig. 2).

고 찰

전자의무기록이 점차 보편화되고 환자들은 인터넷이나 신문방송을 통해 자신의 질환에 대해 상당한 정보를 가지고 있어 의사의 자세한 설명과 근거를 요구하는 경향이 커지고 있다. 따라서 외래 진료 시 단순히 의사의 눈으로만 보고 문자로 기술하기 보다는 영상을 녹화하거나 캡처하여 환자에게 보여주고 사진 또는 동영상으로 기록하는 일이 잦으며 이 때 후두경에 직접 눈을 대고 보기보다는 의사가 모니터 영상을 보며 판단을 하는 일이 흔하다.

모니터에 나타나는 발적 상태는 후두의 상태뿐만 아니라 후두경 광원의 색온도 및 밝기, CCD 카메라, 화이트밸런스, 모니터의 특성 등에 의해 영향을 받는다. 분석을 하기 위해서는 모든 조건을 표준화하는 것이 이상적이겠지만 실제

진료 환경에서도 표준화가 되지 못한 상태에서 모니터를 보고 의사가 판단하므로 모니터에 나타난 화상 그대로를 분석하는 것이 보다 현실적이라고 할 수 있다. 본 연구에서는 진료 시 캡처한 영상을 분석하여 발적의 정도를 계량화하였고 의사의 주관적 판단과 상관관계가 높음을 증명하였다. 또한 의사의 주관적 판단은 검사자간 혹은 검사자내 재현성이 높다는 결과를 얻어 주관적 판단 결과를 참 값으로 간주하고 산출한 척도와 비교한 연구의 타당성을 입증하였다.

본 연구 결과, 시각에 의해 발적을 판단할 때는 색상 자체뿐만 아니라 화상의 명도(luminance)에 많이 영향을 받는 것이 밝혀졌는데, 같은 색상에서도 화상이 어두울수록 발적의 정도가 심하다고 판정하는 경향이 있었다. 색상이 RGB의 조합에 의해 결정된다는 것을 생각할 때 R 값이 높을수록 발적의 정도가 심할 것이라고 예상할 수 있지만 예상과 달리 R 값 자체는 발적의 주관적 grade와 역비례관계를 보였다. R 값이 높으면 색상은 붉어지지만 명도가 증가하게 되어 주관적 grade에는 상반된 영향을 주며 G, B 값과의 상대적 관계에 따라 시각적 판단이 결정된다고 생각할 수 있다. 여러 척도들 가운데 Relative redness와 Red composite value가 특히 주관적 판단과 상관성이 높게 나타났다는데 이 두 척도는 G, B에 대한 상대적 R 값과 명도의 개념을 모두 포함하였기 때문으로 생각된다.

성대와 피열부를 나누어 분석하였을 때 주관적 검사의 재현성이 피열부보다 성대에서 더 높은 것으로 나타났는데, 이는 피열부는 정상상태에서도 어느 정도 붉은 색을 가지고 있어 발적의 유무나 grade에 대한 판단이 성대에 비해 부정확한 것으로 생각된다. 또한 피열부에 비해 성대에서

Table 3. Correlation between the subjective assessment and the computerized measurement of the laryngeal erythema

Numerical parameters	True vocal cord		Arytenoid	
	R	p-value	R	p-value
Red	-.684	<.001	-.555	<.001
Green	-.838	<.001	-.820	<.001
Blue	-.830	<.001	-.720	<.001
Relative redness	.678	<.001	.778	<.001
Red composite value	.827	<.001	.764	<.001
Brightness	-.815	<.001	-.754	<.001

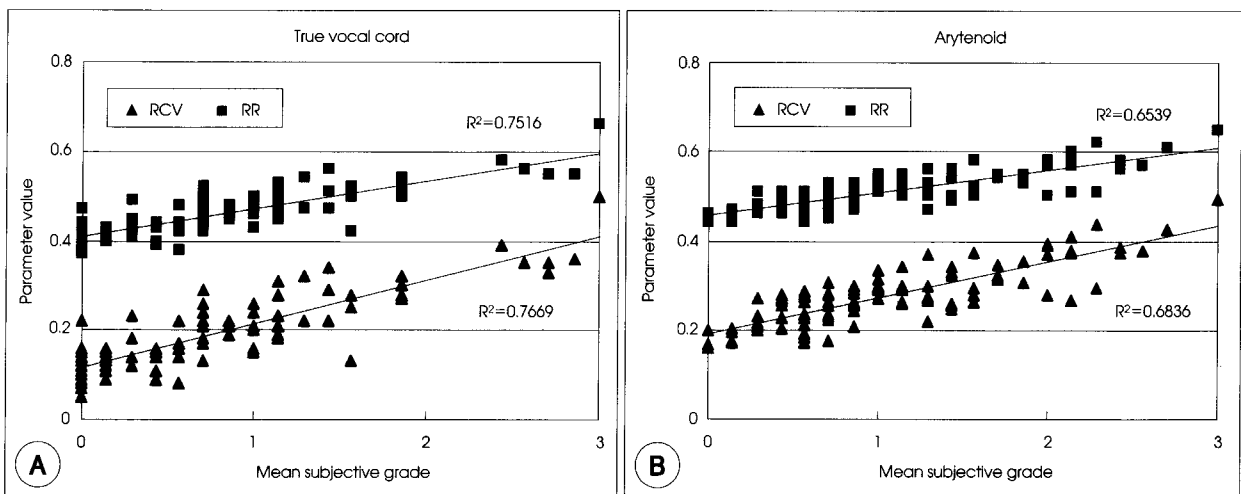


Fig. 2. Comparison of the true vocal cord and the arytenoid area. The relative redness(RR) and the red composite value(RCV) in the true vocal cord(A) are more proportional to the subjective grade than in the arytenoid(B).

주관적 grade변화에 대한 산출 척도의 변화가 큰 것으로 나타나 컴퓨터를 이용한 발적 평가는 피열부보다 성대에서 더 적합하다고 생각되었다.

본 연구는 컴퓨터에 의한 후두 발적 평가 시스템을 개발하고 이 시스템에서 산출한 척도와 주관적 평가와의 상관관계를 증명함으로써 시스템의 타당성을 증명한 것이다. 실제 외래에서 진료한 영상을 이용하였기 때문에 어두운 영상도 있고 밝은 영상도 있으며 때로는 화이트밸런스가 잘 맞지 않은 영상도 있었으나 그럼에도 불구하고 전문의의 시각에 의한 판단과 일치하는 결과를 얻어 컴퓨터에 의해 계산된 척도가 인간의 판단에 근사함을 확인하였다. 후두 질환 진단의 expert system 개발을 위해서는 추후 더욱 적합한 척도를 개발하고, 성대 변연의 윤곽선 추적이나 texture analysis 등의 영상처리 기법에 대한 연구가 필요하리라 생각된다.

결 론

후두의 발적 상태를 평가하는데 있어 전문의에 의한 주

관적 평가는 높은 재현성을 보이며, 영상의 색상 분석에 의해 산출한 척도 중 특히 Relative redness와 Red composite value가 주관적 평가와 높은 상관관계를 보여 가장 적합한 척도로 생각된다. 컴퓨터에 의한 후두 발적 평가 시스템은 적합한 척도를 적용한다면 전문의에 의한 평가와 유사한 결과를 얻을 수 있다.

중심 단어 : 후두 발적 · 색상 분석.

REFERENCES

- 1) Branski RC, Bhattacharyya N, Shapiro J. *The reliability of the assessment of endoscopic laryngeal findings associated with laryngopharyngeal reflux disease. Laryngoscope* 2002;112 (6):1019-24.
- 2) Sung MW, Kim KH, Koh TY, Kwon TY, Mo JH, Choi SH, et al. *Videostrobokymography: a new method for the quantitative analysis of vocal fold vibration. Laryngoscope* 1999;109 (11):1859-63.
- 3) Ilgner JF, Palm C, Schutz AG, Spitzer K, Westhofen M, Lehmann TM. *Colour texture analysis for quantitative laryngoscopy. Acta Otolaryngol* 2003;123 (6):730-4.
- 4) Hanson DG, Jiang J, Chi W. *Quantitative color analysis of laryngeal erythema in chronic posterior laryngitis. J Voice* 1998;12 (1):78-83.
- 5) Fieguth P, Simpson T. *Automated measurement of bulbar redness. Invest Ophthalmol Vis Sci* 2002;43 (2):340-7.