Reliability and Validity of the Toowoomba Adult Trauma Triage Tool

Anthony Wollaston RN, BN, GradCertMid, GradDipHSc
Clinical Nurse
Toowoomba Health Service Emergency Department, Toowoomba, Queensland, Australia 4350

Mr. Paul Fahey, BSc, MMedStat
Lecturer, Department of Mathematics and Computing
University of Southern Queensland, Toowoomba, Queensland, Australia, 4350

Ms Michelle McKay RN, BHSc(Nursing), GradCertHMan
Nursing Unit Manager
Toowoomba Health Service Emergency Department, Toowoomba, Queensland, Australia, 4350

Prof Desley Hegney BA(Hons), DNE, PhD, FRCNA
Professor of Rural Nursing
University of Southern Queensland, Toowoomba, Queensland, Australia, 4350

Dr Peter Miller MBBS, FACEM, FRACGP
Director of Emergency Department
Toowoomba Health Service Emergency Department, Toowoomba, Queensland, Australia, 4350

Mr James Wollaston BChEng, MBA
Principal, Wollaston Consulting
11 Chandler Grove, Burwood, Victoria, Australia, 3125

Reprint requests to Anthony Wollaston

Keywords: Triage, Decision Support Systems Clinical, Decision Support Computerised

Abstract

Introduction
In Australia a nationally adopted five tiered triage scale called the Australasian Triage Scale (ATS) is used for the purpose of differentiating patient acuity levels for all patients that present to an Emergency Department (ED). The scale was formulated with the aim of promoting a standardized approach to triage. Numerous studies now suggest that the ATS has not been successful in achieving this intention. The Toowoomba Adult Trauma Triage Tool (TATTT) seeks to address this deficiency by providing a reproducible, reliable and valid method of triage categorisation, albeit in a select group of patients.

Method
Ten triage nurses from Toowoomba Health Service (THS) and five from Princess Alexandra Hospital (PAH) were enrolled in the study. Each participant rated 5 written simulations before receiving training in the TATTT. Each participant then rated nine written scenarios, five simulated scenarios on video and one computer simulated scenario with the TATTT. Parallel triaging of actual adult injury based presentations to THS ED was also undertaken.

Results
The overall percentage agreement for all simulations for which data was available was 87%. The overall kappa statistic was 0.82 with a 95% confidence interval from 0.76 to 0.88. The overall percentage agreement between the triage category assigned using existing triage practices and the triage category assigned by the TATTT was approximately 47% with an estimated kappa of $\hat{\kappa} = 0.19$ (95% confidence interval [-0.02, 0.40]).

Conclusion

Preliminary evidence thus far suggests that the TATTT provides systematically different results compared to current triage practices utilising the ATS and has greater reliability than current triage practices.

INTRODUCTION

Triage in an Emergency Department (ED) context constitutes the formal process of immediate assessment and categorisation of all patients who present seeking treatment (Commonwealth Department of Health and Family Services and the Australasian College of Emergency Medicine, 1997). In Australia a five tiered scale for the purpose of differentiating patient acuity levels, called the National Triage Scale (NTS), has been widely adopted to facilitate this process. The NTS was formulated in 1993 by the Australasian College for Emergency Medicine (ACEM) with the aim of promoting a standardized approach to triage in Australian EDs (Considine, Ung, Thomas, 2001). In 2000 the NTS was revised and renamed the Australasian Triage Scale (ATS).

Consequently, since the introduction of the NTS, a lack of standardisation in its application still exists (Considine, Ung, Thomas, 2000).

If the ATS was effective in providing a standardised approach to triage, a patient should be able to present to any triage nurse, in any ED, at any time of the day with a specific problem and be allocated the same triage category each time (Considine, Ung, Thomas, 2000). A number of studies have identified experimentally that this is unlikely to occur.

Doherty (1996) conducted a study into the uniformity of triage between medical and nursing staff and between four different hospitals utilising the NTS. The study comprised 12 written patient scenarios. He found that no one scenario was triaged the same by all participants and 10 of the 12 scenarios received triage scores encompassing three or more categories (Doherty, 1996).

Dilley and Standen (1998) assessed the level of uniformity among Victorian public hospital triage nurses in utilizing the NTS. The study comprised 20 written patient scenarios administered to 188 nurses from 14 different Victorian hospital EDs. The authors found that no one patient scenario was triaged to the same category by all 188 triage nurses and 75% of scenarios were triaged to four different triage categories (Dilley and Standen, 1998).

Considine, Ung and Thomas (2000) examined triage nurses level of agreement in the allocation of triage categories using the NTS. They surveyed 31 triage nurses from two institutions with 10 written scenarios. The authors found that no one scenario was triaged the same by all participants. Six of the scenarios were triaged to two triage categories and four of the scenarios were triaged to three triage categories (Considine, Ung, Thomas, 2000).

The implications of this variability in triage categorisation as identified in these studies are numerous. If triage nurses fail to categorise patients appropriately then medical attention may be unacceptably delayed or unnecessarily expedited (Goodacre, Gillet, Harris, et al 1999). As a consequence there will be inequity in patient access to ED services, potential risks to patient safety, inefficiencies in the management of patient workload and difficulty in making meaningful comparisons between institutions based on

triage categorisation data (Considine, Ung, Thomas, 2000; Goodacre, Gillet, Harris, et al 1999; Doherty, 1996; Monash Institute of Health Services Research, 2001).

Algorithms can be used to provide decision support in the assessment of clinical urgency. The benefit of such an approach lies in the direction and structure they provide in guiding the decision for the triage nurse (Gerdtz and Bucknall, 1999). This concurs with the thoughts of other authors who suggest nurses need to use standard guidelines when assessing patients so that all patients are assessed similarly and according to standards of practice (Stephens, Pokorny, Bowman, 1997). This need for a systematic method of assessment which is easy to understand and quick to use has been widely recognised (Williams, 1992).

The TATTT seeks to provide a reproducible, reliable and valid method of triage assessment and categorisation, albeit in a select group of patients. This paper outlines the preliminary investigation into the validity and reliability of the TATTT.

The study was funded by the Queensland Nursing Council and ethics approval was obtained from the Toowoomba Health Service District Ethics Committee and the University of Southern Queensland's Human Research and Ethics Committees.

Method

A computer package for the TATTT was written in Visual Basic on a Pocket PC 2002 device running the Windows CE (Compact Edition) 3.0 operating system.

As a random sample of participants for this initial investigation of the TATTT was not required, volunteers were sought from among qualified triage nurses employed at the two hospitals involved in this study. Ten triage nurses from Toowoomba Health Service (THS) and five from Princess Alexandra Hospital (PAH) were enrolled.

A self directed training package on the TATTT and the pocket PC application was developed. The training package required approximately 20 – 30 minutes self directed reading time by participants prior to using the TATTT. The training package was provided to participants following administration of five pre session written scenarios for rating. Opportunity to clarify any misunderstandings with a member of the research team (AW) was provided prior to the rating session. This was followed by a personal demonstration in the use of the pocket PC software application.

Each participant then rated nine written scenarios, five simulated scenarios on video and one computer simulated scenario with the TATTT. One month later each participant rated the same nine written scenarios, five simulated scenarios on video and one computer simulated scenario again with the TATTT.

After the completion of the repeat rating session, all 15 participants then underwent an individual semi structured interview which focussed on their opinions of the TATTT; the ease of use of the TATTT software; the ease of use of the pocket PC; and their participation in the research project in general. All interviews were conducted by a

member of the research team (DH) and lasted approximately 30 minutes. .Interviews were tape-recorded and later transcribed verbatim.

Inspiration for the scenarios was obtained from actual case records of patients presenting to THS ED. Details were altered to protect the confidentiality of actual patients. Vital signs and assessment findings incorporated into the scenarios were utilised verbatim from the triage record. All scenarios were reviewed by the research team in relation to the level of information provided and consensus of the expected triage category. The variation in presentation format of the scenarios was performed in order to examine the relative merits of providing different approaches to written presentations of scenarios.

In order to gain preliminary information on the validity of the TATTT parallel triaging of actual adult injury based presentations to THS ED was undertaken. Between April 8, 2003 and May 15, 2003 one member of the research team (AW) parallel coded 58 adult trauma patients presenting to the ED at THS. Patients were eligible for inclusion if they were over the age of 15 years and presented as a consequence of an injury. Envenomations, self poisonings, eye injuries, and retained foreign bodies were excluded.

Patients presenting to THS ED were triaged and given their urgency categorisation by a study participant via conventional means utilising the ATS. The patient was then triaged using the TATTT on pocket PC. The primary triage nurse was blinded to the TATTT assessment criteria and results.

Data Analysis

A first impression of validity of the TATTT was obtained by examining the agreement between ATS and TATTT. Quantitative measures of agreement were obtained using percentage agreement and kappa statistics (with associated 95% confidence intervals). Further information was provided by a detailed clinical review of every patient in which the TATTT and ATS codes differed. Where Kappa values were low, indicating lack of agreement, a test for bias (an extension of McNemar's test) was used to determine the direction of the disagreement.

Inter-rater reliability was described using descriptive statistics and measured against the 'expected' triage category for each simulation provided by the research team. Inter-rater reliability was measured using percentage agreement, kappa statistics and 95% confidence intervals for kappa

Intra-rater reliability was measured by comparing initial codes with one-month re-codes of the same scenarios for each of the 15 coders. Intra-rater reliability was measured using percentage agreement, kappa statistics and 95% confidence intervals for kappa (Fleiss, 1981).

Percentage agreement overestimates reliability because it does not correct for chance (accidental) agreement between coders. The Kappa statistic quantifies agreement after correction for chance. Kappa values of 0.6 or above are deemed by the ACEM to be desirable in triage (Australasian College for Emergency Medicine, 2000).

In this study there were too few pairs differing by more than one triage category, to justify the use of weighted kappas.

Results

In total, the fifteen coders were required to code each of the fifteen simulations on two separate occasions one month apart. It appears one coder accidentally missed coding one of the written simulations on the first occasion. Table 1 summarises the overall test-retest reliability for all simulations for which data was available. The overall percentage agreement is 87%. The overall kappa statistic is 0.82 with a 95% confidence interval from 0.76 to 0.88.

The overall agreement attained using the TATTT on the written simulations is summarised in Table 2. The overall percentage agreement is 90.3% and the kappa statistic is 0.86 with an associated 95% confidence interval of [0.79, 0.94].

The overall agreement attained using the TATTT on the videoed simulations is summarised in Table 3. The overall percentage agreement is 81.3% and the kappa statistic is 0.76 with an associated 95% confidence interval of [0.65, 0.87]. From Table 3 it can be seen that there were disagreements between the expected triage code and those actually produced by the triage nurses. Estimated kappa values are provided for the test-retest reliability on each video. However, as the width of the confidence intervals demonstrates, the sample size is too small to produce any reliable estimates.

The overall agreement attained using the TATTT on the computer simulation is summarised in Table 4. From Table 4 we can see: 13 of the 15 coders (87%) rated the simulation as category 3 on the initial triage; 14 of the 15 coders (93%) rated the simulated patient as category 3 on the re-test. The overall test-retest agreement was 80% but the kappa statistic was only $\hat{\kappa} = -0.10$ with an associated 95% confidence interval of -0.53 to 0.33. The low kappa is a product of the lack of variability in the scenario. That is, as there is strong agreement that the 'correct' rating is 3, any deviation from this has a large impact on the kappa statistic.

All 15 coders were asked to code five written scenarios using their existing techniques and knowledge before they received any training on the TATTT. The results are shown in Table 5. The data in Table 5 shows that there is considerable difference in coding of written scenarios under the current ATS system. None of the five simulations received full agreement across all 15 coders. Three of the five simulations showed close to 50:50 spreads across two triage categories and one simulation had results spread across 3 triage categories. In comparison, data from the TATTT coding of the other nine written scenarios showed five of the nine simulations recorded full agreement across all 15 coders and none varied across more than two triage codes.

The agreement between the triage category assigned using existing triage practices and the triage category assigned by the TATTT (operated by AW) during parallel coding are summarised in Table 6. The overall percentage agreement between the two systems was approximately 47% with an estimated kappa of $\hat{\kappa} = 0.19$ (95% confidence interval [-

0.02, 0.40]). In total 27 of the 58 patients (47%) received the same triage rating on both methods, 22 (38%) received a more urgent rating on the TATTT than by conventional methods and nine (16%) received a less urgent rating on the TATTT than by conventional methods. In all but one case the level of difference between the TATTT and existing practice was only by a factor of 1 category. The test for bias produced statistically significant evidence of a bias towards more urgent coding on the TATTT (x^2=24, df=10, p<0.01).

An example of a patient whose triage score was increased is: an elderly woman presented in a wheel chair following a fall off a chair and landed on her left buttock. The accident was attributable to incoordination. She has difficulty mobilising, has a slight tachycardia and rated her pain as 5/10 on a numeric rating scale. This patient was then given a triage categorisation of 4 where as the TATTT gave the patient a triage category of 3.

An example of a patient whose triage score was decreased is: a young male patient presented with a painful knee joint following a football match the previous day. He was partial weight bearing, neurovascularly intact, nil obvious deformity, had utilised 'RICE' appropriately, rated his pain as 2/10 at rest, declined analgesia and had normal vital signs. This patient was then given a triage categorisation of 4, which is consistent with the traditional notion of acute presentation (within 24 hours) of a minor orthopaedic complaint, whereas the TATTT gave the patient a triage category of 5.

Following a semi structured interview of all participants involved in the study the following general perspectives were able to be made:
- The TATTT application was viewed positively;
- The pocket PC was found to be easy to use;
- Participants believed that the TATTT provided clear direction in the triage assessment process;
- The TATTT increased the level of confidence participants felt with the decision reached; and
- Participants felt they would be comfortable adopting the TATTT in clinical practice.

Discussion

Inter-rater Reliability of the TATTT

Overall the results have shown the inter-rater reliability of the TATTT to be high. The study's overall (test retest) kappa statistic was 0.82 for 448 ratings of the simulated triage. This level of inter-rater agreement exceeds that of previous studies on the NTS, Canadian Triage and Acuity Scale and by the Emergency Severity Iindex (Whitby, Ieraci, Johnson, et al, 1997; Dilley and Standen, 1998; Beveridge, Ducharme, Janes, et al, 1999; Wuerz, Milne, Eitel, et al, 2000; Wuerz, Travers, Gilboy, et al, 2001). This result also exceeds the ATS reproducibility standard of a weighted kappa statistic of at least 0.60, specified by the ACEM (Australasian College for Emergency Medicine, 2000). In making this comparison it is important to note that this study only involved one patient presentation

group. In contrast, all of the studies previously cited involved the full gamete of adult ED presentations. There is also the possibility that this high level of reliability is perhaps an artefact of the use of simulations. None of the three forms of simulation used were able to fully test 'real life' clinical observation or communication skills. In contrast, some of the studies cited also involved real patient presentations.

The scenarios tested prior to the education session also provided for some comparison between the ATS and the TATTT. The results show that the TATTT appears to promote more consistent results than conventional practice utilizing the ATS. In the pre-education session results, none of the five simulations received full agreement across all 15 coders. Three of the five simulations showed close to 50:50 spreads across two triage categories and one simulation had results spread across 3 triage categories. These results are consistent with those of other studies on the NTS (Considine, Ung, Thomas, 2000; Dilley and Standen, 1998; Doherty, 1996). In comparison, data from the TATTT coding of the other nine written scenarios showed five of the nine simulations recorded full agreement across all 15 coders and none varied across more than two triage codes.

These encouraging results suggest that the tool is easily understood and can readily be applied with a high degree of consistency following limited training.

Simultaneous Parallel Coding

Clinical audit of all cases of variance was undertaken by the research team and it was concluded that the TATTT gave more appropriate scores than current practice in each case.

In the provision of this audit process the research team was mindful that retrospective evaluation of triage decisions is problematic and open to charges of bias. As Goodacre, Gillet, Harris, et al (1999) identified, when they performed a retrospective audit of nursing triage decisions using the NTS, there was only fair to moderate consistency between physician reviewers auditing triage decisions. Equally, as Brillman, Doezema, Tanberg, et al (1996) states, verifiable gold standards for triage do not exist. Therefore the consensus of expert opinion by the research team is the only available option for deriving what is believed to be the appropriate triage score for that patient based on the triage information available. This consensus view of expert opinion has been utilised previously in other research to identify the expected triage category (Considine, Ung, Thomas, 2000).

Algorithms and Triage

Algorithms can be utilized to provide decision support in the allocation of clinical urgency (Gerdtz and Bucknall, 1999). Stephens, Pokorny and Bowmen (1997) suggest this is required so that all patients are assessed similarly and according to standards of practice. This need for a systematic method of assessment has been widely recognised (Williams, 1992). The TATTT aims to address this need by its provision of a consistent, systematic approach to the assessment of patients which also supports the decision making process. The parallel triage component of the project highlighted this need as it

was evident for example, that pain assessment was not performed on every patient in the sample group. The TATTT's systematic approach mandates the assessment of a patient's level of pain on a numeric rating scale as part of the triage assessment process. It was also shown that even when pain was assessed, the TATTT appeared to be more responsive to the patients' self reported pain rating than current triage practice. For example a patient from a motor bike accident presented with a clearly deformed forearm which he rated as being 9/10 on a numeric rating scale for pain and was allocated a triage score of 3. This acuity rating may be due to this triage nurse's interpretation of the ATS category 3 clinical discriptor pertaining to pain of, "moderately severe pain any cause – requiring analgesia" (Australasian College for Emergency Medicine, 2000). It may also be consistent with the well documented tendency of healthcare providers to under assess pain relative to the patient's self report (Tait and Chinball, 2002). The patient received a triage score of 2 using the TATTT. It is the opinion of the research team that this triage score is more commensurate with this patients needs and consistent with the ATS category 2 clinical discriptor of, "very severe pain - any cause" (Australasian College for Emergency Medicine, 2000).

There are a number of reported limitations of algorithmic approaches to triage cited in the literature. It is important to consider these limitations individually in relation to the TATTT.

1. The use of lengthy algorithms may unnecessarily delay a time critical patient at triage (Gerdtz and Bucknall, 1999). It is anticipated that the TATTT will require no longer a time frame to complete in entirety than current practice affords for the purposes of triage assessment. Equally it will not delay the time to treatment in the critical patient as the application allows the operator to select the most glaringly obvious presenting feature equivalent to a category one at which point the application closes such that the entire tool does not have to be completed. Also the tool's internal accumulative scoring feature means that when a threshold number of critical features have been reached the tool automatically displays a triage category score of one without looking for any further information.

2. Algorithms impose restrictions on performance by limiting the development of more flexible ways of assessing patients at triage (Gerdtz and Bucknall, 1999). Measures have been taken to allow for some flexibility in the tool which reflects the complexity of the clinical practice environment and to respect the individual clinical judgement of the triage nurse. For example, although the TATTT provides some clear threshold criteria for the identification of a high risk mechanism of injury, it also allows for the assignment of a high risk mechanism of injury on the basis of the impression formed by the triage nurse.

3. Failure of appropriate application was reported as the chief reason for a protocol driven triage system to detect critically ill patients (Cooke and Jinks, 1998). In order to address the "failure of application" issue a very detailed training package about each individual element of the TATTT was produced and provided to each participant for self directed perusal. Given the study's overall high level of inter-rater reliability, the TATTT's universal nature and its bias toward 'up triage' it

can be argued that the critically ill patient will always be detected. Equally, given the limited level of self directed training that occurred, it is anticipated that the level of inter-rater reliability for the TATTT can be improved upon with a face to face teaching session to augment this self directed component.

4. Patients with more than one presenting complaint may be triaged by more than one algorithm, each leading to a different category (Goodacre, Gillet, Harris, et al 1999). The TATTT is universally applied to all adult injury based presentations regardless of the aetiology of the patients presenting complaint. This nullifies the potential for conflicting choices of algorithms which could lead to differing results.

The limitations on this study mainly relate to scope: only 15 coders from only 2 hospitals rated only 1 computer 5 video and 9 written simulations. The coders were not randomly selected. Being interested in research, they may have been more open to new ideas.

Conclusion

Preliminary evidence thus far suggests that the TATTT provides systematically different results compared to current triage practices utilising the ATS; has greater reliability than current triage practices; will be safe for use in the clinical environment; and is acceptable to users. As such, it can viewed as a viable alternative to current triage practice worthy of further investigation.

Further evaluation of the TATTT application in a larger prospective trial is required to further validate the reproducibility of the TATTT, its sensitivity to patient acuity and stratification of patient presentations.

Table 1: Overall test-retest agreement across all coders and all simulations

		Triage category one month later					Total
		1	**2**	**3**	**4**	**5**	
Initial	1	22	6	0	0	0	28
triage	2	1	44	1	0	0	46
category	3	1	0	62	15	0	78
	4	0	0	6	52	0	58
	5	0	0	0	0	14	14
Total		24	50	69	67	14	224*

* one observation was missed

Table 2: Summary of agreements obtained on the written scenarios using the TATTT.

Written scenario number	Expected triage	Agreement, time 1	Agreement, time 2	Reliability (Kappa)	95% confidence interval
1	3	15/15	11/15	0.00	-1.00 to 1.00
2	4	10/15	12/15	0.33	-0.21 to 0.88
3	3	14/14	14/14	1.00	n/a
4	4	15/15	15/15	1.00	n/a
5	5	15/15	15/15	1.00	n/a
6	2	14/15	15/15	0.00	-1.00 to 1.00
7	1	15/15	15/15	1.00	n/a
8	4	11/15	13/15	0.19	-0.44 to 0.81
9	2	15/15	15/15	1.00	n/a

Table 3: Summary of agreements obtained on the video simulations using the TATTT

Video number	Expected triage	Agreement, time 1	Agreement, time 2	Reliability (Kappa)	95% confidence interval
1	1	10/15	7/15	0.29	-0.16 to 0.74
2	2	13/15	13/15	0.42	-0.24 to 1.00
3	3	6/15	6/16	0.44	-0.02 to 0.91
4	4	14/15	12/15	0.44	-0.23 to 1.00
5	5	14/15	14/15	1.00	n/a

Table 4: Summary of agreements obtained on the computer simulation using the TATTT.

		Triage category one month later		Total
		3	**4**	
Initial triage	3	12	1	13
category	4	2	0	2
Total		14	1	15

Table 5: Coding of pre session written scenarios using the ATS

Simulation	ATS triage code					Total
	1	**2**	**3**	**4**	**5**	
w10	0	9	6	0	0	15
w11	0	0	8	7	0	15
w12	10	4	1	0	0	15
w13	13	2	0	0	0	15
w14	0	0	7	8	0	15
Total	23	15	22	15	0	75

Table 6: Comparison of triage ratings from parallel coding using existing practices and the TATTT

		TATTT Triage (AW)					
		1	**2**	**3**	**4**	**5**	**Total**
Existing triage practice (Multiple coders)	1		1				1
	2		1				1
	3		7	9	2		18
	4		1	14	17	6	38
	5						
	Total		10	23	19	6	58

References

Australasian College for Emergency Medicine (2000). Guidelines for Implementation of The Australasian Triage Scale in Emergency Departments, Australasian College for Emergency Medicine.

Australasian College for Emergency Medicine (2000). Policy Document – The Australasian Triage Scale.

Beveridge R, Ducharme J, Janes L, Beaulieu S, Walter S (1999). "Reliability of the Canadian Emergency Department Triage and Acuity Scale: interrater agreement." Annals of Emergency Medicine 34(2): 155-159.

Brillman J, Doezema D, Tanberg D, Sklar D, Davis K, Sims S, Skipper B (1996). "Triage: Limitations in predicting need for emergent care and hospital admission." Annals of Emergency Medicine 27(4): 493-500.

Commonwealth Department of Health and Family Services and the Australasian College of Emergency Medicine (1997). The Australian National Triage Scale: a user manual.

Considine J, Ung L, Thomas S (2000). "Triage nurses' decisions using the National Triage Scale for Australian emergency departments." Accident & Emergency Nursing 8(4): 201-209.

Considine J, Ung L, Thomas S (2001). "Clinical decisions using the National Triage Scale: how important is postgraduate education?" Accident & Emergency Nursing 9(2): 101-108.

Cooke M, Jinks S (1998). "Does the MAnchester triage sytem detect the critically ill?" Journal of Accident and Emergency Medicine 16: 179-181.

Dilley S, Standen P (1998). "Victorian nurses demonstrate concordance in the apllication of the National Triage Scale." Emergency Medicine 10: 12-18.

Doherty S (1996). "Application of the National Triage Scale is not Uniform." Australian Emergency Nursing Journal 1(1): 26.

Fleiss J (1981). Statistical Methods for Rates and Proportions. NewYork, Wiley.

Gerdtz M, Bucknall T (1999). "Why we do the things we do: applying clinical decision-making frameworks to triage practice." Accident & Emergency Nursing 7(1): 50-57.

Goodacre S, Gillet M, Harris R, Houlihan K (1999). "Consistency of retrospective triage decisions as a stadardised instrument for audit." Journal of Accident and Emergency Medicine 16: 322-324.

Monash Institute of Health Services Research (2001). Consitency of Triage in Victoria's Emergency Departments, Literature Review: 1-41.

Stephens G, Pokorny M, Bowman J, (1997). "The Effects of Inservice Education on the Institution of Triage Protocols." Journal of Nursing Staff Development 13(4): 189-192.

Tait R, Chinball J (2002). "Pain in older subacute care patients: Associations with clinical status and treatment." Pain Medicine 3(3): 231-239.

Whitby S, Ieraci S, Johnson D, Mohsin M (1997). Analysis of the process of triage: the use and outcome of the National Triage Scale. Liverpool, Liverpool Health Service.

Williams, G. (1992). "Sorting Out triage." Nursing Times 88(30): 34-36.

Wuerz R, Milne L, Eitel D, Travers D, Gilboy N, (2000). "Reliability and validity of a new five-level emergency department triage instrument." Academic Emergency Medicine 7: 238-240.

Wuerz R, Travers D, Gilboy N, Eitel D, Rosenau A, Yazhari R (2001). "Implementation and Refinement of the Emergency Severity Index." Academic Emergency Medicine 8(2): 170-176.