February 2024

Emergency Triage Education Kit, 2nd edition, scenario validation final report

Professor Marie Gerdtz, the University of Melbourne

Introduction

The Emergency Triage Education Kit (ETEK) was developed and implemented in 2007 to optimise consistency of triage using the Australasian Triage Scale (ATS). The program was funded and developed by the Commonwealth Department of Health and Ageing Acute Care Division (DoHA) in collaboration with national peak professional bodies and stakeholders.

The University of Melbourne was contracted by DoHA to write the original ETEK content including:

1. The development of course learning outcomes, course syllabus, approaches to teaching and learning
2. Assessment.

In October 2022 the Commonwealth Department of Health engaged the Australian Commission on Safety and Quality in Health Care (the Commission) to update the ETEK. A 12-month project was established to produce revised content in line with clinical guidelines and user feedback (ETEK, 2nd edition). A Project Advisory Group was established to support this work updating the content in line with current evidence, clinical practice guidelines and national standards. As part of this project the Commission engaged The University of Melbourne to expand and re-validate the existing set of clinical scenarios for teaching and assessment purposes.

Methodology for this work was established as part of the original 2007 project. The method was published in two peer reviewed papers[1,2] and involves the validation and testing inter-rater reliability of scenarios for the purposes of teaching consistent application of the Australasian Triage Scale (ATS) based on known predictors of urgency.

Aim

The objective of this project was to develop a valid and reliable set of clinical scenarios for teaching and assessment purposes for inclusion in the ETEK, 2nd edition.
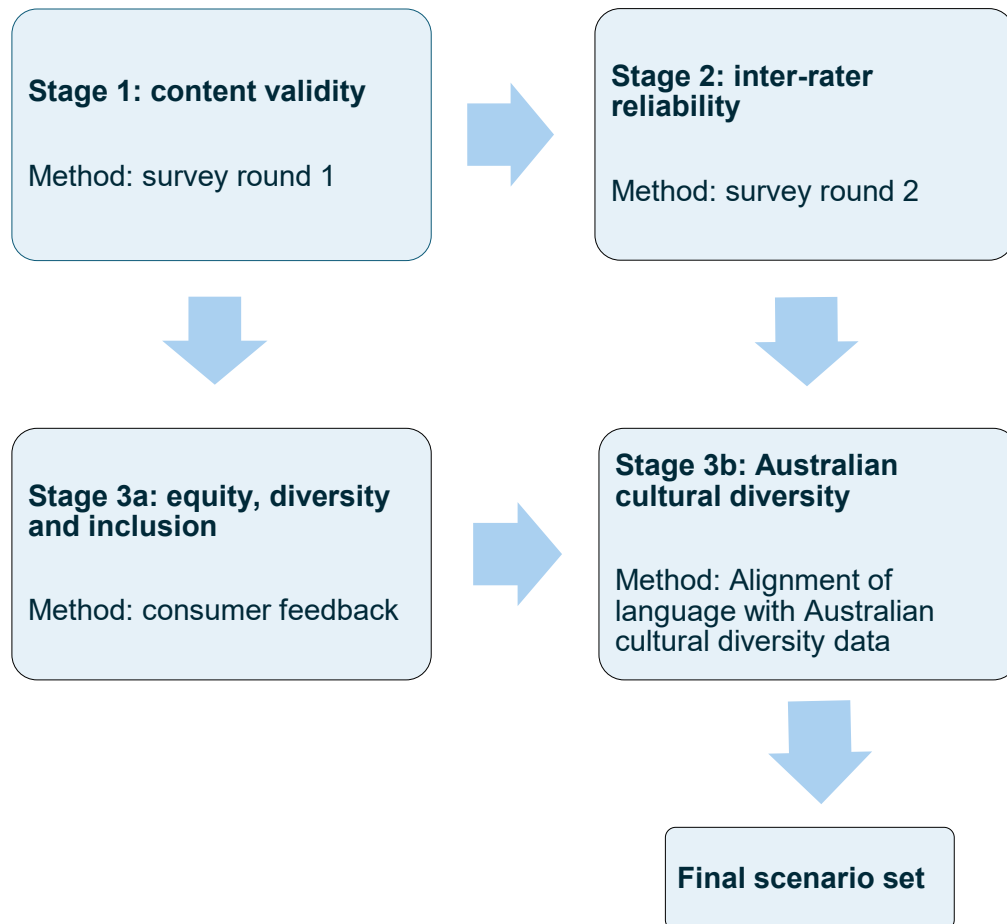
Approach

Ethical approval

Ethical approval to conduct this low-risk project was sought via The University of Melbourne Human Research and Ethics Committee. Approval was granted on the 23 October 2023 (reference number: 2023-27536-46148-5). The project was performed in accordance with NHMRC Guidelines. Participants were given a plain language statement and provided written consent.

Design

A cross-sectional survey design was used. This involved validation of a large set of clinical scenarios (N=237) by expert triage nurses. The original scenarios used were those developed by The University of Melbourne for the original ETEK project. These were re-tested and revised in accordance with the Australasian College for Emergency Medicine's Guidelines on the implementation of the ATS.[3,4]

A systematic mapping exercise was conducted to align the original scenario set to the new syllabus. The study consisted of three stages, shown in Figure 1.

Figure 1. Summary of the clinical scenario revision process

```
┌────────────────────────────┐        ┌────────────────────────────┐
│ Stage 1: content validity  │───────▶│ Stage 2: inter-rater       │
│                            │        │ reliability                │
│ Method: survey round 1     │        │                            │
│                            │        │ Method: survey round 2     │
└────────────────────────────┘        └────────────────────────────┘
             │                                      │
             ▼                                      ▼
┌────────────────────────────┐        ┌────────────────────────────┐
│ Stage 3a: equity, diversity│───────▶│ Stage 3b: Australian       │
│ and inclusion              │        │ cultural diversity         │
│                            │        │                            │
│ Method: consumer feedback  │        │ Method: Alignment of       │
│                            │        │ language with Australian   │
│                            │        │ cultural diversity data    │
└────────────────────────────┘        └────────────────────────────┘
                                                    │
                                                    ▼
                                       ┌────────────────────────────┐
                                       │ Final scenario set         │
                                       └────────────────────────────┘
```

Stage 1: content validity

The purpose of survey round 1 was to establish the content validity of the scenarios. Survey round 1 used the existing ETEK scenarios. To reduce respondent burden, the triage scenarios were administered via two surveys – one comprising scenarios 1-119 (part A) and another comprising scenarios 120-237 (part B). For each scenario, participants were asked to determine the ATS triage category, select the category of the presenting problem, and were provided the option of answering four questions:

- Does the information presented in this scenario reflect the 2016 ATS Guideline?
- Is the information presented in this case sufficient for you determine a triage category using the ATS?
- List and describe the top three factors in the scenario that influenced your decision
- Enter any other comments about this scenario.

Participants were also required to enter demographic details regarding their age, gender, years of experience as a nurse and as a triage nurse, qualification, state/territory and place of work (to allow hospital peer group to be determined).

The surveys were developed and administered via RedCap (Research Electronic DataCapture). RedCap is a secure, web-based application designed to support data capture, providing:

- An intuitive interface for validated data entry
- Audit trails for tracking data manipulation and export procedures
- Automated export procedures for seamless data downloads to common statistical packages
- Procedures for importing data from external sources.

The survey was accessed by participants (for details, see section about participants) via an anonymous link.

Following survey round 1, for each scenario the frequencies of each urgency category were calculated, from which the modal category was determined. For survey round 1, partial survey responses (i.e. surveys only partially completed by participants) were included in the analysis.

Scenarios for which there was less than 70% agreement were considered by the research team, together with survey responses to the qualitative questions. The clinical research team amended scenario wording and clinical content in line with the ACEM guideline.

The research team also reviewed scenarios for which agreement on the urgency category was greater than 70% and amended wording or clinical content, where deemed necessary in line with the ACEM guideline.

Stage 2: Inter-rater reliability

The purpose of survey round 2 was to assess the inter-rater reliability of scenarios for inclusion as teaching and assessment tools within the ETEK, 2nd edition. Survey round 2 used the scenarios revised after survey round 1. The output of this process was a large validated and reliable scenario set aligned to the ETEK, 2nd edition, syllabus.

To reduce respondent burden, the triage scenarios were administered in two parts – one comprising scenarios 1-119 (part A) and another comprising scenarios 120-237 (part B). For each scenario, participants were asked to determine the ATS triage category, select the category of the presenting problem, and were provided the option of answering four questions:

- Does the information presented in this scenario reflect the 2016 ATS Guideline?
- Is the information presented in this case sufficient for you determine a triage category using the ATS?
- List and describe the top three factors in the scenario that influenced your decision
- Enter any other comments about this scenario.

The survey was accessed by participants via an anonymous link. For survey round 2, only complete survey responses (i.e. surveys in which participants rated all triage scenarios) were included in the analysis because calculation of the kappa statistic for interrater reliability requires the number of participants for the entire response set.

For all scenario sets we aimed to attain moderate to excellent levels of agreement using kappa ($κ > 0.6$). The sample size for the study is based on the primary study objective to assess the reliability of the triage scenarios. The number of participants is determined by $n > 2K^2$ where K is the number of ordinal points on a rating scale. For a 5-point triage scale a sample size of 50 has been determined.[5] The total number of scenarios contained within the survey (m) will be determined based on a mapping exercise that will ensure the scenario set contains both a sufficient set of common presentations and enough cases in each triage category and presentation types so as to limit the influence of case-specific effects on the outcome.

Scenarios for which there was less than 70% agreement were considered by the research team, together with survey responses to the qualitative questions. The clinical research team amended scenario wording and clinical content.

The research team also reviewed scenarios for which agreement on the urgency category was greater than 70% and amended wording or clinical content, based on respondent feedback and in line with the ATS guidelines.

Stage 3a: Consumer input

The purpose of stage 3a was to identify potential biases in the written presentation of information in the scenarios. A consumer representative was provided with the scenarios resulting from stage 1 and asked to provide recommendations on how the language could be improved in terms of recommendations on basic principles of equity, diversity and inclusion.

The research team reviewed and summarised the consumer's comments, and recommendations are provided about how the Commission may wish to modify the language of some of the scenarios.

Round 3b: Reflection of Australian cultural diversity

To ensure that the names used in the triage scenarios were reflective of Australia's cultural diversity in 2023, data published by the Australian Bureau of Statistics on the Australian population by country of birth (Australian Bureau of Statistics, 2022) were interrogated to identify the top 10 countries of birth, which were used to inform the proportions of names in the scenario set that should be altered to reflect these countries. Recommendations are provided regarding scenario modification to reflect Australian cultural diversity.

Participants

Participants were currently practicing, experienced (more than 5 years) emergency nurses, and were sourced from each of the Australian states and territories.

Participants in round 1 (n=22 total invited, 11 for each half of the survey) were experienced triage nurse educators and clinicians. Table 1 shows the sampling frame used and participant numbers for survey round 1.

Participants in round 2 (n=121; 60-61 invited for each half of the survey) were experienced and currently practicing triage nurses. Participants were recruited by the Commission using an expression of interest process and appointed using stratified sampling technique to ensure representation of members from all Australian states and territories.

Table 2 shows the sampling frame used and actual participant numbers for survey round 2.

Table 1. Sampling frame for survey round 1

State or territory	Target n per state	Actual n (scenarios 1-119)	Actual n (scenarios 120-237)	Total actual participants per state
NSW	3	1	1	2
Qld	8	3	3	6
SA	2	2	2	4
Tas	1	1	0	1
WA	3	1	2	3
ACT	2	0	1	1
Vic	2	0	1	1
NT	1	0	1	1
Total	**22**	**8**	**11**	**19**

Table 2. Sampling frame for survey round 2

State or territory	Target n per state	Actual n (scenarios 1-119)	Actual n (scenarios 120-237)	Total actual participants per state
NSW	34	5	7	12
Qld	30	2	2	4
SA	9	5	1	6
Tas	5	1	1	2
WA	13	3	6	9
ACT	8	0	0	0
Vic	18	9	8	17
NT	4	0	1	1
Total	**121**	**25**	**26**	**51**

Survey analysis

Raw data was exported from RedCap as a csv file. Data were imported into SPSS version 29 (IBM Corp., Armonk, N.Y., USA) and frequencies for urgency ratings were determined for each scenario.

Descriptive analysis was performed including calculation of frequencies, mean and standard deviation for demographic variables. The data was also explored descriptively by determining concurrence; that is, the percentage of responses for each case scenario in the modal category and spread.

Raw percentage agreement was calculated for the 237 scenarios for the modal response category.

Scenarios were analysed as an entire set and were also categorised according to the chapters contained in the ETEK: Triage fundamentals, Psychological distress and behavioural disturbance, Paediatrics, Pregnancy.

To explore the data, all scenarios in which the modal response category was greater than 60%, and those in which it was greater than 70%, were included in a model to calculate chance-corrected agreement in each group (unweighted kappa).

This approach was taken to identify the maximum number of scenarios appropriate for testing according to the predetermined criteria (kappa > 0.6). The formula for calculating kappa is by Fleiss et al:[6]

$$Kj = \frac{\bar{P}j - pj}{1 - pj}$$

$$= \frac{\sum_{i=1}^{N} nij^2 - Nnpj[1 + (n-1)pj]}{Nn(n-1)pjqj}$$

Notation: N represents the total number of subjects, n is the number of ratings per subject and k is the number of categories into which assignments are made. Let the subscript I where i=1,… N, represent the subjects and the subscript j, where j=1,…k, represent the categories of the scale. Define nij the number of raters who assigned the ith subject to the jth category. The quantity pj is the proportion of all assignments that were to the jth category.[6]

The formula for calculating kappa variance was also from Fleiss et al:[6]

$$Var(k) = \frac{2}{Nn(n-1)(\sum pjqj)^2}$$

$$\times \left| (\sum pjqj)^2 - \sum pjqj(qj - pj) \right|$$

Analysis of kappa statistics was done by programming the above formula into Microsoft Office Excel (2003). Frequencies were entered by scenario and triage code.

For survey round 1, all survey participants were included in the analysis. For survey round 2, only participants who rated all triage scenarios were included in the analysis.

Results

Survey round 1

Following review of participants' feedback in round 1, the research team revised.

The final set of scenarios resulting from round 1 is in Appendix A.

Survey round 2

To reduce participant burden and in an attempt to increase both the response rate and the rate of survey completion, round 2 was conducted via two parts: part A assessing scenarios 1-119, and part B assessing scenarios 120-237.

Only surveys in which participants rated every scenario were included in the analysis, because calculation of the kappa statistic for interrater reliability requires the number of participants for the entire response set. Table 3 summarises participant demographics.

Survey round 2 part A was completed by 25 participants; survey round 2 part B was completed by 26 participants. For each part of survey round 2, the overall kappa statistic is summarised in Table 4. For both parts of survey round 2, the kappa statistic was over 0.45, indicating moderate interrater reliability.

Considering only the scenarios for which there was at least 70% agreement on the triage rating (n=132 in total, or n=78 scenarios for survey part A, and n=54 scenarios in survey part B), the interrater reliability kappa statistic was over 0.6, indicating 'Good' reliability. Therefore, in total, there are m=132 scenarios in which the kappa statistic is >0.6, or 'Good'. The scenarios with at least 70% agreement are presented in Appendix A; those with less than 70% agreement are presented in Appendix B.

Table 3. Demographics of survey round 2 participants

	Part A – scenarios 1-119	Part B – scenarios 120-237	All scenarios
Age – mean (SD)	39.48 (6.81)	37.71 (7.04)	38.59
Years worked as an RN – mean (SD)	14.24 (6.31)	14.31 (7.31)	14.27
Years worked as a triage nurse – mean (SD)	10.12 (5.83)	10.0 (6.85)	10.06
Appointment			
Registered nurse	3 (12%)	5 (19%)	8 (16%)
Clinical nurse specialist	5 (20%)	10 (38%)	15 (29%)
ANUM	0 (0%)	2 (8%)	2 (4%)
Nurse practitioner	1 (4%)	0	1 (2%)
Nurse educator	13 (52%)	8 (31%)	21 (41%)
Clinical nurse consultant	1 (4%)	1 (4%)	2 (4%)
Other	2 (8%)	0 (0%)	2 (4%)
Highest qualification			
Registered nurse certificate	0	1 (4%)	1 (2%)
Bachelor of nursing	2 (8%)	8 (31%)	10 (20%)
Certificate/graduate certificate in emergency nursing	15 (60%)	7 (60%)	22 (43%)
Masters	8 (32%)	10 (27%)	18 (35%)
Total participants with complete surveys	**25**	**26**	**51**

Table 4. Summary of kappa statistics and number of scenarios for survey round 2, parts A and B

	PART A - scenarios 1-119			PART B - scenarios 120-237		
	Number of scenarios (m)	Number of nurses (n)	Kappa*	Number of scenarios (m)	Number of nurses (n)	Kappa*
Overall	119	25	0.49	118	26	0.46
≥70% agreement	78	25	0.61	54	26	0.65
Topic						
Paediatrics	36	25	0.48	5	26	0.65
Psychological distress and behavioural disturbance	24	25	0.37	16	26	0.25
Pregnancy	17	25	0.42	4	26	0.5
Triage fundamentals	18	25	0.65	45	26	0.5
Pain	24	25	0.49	48	26	0.38

*Values of kappa correspond to the following strength of agreement:

0.21-0.40	Fair
0.41-0.60	Moderate
0.61-0.80	Good
0.81-1.00	Very good

Round 3 – Consumer input

Consumer feedback is presented in Appendix C; the research team made adjustments based on this feedback where deemed appropriate.

Conclusions

- 132 of 237 (or 55%) of scenarios achieved agreement of at least 70% on triage category. These scenarios can be used for both teaching and assessment of triage decision making (Appendix A).
- The remaining 105 scenarios, for which agreement on triage category is less than 70%, may be used for teaching purposes where cases are presented and discussed with an experienced triage nurse (Appendix B).

Recommendations

- To optimise scenarios with less than 70% agreement, a further validation round might be considered since these scenarios have undergone further revision post round 2. This approach could be undertaken under an amendment to the existing HREC and may yield a larger scenario set for assessment purposes.
- It is recommended that in future validations of the ETEK scenarios, participants are compensated for their time; this is likely to improve the response rate.
- During the survey round 2 data collection phase, the researchers noted that the ACEM Guideline on the Implementation of the Australasian Triage Scale had been updated by ACEM (November 2023).[3] Correspondence with ACEM Acting Manager, Policy and Advocacy, confirmed no material change to predictors of urgency. Specifically, the difference between the 2016 and the 2023 documents are stylistic in nature, meaning there have been no changes to ATS categories that would affect the scenarios. Breakdown of the key changes to the guideline include:
  - Retaining key messages with some structural changes to align with ACEM's new document structure
  - Purpose and background section added
  - Definitions and explanatory notes spread throughout the 2016 version have been placed into a 'terminology' section
  - Section 3.2 The Triage Assessment – bullet point added regarding vital signs to align with the recently published joint statement between ACEM and CENA
  - The addition of the headings 'geriatrics' and 'pregnancy' have been added under Section 5 – Specific Conventions.

- Considerations for revising the pseudonyms used in scenarios: There is no established methodology for ensuring that names used in educational material are representative of cultural and ethnic diversity. One possible approach could be to determine the proportion of the Australian population that is comprised of foreign-born people, identify common names in these countries, and then allocate such names to a corresponding proportion of scenarios. For example, according to 2022 ABS data, the largest groups of overseas-born Australians were English-born (3.7%), Indian-born (2.9%), Chinese-born (2.3%), and New Zealand-born (2.3%).[7] Common names in China and India could be identified, and then these names could be applied to a proportionally appropriate number of scenarios. This approach is associated with several limitations, including that it is difficult to identify reputable sources of naming data for these countries; foreign-born Australians do not

necessarily use their birth name and may use an Anglicised name; using 'common' names from other countries may introduce other forms of bias. In light of these limitations, it is recommended that the ETEK triage scenarios be accompanied with an explanatory note to the effect of:

> *The names, characters, places and incidents used in the scenarios fictitious and do not represent specific individuals. Pseudonyms are used to encourage a person-centred approach.*

References

1. Gerdtz MF, Chu M, Collins M, Considine J, Crellin D, Sands N, et al. (2009). Factors influencing consistency of triage using the Australasian Triage Scale: implications for guideline development. Emergency Medicine Australasia 2009;21(4);277-285.

2. Gerdtz MF, Collins M, Chu M, Grant A, Tchernomoroff R, Pollard C, Harris J, Wassertheil J. Optimising triage consistency in Australian emergency departments: the Emergency Triage Education Kit. Emerg Med Australas 2008;20(3):250-259.

3. Australasian College for Emergency Medicine. Guidelines on the Implementation of the Australasian Triage Scale in emergency Departments (V6). ACEM, 2023. https://acem.org.au/getmedia/51dc74f7-9ff0-42ce-872a-0437f3db640a/G24_04_Guidelines_on_Implementation_of_ATS_Jul-16.aspx

4. Australasian College for Emergency Medicine. Guidelines on the Implementation of the Australasian Triage Scale in emergency Departments (V5). ACEM, 2016.

5. Cicchetti DV. Assessing inter-rater reliability for rating scales: resolving some basic issues. Br J Psychiatry 1976 Nov;129:452-456.

6. Fleiss JL, Nee JCM, Landis JR. Large sample variance of kappa in the case of different sets of raters. Psychol Bull 1979;86(5):974-7.

7. Australian Bureau of Statistics. Australia's Population by Country of Birth. ABS 2022. https://www.abs.gov.au/statistics/people/population/australias-population-country-birth/latest-release.