대한응급의학회지 제 13 권 제 3 호
Volume 13, Number 3, September, 2002

중증도 분류도구로서 변형 Emergency Severity Index의 평가자간 신뢰도

연세대학교 의과대학 응급의학교실, 건국대학교 의과대학 응급의학과[1]

박유석 · 조진경 · 윤천재 · 박인철 · 이경룡[1] · 김승호

Inter-rater Reliability of the Modified Emergency Severity Index as a Triage Tool

Yoo Seok Park, M.D., Jin Kyung Cho, M.D., Cheon Jae Yoon, M.D., In Cheol Park, M.D., Kyeong Ryong Lee, M.D.[1], and Seung Ho Kim, M.D.

Purpose: Triage in the emergency departmen (ED) is the preliminary clinical sorting process before full disclosure of patients' problems so that patients with the highest acuity are treated first in the setting of resource constraints. To overcome the inter-rater variability of existing triage tools, the Emergency Severity Index (ESI) was developed and was shown to be both valid and reliable in practice. Because of the disparity in practice patterns and some inappropriate criteria in the original ESI, the authors have modified the ESI and determined its inter-rater reliability.

Methods: We applied the modified ESI to a convenient sample of adults who visited an urban academic ED between July 24, 2001, and August 5, 2001. After completion of a short, 4-hour training course on the modified ESI, an intern and emergency medicine resident pair triaged the patients independently. The inter-rater reliability was measured using a weighted kappa analysis and was categorized as excellent (≧0.8), good (0.60-0.79), or fair (≦0.59).

Results: Five hundred forty-two patients were enrolled. The overall weighted kappa between the intern and the resident was 0.82 (95% CI : 0.78-0.86). Among the 542 patients, 469/542 (87%) pairs agreed exactly, 67 (12%) pairs disagreed by 1 level, and 6 (1%) pairs by 2 levels.

Conclusion: In this study, the modified ESI demonstrated excellent inter-rater reliability when used by residents and interns for our ED patients.

Key Words: Triage, Emergency department, Reliability

Department of Emergency Medicine, Yonsei University College of Medicine, Seoul, Korea
Department of Emergency Medicine, Kon-kuk University College of Medicine, Chungju, Korea[1]

서 론

우리나라의 3차 진료 기관의 응급의료센터에는 환자의 편중과 입원 대기 환자의 적체로 과밀화 현상이 일반화되고 있다[1]. 이로 인해 즉각적인 가료가 필요한 응급환자가 간과되어져 대기실로 보내지고 있는 것이 현실이다[2]. 응급실 과밀화는 의료진의 업무가중뿐 아니라 응급의료 서비스의 질 저하, 진료 지연과 대기시간의 연장으로 이어져 환자 및 보호자의 불만을 증가시키고, 심지어는 폭력 행사까지 나타나고 있다[1]. 따라서 공간, 인력, 장비 등의 제약 속에서 위와 같은 문제의 해결을 위해서는 응급의료센터에 내원하는 환자들에게 중증도 또는 급성도(acuity)에 기준한 선별적이면서도 신속하고 적절한 치료가 이루어져야 할 것이다. 중증도 분류(triage)는 응급의료에서 충분한 진단과 치료를 수행하기 전인 초기단계에서 임상 평가를 통해 환자의 중증도를 분류하고, 이에 따라 한정된 자원(제한된 병상 수, 의료진, 장비)안에서 급성도가 가장 높은 환자에게 가장 먼저 의료를 제공하게 하는 과정이다[3].

중증도 분류에는 3단계 분류가 현재 가장 널리 사용되고 있으며, 이를 좀 더 자세히 분류하는 4단계 분류와 호주, 캐나다 등에서 사용되고 있는 5단계 분류법이 있다[4]. 기존의 3단계 또는 4단계의 중증도 분류도구의 문제점, 특히 비응급 환자를 정확히 분류하지 못하는 점[5,6]을 보완하기 위하여 1999년 Wuerz 등[7]은 5단계의 Emergency Severity Index(ESI)를 개발하였다. ESI는 다른 분류도구와는 달리 증상, 생리지표를 통한 급성도 뿐만 아니라 예상되는 자원의 사용까지를 포함하는 포괄적인 개념의 도구로서 평가자간 높

책임저자: 김 승 호
서울시 서대문구 신촌동 134번지
연세대학교 의과대학 응급의학교실
Tel: 02) 361-5790, 6660, Fax: 02) 392-3715
E-mail: edksh@yumc.yonsei.ac.kr
접수일: 2002년 6월 30일, 게재승인일: 2002년 8월 30일

* 본 논문의 요지는 2001년 대한응급의학회 추계학술대회에서 구연되었음.

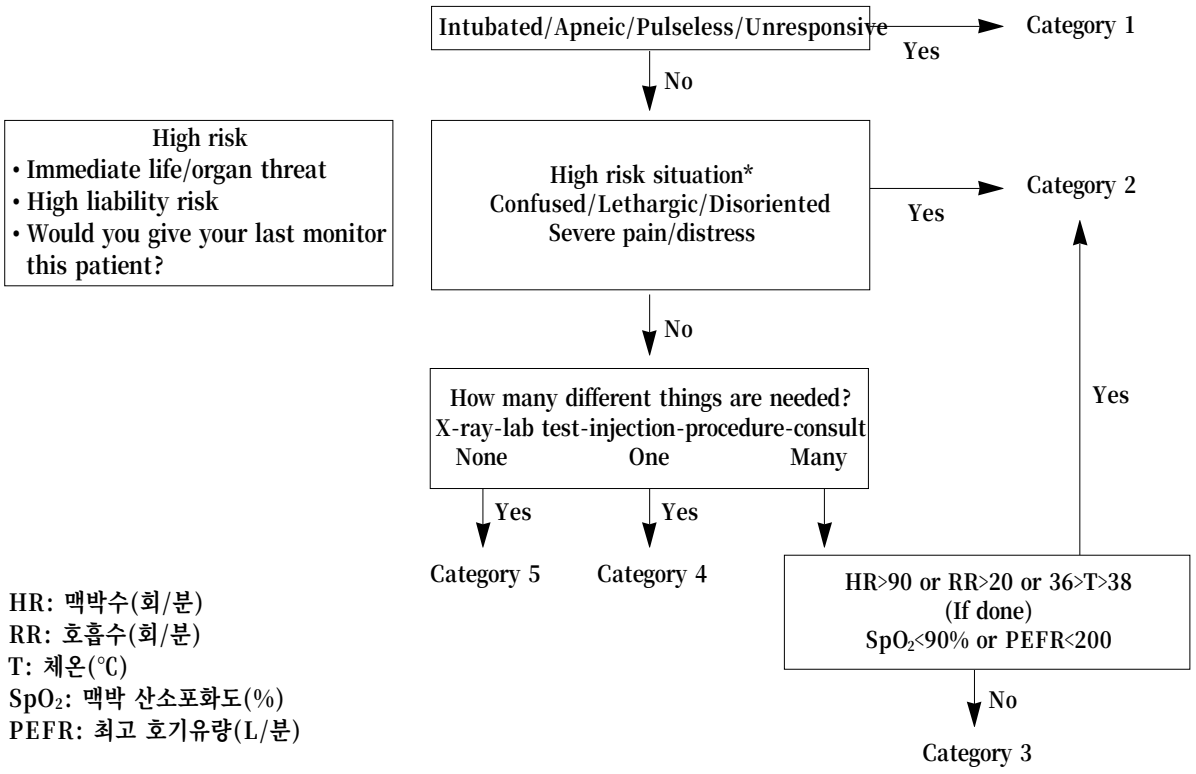

그림 1. 기존의 Emergency Severity Index

은 일치도가 보고 되고 있다[3]. ESI는 알고리듬(algorithm) 형식의 중증도 분류도구로서 환자는 5범주(category)로 분류된다(그림1)[7]. 각각의 범주는 환자의 생체 기능의 안정성, 고통의 정도, 예상되는 자원의 양(intensity), 기대되는 의료진의 반응에 의해 결정된다[3].

하지만, 연구자들이 Wuerz의 ESI를 실제로 적용하여 예비실험을 한 결과, 주관이 많이 관여되는 통증의 정도판단, 생체징후에 응급의료에서 사용하기에는 적당하지 않은 전신성 염증성 반응 증후군(SIRS: systemic inflammatory response syndrome)의 기준사용, 그리고 외국과는 달리, 방사선 촬영과 임상병리검사를 선별적이 아닌 대부분의 환자에서 시행하는 진료형태 등의 요인으로 인해 대부분 환자의 중증도가 높게 분류되는 경향을 볼 수 있었다. 이에 연구자들은 해당 병원의 응급진료 현실에 맞게 ESI를 변형시키고, 변형된 ESI의 임상 적용가능성을 평가하기 위해 중증도 분류도구로서 필요한 요소 중 하나인 평가자간의 신뢰도(reliability)를 알아보고자 하였다.

대상과 방법

연간 접수환자 약 30,000명인 서울 소재 대학부속병원 응급의료센터에 2001년 7월 24일부터 8월 5일까지 내원한 15세 이상 성인 환자를 대상으로 모두 8명의 응급의학과 인턴과 4명의 전공의가 변형된 ESI를 이용한 중증도 분류를 시행하였다. 변형된 ESI는 original ESI와 몇 가지 차이가 있는데 "심한 통증"을 뇌출혈이 의심되는 두통, 급성심근경색이나 협심증이 의심되는 심인성 흉통, 그 외 생명을 위협할 만한 통증(예를 들어 대동맥 박리증에 의한 복통 등)으로 제한하였다. "생체 징후의 이상"은 수축기혈압이 90 mmHg 미만, 심한 호흡곤란, 출혈이 의심되는 환자의 맥박이 분당 100회 이상인 경우로 정의하였다. 그밖에 잘 사용되지 않는 최고 호기유량이나 산소포화도는 제외시켰다. 그리고 자원사용에서 기본적인 방사선 촬영과 임상병리검사를 하나로 묶었고, 외상환자에게는 방사선 촬영과 술기를 하나로 묶어서 크게는 진단, 술기, 협진의 3가지로 구분하였다 (그림 2).

연구 참여에 동의한 인턴과 전공의에게는 변형된 ESI 도구에 대한 이해와 실제적인 적용에 대한 4시간의 교육을 실시하였다. 교육은 중증도 분류와 변형된 ESI에 대한 교육, 가상 시나리오 및 실제 환자에 대한 적용과정으로 진행하였다. 연구에서는 임의로 선택된 인턴과 전공의 각 1명이 동일 환자에 대하여 변형된 ESI를 이용하여 각각 중증도 분류를 시행하되 서로의 결과를 모르도록 하였고, 그 결과를 가지고 5*5의 평가자간의 일치도 매트릭스를 작성하였다.

평가자간의 일치도는 SAS version 6.12 (SAS institute, Cary, NC)의 weighted kappa analysis를 통해 측정하였고 95%의 신뢰구간을 사용하였다. Weighted kappa

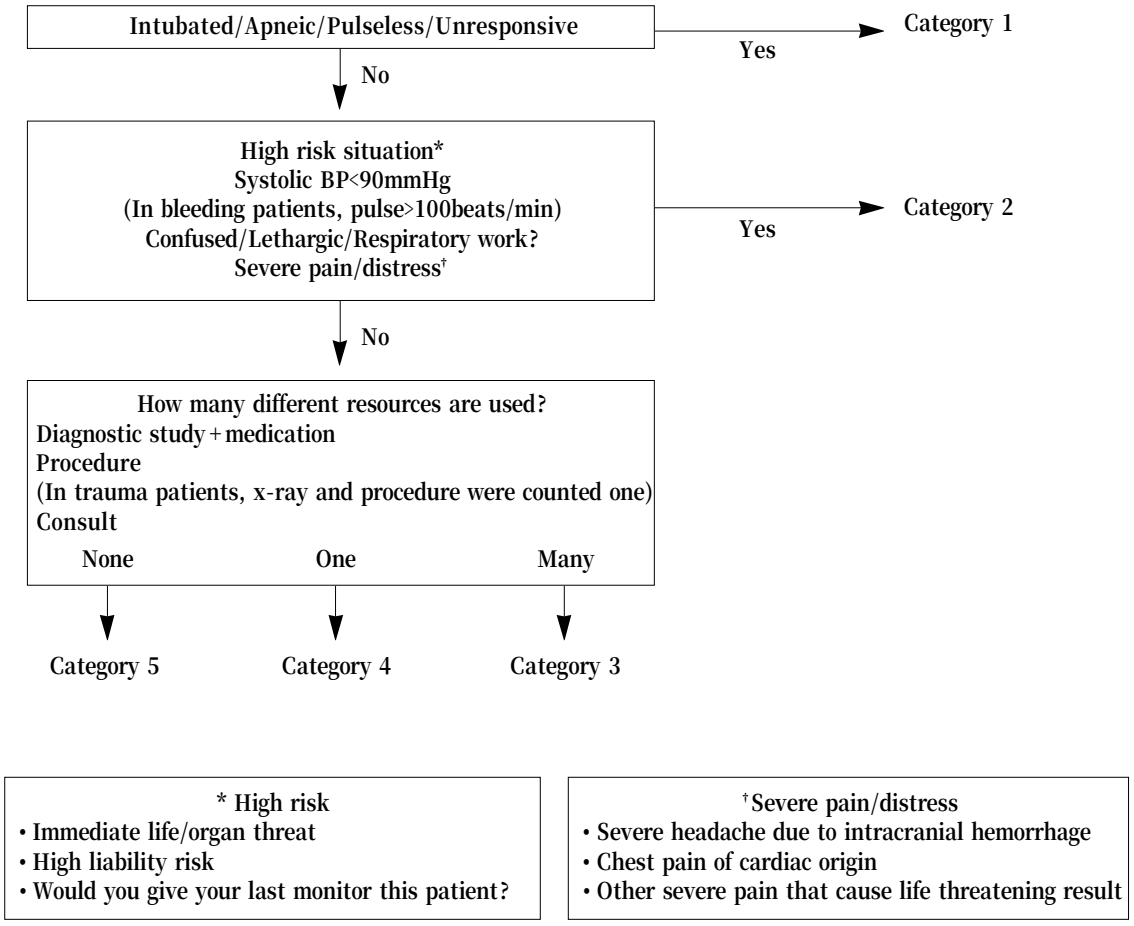

그림 2. 변형된 Emergency Severity Index

analysis에서는 Cohen's simple kappa 분석과 유사하게 0 (완전 불일치)에서 1 (완전 일치) 까지의 값이 나오나, 쌍이 되는 등급의 일치하지 않는 정도를 교정하여 등급간의 차이가 클수록 weighted kappa 값이 떨어지게 된다. 측정된 weighted kappa 값은 0.80 이상은 매우 우수 (excellent), 0.60에서 0.79사이는 우수 (good), 0.59이하는 보통 (fair)으로 평가된다.

결 과

모두 542명의 환자가 대상이 되었고 이중 남자는 312명, 여자는 230명이었고, 평균 연령은 50.1±17.9세 였다. 위장관계, 신경계, 심혈관계, 호흡기계 등의 순으로 내원 증상 및 질환이 의심되었다 (표 1). 평가자간의 weighted kappa 값은 0.82 (95% 신뢰구간 : 0.78~0.86)이었다 (표 2).

542명의 환자 중 34명 (6%)에 대하여 전공의가 인턴에 비해 보다 중증의 범주로 분류하였고, 37명 (7%)에서 인턴이 보다 중증의 범주로 평가하였다. 전체적으로 469명 (87%)의 환자의 중증도 분류가 일치하였고, 67명 (12%)에서 한 수준, 6명 (1%)에서 두 수준의 불일치가 있었다. 한 수준의 차이가 나는 67명에 대하여 인턴 또는 전공의가 보다 중증의 범주로 평가하는 데에는 어떤 연관성은 없었다. 두 수준의 불일치가 나타난 환자를 살펴보면, 요측부 통증을 주소로 내원한 요로결석, 두통을 주소로 내원한 긴장성 두통, 심계항진을 주소로 내원한 동성 빈맥 환자를 인턴은 2범주, 전공의는 4범주로 분류하였다. 또, 두통을 주소로 내원한 지주막하 출혈환자와 상복부 통증을 주소로 내원한 급성심근경색 환자를 인턴은 4 범주, 전공의는 2범주로 분류하였고, 생체징후가 안정한 상부 위장관 출혈환자를 인턴은 1 범주, 전공의는3범주로 분류하였다.

고 찰

중증도 분류란 불어 'trier'에서 유래된 것으로 영어로는 'to sort or to choose'란 말로 풀이된다. 역사적으로 이는 1800년대 군대에서 시작되었으며, 우선 순위에 따라 부상자를 분류, 수송하기 위하여 사용되었다[7]. 미국에서 응급의료센터에 중증도 분류의 개념이 처음 도입된 것은 1950년대

후반에서 1960년대 초반이다. 당시 응급의료센터에서는 환자수의 급증이 일어났으며, 상당수의 환자는 응급처치가 필요하지 않는 상태에서 내원하였다. 또, 의료를 중환자나 외상 환자에게만 필요한 것이 아닌 모든 사람이 이용할 수 있는 사회적 자원으로 생각하게 되었고, 많은 의사가 개원을 하고, 의료가 전문화되면서 응급의료센터는 일차적인 의료제공의 창구가 되었다[8]. 이에 환자의 우선 순위를 결정하고, 응급의료센터의 흐름을 원활히 유지하기 위하여 중증도 분류 원칙을 적용하는 것은 필수적이었다. 따라서, 진단과 치료가 시행되기 전에 대기할 수 있는 환자로부터 즉각적인 의학적 처치가 필요한 환자를 구별할 수 있는 빠르고 정확한 도구가 필요하였다. 현재 미국에서는 96%의 응급의료센터에서 몇 가지 종류의 중증도 분류 방법이 사용되고 있다[4]. 이중 61%의 센터에서는 응급(emergent), 긴급(urgent), 비응급(nonurgent)으로 분류하는 3단계 중증도 분류 방법을 사용하고 있다. 하지만 이는 재현가능성(reproducibility), 즉 신뢰도가 떨어지고[9,10], 임상 결과에 대한 경험적 타당성이 부족하여 이를 통해 응급의료센터의 자원의 소모를 예측할 수 없다[5]. 그러나, 이와는 달리 호주, 캐나다, 영국에서 사용되고 있는 5단계 분류법들은 평가자간의 높은 일치도와 타당성을 보고하고 있다[11~13]. 그리고, Wuerz 등[3,14,15]은 급성도 뿐만 아니라, 자원의 사용 등 포괄적 개념의 5단계 분류방법인 ESI를 개발하여 평가자간의 높은 일치도와 입원율, 응급실 재원시간, 비용 등 임상 결과와의 연관성 및 6개월간의 생존율과의 연관성에 대하여 발표하였다. 신뢰도와 타당성이 높은 중증도 분류 도구는 입원율, 재원기간, 사망률을 예측할 수 있으며[13,16], 응급의료센터의 자원 소모의 예측 및 예산 책정 등 응급의료센터의 운영계획을 세우는데 도움을 줄 수 있으며, 정부 기금 책정, case mix (각 범주에 해당하는 환자수) 자료를 만드는 데에도 사용될 수 있다[17]. 연구자들은 예비 실험에서 기존의 ESI를 적용해 보았는데 104명의 대상 환자 중 32명 (30.8%)의 환자가 2범주로, 60명 (57.7%)의 환자가 3범주로 분류되어 대부분의 환자 (88.5%)가 2 또는 3 범주로 분류되는 문제점이 발견되었다. 연구자들은 기존의 ESI에서 통증의 판단에 평가자의 주관이 많이 관여되고, 생체징후의 판단에 응급실에서 사용하기에는 적당하지 않은 전신성 염증성 반응 증후군의 기준이 사용되고 있으며, 외국과는 다르게 방사선 촬영과 임상병리검사를 선별적이 아닌 대부분의 환자에서 시행하는 진료형태 등의 요인 때문에 대부분 환자의 중증도가 높게 분류되었다고 판단하였다. 이에 이러한 ESI의 단점을 보완하기 위해 심한 통증을 몇 가지로 제한하고, 생체 징후를 기존 ESI보다 더 높은 기준으로 정하고, 자원의 사용을 크게 진단, 술기, 협진으로 구분하는 등 ESI를 변형하여 국내 병원에 보다 적합한 형태로 변형하였다.

중증도 분류도구로서 높은 신뢰도를 확립하는 것은 꼭 필요한 첫 번째 단계이다. 변형 ESI는 4시간의 비교적 짧은 기간의 교육과정에도 불구하고, 인턴과 전공의간의 매우 우수한 일치도(weighted kappa = 0.82 (95% CI: 0.78 ~ 0.86))를 보이고 있으며 이는 Wuerz 등의 연구와 큰 차이가 없었다. 이는 이 분류도구가 기존의 다른 분류 방법과는 달리, 알

표 1. 대상 환자 및 의심되는 질환 분포(n = 542)

성별	
남	312(57.6%)
여	230(42.4%)
연령(평균±표준편차)	50.1±17.9
진단	
위장관계 질환	102(18.8%)
신경계 질환	67(12.4%)
심혈관계 질환	62(11.4%)
호흡기계 질환	55(10.1%)
신장/비뇨기계 질환	55(10.1%)
혈액/종양계 질환	27(5.0%)
안과계 질환	20(3.7%)
피부계 질환	14(2.6%)
정신신경계 질환	13(2.4%)
산부인과계 질환	11(2.0%)
구강/악안면계 질환	9(1.7%)
내분비계 질환	8(1.5%)
근골격계(비외상성) 질환	6(1.1%)
경증 외상	38(7.0%)
중증 외상	13(2.4%)
기타	42(7.7%)

표 2. 변형 Emergency Severity Index를 사용한 중증도 분류의 평가자간 일치도(n = 542)

레지던트	인턴					
	1 범주	2 범주	3 범주	4 범주	5 범주	합계
1 범주	10	1	0	0	0	11
2 범주	3	84	20	2	0	109
3 범주	1	7	252	14	0	274
4 범주	0	3	19	110	2	134
5 범주	0	0	0	1	13	14
합계	14	95	291	127	15	542

Weighted kappa = 0.82(95% CI : 0.78-0.86)

고리듬 형식으로 되어있어 분류가 용이하고, 기존의 ESI에서 일치도가 떨어지는 긴급에 해당하는 3, 4범주를 예상되는 자원 사용을 기준으로 분류하고, 또 이 기준을 예비실험을 통하여 해당병원의 현실에 맞게 변형시켰기 때문이라고 생각된다. 그리고 연구에 포함된 분류자들이 모두 의사이기 때문에 일치도가 더 높았을 가능성도 있다.

본 연구의 제한점으로 첫째, 이번 연구는 서울 소재 1개의 대학병원에 실시된 것으로 주로 내과 계열의 환자가 많은 부분을 차지하고, 만성 질환 환자가 많으며, 입원 대기가 많은 특성을 지닌 응급의료센터의 결과로서, 여러 가지 다른 환경을 가지고 있는 다수의 응급의료센터에서 후속 연구를 시행하여 비교해 보는 것이 필요하다. 둘째, 이번 연구에 포함된 분류자들은 모두 의사로, 다른 나라에서 주로 중증도 분류를 담당하는 간호사와는 차이가 있으므로 향후 의사-간호사간의 일치도에 관한 연구가 필요할 것으로 생각된다. 셋째, 이번 연구에서는 타당성에 대한 연구를 시행하지 않았는데 이는 아직 중증도의 타당성을 정확히 반영하는 지표가 없기 때문이다. Wuerz 등[3,14,15]은 타당성 검증을 위해 입원율, 응급실 재원시간, 비용 등의 임상 결과와의 연관성 및 6개월간의 생존율과의 연관성에 대하여 발표하기도 하였지만 분류에서 응급검사가 필요하다고 생각되어 높은 범주에 포함되었던 환자가 검사결과 이상 없다고 판단되면 입원할 필요가 없고(예를 들어 뇌출혈의 가능성을 시사하는 두통 환자의 진단, 저혈당에 의한 의식변화, 천식에 의한 호흡곤란), 반대로 중증도는 낮게 분류된 환자가 여러 가지 다른 사정으로 입원하게되는 경우도 있으며 연구를 시행한 병원과 같이 병실 부족으로 입원을 못하는 경우도 있기 때문이다. 재원 기간 역시 여러 가지 다른 변수, 즉 대기시간, 병실의 가동 가능한 침상 수, 검사에 걸리는 시간, 협진에 걸리는 시간, 중환자실 부재 등 환자의 중증도와는 관련 없는 다른 요인에 많은 영향을 받는다. 앞으로 타당성을 반영하는 지표 및 이를 토대로 변형된 ESI의 타당성에 관한 연구가 필요하다. 넷째, 본 연구는 15세 이상 성인을 대상으로 이루어 졌으며, 앞으로 소아와 성인을 포함한 모든 내원환자에게 확대 실시가 필요할 것이다.

결 론

변형된 ESI는 매우 우수한 평가자간 일치도를 보여 응급의료센터에서 중증도 분류도구로서 높은 신뢰도를 가지고 사용될 수 있을 것으로 사료된다.

참고문헌

1. 박인철, 이경룡, 구홍두, 김승환, 장석준, 심호식: 응급의료센터 과밀화 해소를 위한 환자 전원. *대한응급의학회지* 6: 403-10, 1995.
2. Gerven RV, Delooz H, Sermeus W: Systemic triage in the emergency department using the Australian National Triage Scale: a pilot project. *Eur J Emerg Med* 8: 3-7, 2001.
3. Wuerz RC, Milne LW, Eitel DR, Travers D, Gilboy N: Reliability and validity of a new five-level triage instrument. *Acad Emerg Med* 7: 236-42, 2000.
4. Zimmermann PG: The case for a universal, valid, reliable 5-tier triage acuity scale for US emergency department. *J Emerg Nurs* 27: 246-54, 2001.
5. Wuerz RC, Fernandes CM, Alarcon J: Inconsistency of emergency department triage. emergency department operations research working group: *Ann Emerg Med* 32: 431-5, 1998.
6. Gill JM, Reese CL, Diamond JJ: Disagreement among health care professionals about the urgent care needs of emergency department patients: *Ann Emerg Med* 28: 474-9, 1996.
7. Gilboy N, Travers D, Wuerz RC: Re-evaluating triage in the new millennium: a comprehensive look at the need for standardization and quality. *J Emerg Nurs* 25: 468-73, 1999.
8. Weinerman ER, Rutzen R, Pearson DA: Effects of medical triage in hospital emergency services. *Public Health Rep* 80: 389-99, 1965.
9. Travers D, Anna W, Michael BJ, Deborah F: Comparison of 3-level and 5-level Triage Acuity Systems. *J Emerg Nurs* 26: 411, 2000.
10. Fernandes CM, Wuerz RC, Clark S, Djurdjev O: How reliable is emergency department triage? *Ann Emerg Med* 34: 141-7, 1999.
11. Hollis G, Sprivulis P: Reliability of the National Triage Scale with changes in emergency department acuity level. *Emerg Med* 8: 231-4, 1996.
12. Robert B, James D, Laurie J, Serge B, Stephen W: Reliability of the Canadian emergency department triage and acuity scale: interrater agreement. *Ann Emerg Med* 34: 155-9, 1999.
13. Cook MW, Jinks S: Does the Manchester triage system detect the critically ill? *J Accid Emerg Med* 16: 179-81, 1999.
14. Wuerz RC, Travers D, Gilboy N, Eitel DR, Rosenau A, Yazhari R: Implementation and refinement of the emergency severity index. *Acad Emerg Med* 8: 170-6, 2001.
15. Wuerz RC, Eitel DR, Gilboy N, et al.: Emergency severity index triage category is associated with six-month survival. *Acad Emerg Med* 8: 61-4, 2001.
16. Cameron PA, Bradt DA, Ashby R: Emergency medicine in Australia. *Ann Emerg Med* 28:342-6, 1997.
17. Cleary MI, Ashby RH, Jelinek GA, Lagaida R: The future of case mix in emergency medicine and ambulatory care. *Med J Aus* 161 (Suppl): S30-3, 1994.