Reliability and validity of the South African Triage Scale in low-resource settings

Mohammed K Dalwai
MBcHB
DLWMOH001

Thesis presented for the degree of:

Doctor of Philosophy (Emergency Medicine)

In the Division of Emergency Medicine, Faculty of Health Sciences at the University of Cape Town

Supervisor: Lee Wallis
Co Supervisors: Michele Twomey & Katie Tayler-Smith

February 2018

Declarations

I declare that this dissertation is my own unaided work. It is being submitted for the degree of Doctor of Philosophy (Emergency Medicine) to the faculty of Health Sciences, University of Cape Town. It has not been submitted before for any degree or examination at the University of Cape Town or any other university.

Plagiarism Declaration: This dissertation has been submitted to the Turnitin module and I confirm that my supervisor has seen my report and any concerns revealed by such have been resolved with my supervisor.

I confirm that I have been granted permission by the University of Cape Town's Doctoral Degrees Board to include the following publications in my PhD thesis, and where co-authorships are involved, my co-authors have agreed that I may include the publications:

1. Dalwai MK, Twomey M, Maikere J, Said S, Wakeel M, Jemmy JP, Valles P, Tayler-Smith K, Wallis L, Zachariah R. Reliability and accuracy of the South African Triage Scale when used by nurses in the emergency department of Timergara Hospital, Pakistan. South African Med J 2014;104:372. doi:10.7196/samj.7604

2. Dalwai M, Tayler-Smith K, Trelles M, Jemmy JP, Maikere J, Twomey M, Wakeel M, Iqbal M, Zachariah R. Implementation of a triage score system in an emergency room in Timergara, Pakistan. Public Health Action; 2013;I:43–5.

3. Dalwai M, Tayler-Smith K, Twomey M, Wallis L. Developing a reference standard for assessing paediatric triage scales in resource poor settings. African Journal of Emergency Medicine , Volume 5 , Issue 4 , 181 – 184

4. Dalwai M, Valles P, Twomey M, Nzomukunda Y, Jonjo P, Sasikumar M, Nasim M, Razaaq A, Gayraud O, Jecrois PR, Wallis L, Tayler-Smith K. Is the South African Triage Scale valid for use in Afghanistan, Haiti and Sierra Leone? BMJ Glob Health 2017;2:e000160. doi:10.1136/ bmjgh-2016-000160

5. Dalwai M, Tayler-Smith K, Twomey M, Nasim M, Popal AQ, Haqdost WA, Gayraud O, Cheréstal S, Wallis L, Valles P. Inter-rater and intrarater reliability of the South African Triage Scale in low-resource settings

of Haiti and Afghanistan. Emerg Med J Published Online First: 16 March 2018. doi: 10.1136/emermed-2017-207062

Signed by candidate

Mohammed K Dalwai

Signed this 18th day of February, 2018

Acknowledgements

I would like to sincerely thank the following people without whom this thesis would not be possible.

Lee Wallis: I know I said I would complete in two years; five years later here we are. Thank you for your unwavering support and quick replies, even after all this time.

Michele Twomey: The origin of this PhD emanated after many of our coffee talks. Thank you for some of the most interesting discussions and your willingness to push triage further in this world.

Katie Tayler-Smith: Thanks for keeping me honest and true in all my writings, for all the red marks and the gentle corrections. As my wife says, you make me sound intelligent.

Stevan Bruijns: Thank you for always making sure we were all still ok mentally.

Thabi Martin: For supporting those who need it, when they need it. Thank you for believing in me.

Yaseen Khan: We have been on such an adventure; thank you for picking up the slack when I couldn't and providing the support when I needed it.

The MSF crew: Miguel, Pola, Renzo and Jean Paul. Thank you for your vision to make this a reality, for your ability to see the need for a validated triage system to be used across low-resource settings. Thank you, Pola, for all your support in the field and for an amazing data collection tool.

The Staff of MSF: For all the individuals who work every day in the field for MSF. I have had the privilege to work alongside you all. Thank you for always wanting to do the right thing for your patients and never giving up.

My family: To my parents for their unwavering support and for pushing me to finish even when all I wanted to do was to give up. To my brothers and sisters, you may now call me dr dr.

My darling wife: Ghizlaine, I have no words to express the gratitude I have for all your support. This has been a labour of love and you have given me the space, time and support to be able to achieve this. I'm looking forward to spending more time with you.

My son: At 12 weeks old, thank you for trying to thwart my attempts to finish my PhD by keeping me awake throughout the night. I look forward to repaying the favour in my old age.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ACS-COT	American College of Surgeons - Committee on Trauma
ATS	Australasian Triage Scale
CTAS	Canadian Triage and Acuity Scale
EC	Emergency Centre
ECS	Emergency Care Systems
ESI	Emergency Severity Index
ETAT	Emergency Triage Assessment and Treatment
FRENCH	French Emergency Nurses Classification in Hospital Scale
GRC	Gondoma Referral Clinic
GRRAS	Guidelines for Reporting Reliability and Agreement Studies
ICC	Intraclass Correlation Coefficient
KOL	Key Opinion Leaders
KTS	Kampala Trauma Score
MSF	Médecins Sans Frontières
MTS	Manchester Triage System
OTT	One-two Triage
QWK	Quadratically Weighted Kappa
RETTS	Rapid Emergency Triage Treatment Scale
SATS	South African Triage Scale
TEM	Triage Emergency Medicine
TEWS	Triage Early Warning Score
TTAS	Taiwan Triage and Acuity Scale
WHO	World Health Organisation

Abstract

Introduction

Emergency medical care (EMC) is proposed by the World Health Organization (WHO) as being one of the core components of a horizontal approach to improving population health in low-resource settings; triage is considered to be a fundamental part of this field. Most studies exploring triage have focused on high-income countries. In 2004, the Cape Triage Group (CTG) developed the South African Triage Scale (SATS) a scale that uses a physiologically based scoring system together with a list of discriminators - designed to triage patients into one of four priority groups for medical attention. The SATS was designed for use in the South African context to mitigate the limited numbers of doctors and professional nurses. The SATS has been implemented and assessed extensively in South Africa, but its performance across a spectrum of different low-resource settings, particularly non-sub-Saharan African and trauma-only settings, has not been adequately assessed. Médecins Sans Frontières (MSF), an international humanitarian organisation, introduced EMC in 2006 into low-resource settings. In 2011, MSF began introducing the SATS in various projects where it was providing EMC.

Methodology

This was a multi-site retrospective cohort study which sought to assess the reliability and validity of the SATS in different low-resource settings.

Aim 1: To implement and evaluate the SATS in Northern Pakistan by describing the steps of implementation and how accurate nurses were in using the triage scale. After one month of implementation, 370 triage forms from a one-week period were evaluated.

Aim 2: To assess the inter- and intra-rater reliability and accuracy of nurse triage ratings when using the SATS in an emergency centre (EC) in Timergara, Pakistan. Fifteen EC nurses assigned triage ratings to a set of 42 reference vignettes (written case reports of EC patients) under classroom conditions. Inter-rater reliability was assessed by comparing these triage ratings; intra-rater reliability was assessed by asking the nurses to re-triage ten

random vignettes from the original set of 42 vignettes and comparing the duplicate ratings. Accuracy of the nurse ratings was measured against the reference standard.

Aim 3: To improve the ability to measure reliability and validity in paediatric settings by developing a set of paediatric paper-based vignettes using the Delphi methodology. In a two-round consensus building process, a panel of EC experts were asked to independently triage 50 clinical vignettes using one of four acuity levels: emergency (patient to be seen immediately), very urgent (patient to be seen within 10 min), urgent (patient to be seen within 60 min), or routine (patient to be seen within four hours). The vignettes were based on real paediatric EC cases in South Africa. Vignettes that reached a minimum of 80% group consensus for acuity ratings on either round one or two were included in the final set of reference vignettes.

Aim 4: To further assess the reliability of the SATS across MSF-supported hospitals using paper-based vignettes in Afghanistan, Haiti and Sierra Leone. Applying the same methodology as in Northern Pakistan, we assessed reliability under classroom conditions between December 2013 and February 2014.

Aim 5: To assess the validity of the SATS across MSF-supported hospitals between June 2013 and June 2014. Validity was assessed by comparing patients' SATS ratings with their final EC outcomes (i.e., hospital admission, death or discharge) across four sites in Afghanistan, Haiti and Sierra Leone.

Findings

The SATS was able to be easily implemented and accurately completed in a low-resource setting of Northern Pakistan. We recommended further implementation and assessment of reliability and validity in low-resource settings.

Across six sites with a total of 87 nurses, including two trauma-only hospitals in Afghanistan and Haiti, a paediatric-only hospital in Sierra Leone and three mixed medical settings in Afghanistan, Pakistan and Haiti, the SATS demonstrated moderate to substantial reliability. Across all settings in which we measured validity using outcome markers, SATS predicted an increase in the likelihood of admission/death when moving from low- to high-triage acuity. In trauma-only settings of Afghanistan and Haiti, the SATS showed a 1-9% under-triage and

a 2-16% over-triage rate. In mixed medical and paediatric settings, under-triage ranged from 0-76% while over-triage ranged from 2-88%.

A more logical standardised approach to assessing validity was put forward when using outcome markers that would allow easier comparisons to be done across validity studies irrespective of the number of levels the triage scale had.

We developed a set of paediatric vignettes for use in low-resource settings but cautioned against its use after measuring reliability using adult reference vignettes. We found that generic vignettes were poor substitutes in a variety of settings based on a lack of contextualisation and understanding by local nurses.

Conclusion

The SATS has reasonable reliability with good validity across different ECs in various low-resource settings. The SATS is a valid triage tool for prioritisation of patients with trauma in low-resource settings. Its use in mixed EC settings seems justified, but in paediatric settings context-specific adjustments and assessments of its performance would be prudent.

Chapter 1: Introduction

Emergency Care Systems (ECS) addresses a wide range of acute conditions, including injuries, communicable and non-communicable diseases and complications of pregnancy.[1] Emergency Care is an essential part of the health system and often serves as the first point of contact for many people around the world.[2] The World Health Organization (WHO) proposes that ECS is one of the core components of a system wide approach to improving population health in low-resource settings.[2] 54% of all deaths in low- to middle-income countries could be affected by ECS. And yet in many low-resource settings, the ECS is not deemed a priority.[3] The World Health Assembly has called on all its member states to prioritise the establishment of integrated ECS.[2]

The three core components of ECS are: i) scene-based care, ii) transportation care, and iii) provision of emergency care on arrival at the receiving health facility.[4] While all three components are equally important, strategies to address the first and second components are futile if the adequate provision of effective emergency care at the receiving health facility cannot be ensured.

One of the major challenges linked to this third component is around a health facility's capacity to deal with emergency patient caseloads.[4] In low-resource settings, triage has been identified as being one of the weakest links in the ECS.[4] Restructuring the intake area of a hospital and implementing a triage system has been shown to decrease mortality.[5] Triage is the systematic process of determining a patient's priority for treatment based on the severity of their condition when there are limited resources. The principal aim of triage is to ensure that patients receive care in a time relative to their clinical need, in order to minimise patient morbidity and mortality and to ensure the most efficient use of emergency centre (EC) resources.[6]

Médecins Sans Frontières (MSF) is an international humanitarian organisation providing emergency medical assistance to populations in danger across 70 countries.[7] Over the last 40 years, MSF has implemented and used triage in high-conflict zones as well as after natural

or man-made disasters, focusing mainly on pre-hospital triage. Over the last decade, the context of MSF's work has evolved to include health facilities and ECs in low-resource settings. They faced as many others in low-resource environments, the problem of overcrowding in the ECs. This often led to compromised patient safety, increased waiting times and further constraints on already limited resources. These challenges highlighted the need to develop or identify and implement a triage scale that could be used reliably and safely in their ECs.[6,8]

However, identifying such a system proved difficult as so many different types of EC triage scales exist. The four most commonly discussed in-hospital EC triage systems include the Manchester Triage Scale (MTS)[9] the Canadian Triage Acuity Scale (CTAS)[10], the Emergency Severity Index (ESI)[11], and the Australasian Triage Scale (ATS).[12] While these triage systems have been widely used and researched in high-resource settings, it was unknown whether these systems were valid for use in low-resource settings. For one thing, these triage scales were complex. For example, the MTS consists of 52 flow charts that nurses need to understand and apply to each patient scenario.[13] ESI, on the other hand, requires the nurse to determine and assign a number of potential resources that could be used by each patient.[14] These triage scales therefore required a high level of training and needed highly qualified staff to use them reliably, which precludes their use in low-resource settings.

The most widely recommended triage tool for low-resource settings has been WHO's Emergency Triage Assessment and Treatment (ETAT) system.[15] However, this system is only applicable for children under five. In fact, there is no way to triage older children or adults with this tool. In 2004, the Cape Triage Group developed the Cape Triage Score, later renamed the South African Triage Scale (SATS) - a physiologically based scoring system and a list of discriminators designed to triage patients into one of four priority levels (with an additional category for dead on arrival patients).[16] The SATS was intentionally designed for use by entry level, relatively unskilled nurses - nursing assistants - due to the limited numbers of doctors and professional nurses in South Africa.[17] The SATS was revised in 2008 and again in 2012.

The advantages of the SATS over the ETAT are that i) the 2008 SATS has three versions - one for adults, children and infants unlike the ETAT, which is for children under five years only; ii)

the SATS uses more objective-based markers to judge triage priority; and iii) staff training for the SATS is only four hours versus just over three days for ETAT.

Premise of Study

This study was born from the challenges faced while working for MSF. Emergency medical care is provided by MSF in a number of hospitals across the countries in which it works, but until recently lacked the use of any formal triage scale. MSF often works in low-resource settings with high patient caseloads. Given their settings they often encountered poorly skilled staff with basic education. Therefore, in 2011, MSF identified the 2008 SATS as an appropriate triage scale and began piloting the tool in a number of settings.

The SATS has been extensively evaluated in South Africa,[18–21] but formal examination of its use in other low-resource settings has received little attention. Key attributes that are essential to assess when introducing a triage tool into a system are its validity and reliability in that setting. With regard to the validity of a triage tool, it should be able to determine an acuity level as closely as possible to the patient's true acuity. For a tool to be reliable, the same results should be generated every time with the same healthcare worker (intra-rater reliability) and there should be agreement among health care workers (inter-rater reliability) regarding a patient's acuity level, irrespective of their true acuity.[22]

A tool's reliability and validity is correlated with the effectiveness of triage. Given the paucity of knowledge in the low-resource settings around effectiveness of triage, further investigation was needed.

Aim

The overall aim of this PhD was to examine the implementation of the SATS by MSF in its emergency centres, and to evaluate the reliability and validity of this tool in these settings.

Objectives

1. To describe the implementation and accuracy of the SATS in a low-resource MSF setting.
2. To develop a set of context-specific reference vignettes to evaluate paediatric triage systems in low-resource settings.
3. To assess the reliability and validity of the SATS when used by EC nurses in a low-resource setting.
4. To assess the inter- and intra-rater reliability of the SATS when used by EC nurses across multiple low-resource settings.
5. To determine the validity of the SATS across multiple low-resource settings.

Ethics approval

The study was approved by the MSF ethics review board and the University of Cape Town Human Research and Ethics Committee. National Ethical review boards in Afghanistan, Pakistan, Sierra Leone and Haiti each gave approval for the respective studies in each country (Appendices 1-5).

Funding

Funding was provided by the Medical Research Council (MRC) to the principal researcher under the MRC Clinicians Researcher Program.

Data Security

All paper-based data were kept at the office of the principal investigator protected by a double locking mechanism. This information will be destroyed after five years of storage. All digital information is kept in a Protection of Personal Information Act (POPI) compliant, two factor authenticated cloud-based storage system. Some information has been anonymised and made available in an open access format as part of some of the publications from this research.

Research structure

Each of the five objectives was completed as a separate study and all are presented in the following chapters. Chapter 2 details the current literature on the reliability and validity of triage in low-resource settings. Chapters 3 to 7 contain peer-reviewed published papers that report the methodology, results and main findings related to the chapter's objectives. Each chapter contains supplemental analysis and discussion of important learning points from each publication including additional relevant but unpublished information. Chapter 8 presents the discussion and brings all the findings of each chapter together in terms of the overall thesis. Chapter 9 highlights the conclusions and further recommendations.

Chapter 2: Literature Survey

Search strategy

An extensive literature search was undertaken at the beginning of the research period in 2013, we did not undertake a formal systematic review but rather regularly repeated the search throughout the years of the PhD, driven by separate publication reviews. The last full literature search was conducted in January 2018 as part of the final write up; however, to keep with the natural contemporaneous nature of undertaking research, studies related to the research question published after 2014, are only discussed in the discussion chapter (chapter 8).

We searched peer-reviewed literature published in the English language from 1980 onwards. We took a three-pronged approach and included databases, electronic journals and bibliographies of articles reviewed. Databases searched included PubMed central, Medline, CINAHL, Cochrane Library and Google Scholar. Search terms included individual and combinations of the following keywords*: triage, high resource settings, Canadian Triage and Acuity Scale, CTAS, Emergency Severity Index, ESI, Manchester Triage System, MTS, Australasian Triage Scale, ATS, Kampala Trauma Score, KTS, World Health Organization, WHO, low resource settings, South African Triage Scale, reliability, validity, low and middle-income countries, LMIC, ETAT, emergency triage, emergency triage assessment and treatment, accuracy, Afghanistan, Sierra Leone, Haiti, Pakistan,* and *Sub-Saharan*.

We also hand searched a number of journals including:
Academic Emergency Medicine
Accident and Emergency Nursing
Acute Medicine and Surgery
African Health Sciences
African Journal of Emergency Medicine
African Journal and Primary Health Care and Family Medicine

American Journal of Emergency Medicine

American Journal of Epidemiology

American Journal of Public Health

Anaesthesia

Annals of Emergency Medicine

Australasian Emergency Nursing Journal

BMC Emergency Medicine

BMC Health Services Research

BMJ Global Health

Bulletin of the World Health Organization

Canadian Journal of Emergency Medicine

Emergency Medicine Australasia

Emergency Medicine Journal

European Journal of Emergency Medicine

International Emergency Nursing

International Journal of Emergency Medicine

International Journal of Nursing Studies

Journal of Emergency Nursing

Pan African Medical Journal

PLoS ONE

South African Medical Journal

The bibliographies of all articles reviewed were also searched for additional relevant studies.

Triage

The best-known reports date the process of triage to between 1792 and 1801. Development of this practice was spearheaded by Dominique Jean Larrey, a French military surgeon serving under Napoleon Bonaparte, and since then, triage has come to have three distinct areas of practice:

1. triage in the field

2. triage to provide effective medical care following a disaster, and
3. triage to provide effective medical care in the emergency centre (EC). [23]

Triage in the field, and triage following a disaster, usually refers to pre-hospital triage. The latter can be done quickly by staff with minimal skills and is usually conducted in response to a mass casualty incident. Triage following a disaster is often in the context of too many patients for the resources available and is used to not only differentiate those who are seen first but includes those that may not receive care at all. It is not necessarily used to direct care at the most critically injured but rather towards those most likely to survive.[24] There are a number of primary and secondary triage tools that have been developed for the pre-hospital setting; these include: Simple Treatment and Rapid Transport, JumpSTART, Care Flight Triage, Triage Sieve, Sacco Triage method, Secondary Assessment of Victim Endpoint and Paediatric Triage tape among others.[24] These are basic three-level systems designed to be used only in a pre-hospital setting. They usually require no equipment and are structured to allocate patients to one of three categories, namely: immediate care, urgent care or delayed care. These categories are normally colour coded to ensure ease of identification of each group. Red, yellow and green are most often used to denote immediate, urgent and delayed care, respectively. The triage systems ask basic broad questions and require basic yes/no answers after each question. For example, “Can the patient walk?” - If yes, the patient is classified as delayed care (green), or “Is the patient breathing?” - If no, then open airway, if breathing after this then patient is classified as immediate care (red).

In-hospital triage scales have developed from these pre-hospital triage systems, and comprise three-, four- or five-level scales. Triage scales in this setting is often used to identify time sensitive illnesses.[25] Typically, the more levels a scale has, the more nuanced it can be in differentiating between patient acuity (i.e., urgency of care) and resource needs. The trade-off however is in the complexity and skill needed to administer the scale accurately. Triage systems involving a higher number of acuity categories usually require greater skill and experience by the staff utilising them. Travers et al. report that five-level scales demonstrate better reliability, improved sensitivity and specificity, and a greater ability to discriminate between patients, than a three-level scale.[26] The ideal triage scale is one that is able to

accurately and reliably differentiate between patients, and is simultaneously simple and easy to implement by staff who have minimal training.

For any given triage scale, each acuity level denotes a specific clinical urgency for care and comes with a recommended 'time to treat' specification. Some scales include a '*dead on arrival' category*, which can lead to some confusion when deciding whether the scale is a four or five level scale.

The SATS is a four-level triage scale, but has five overall levels as follows:

1. Emergency (patient to be seen immediately),
2. Very Urgent (patient to be seen within 10 minutes),
3. Urgent (patient to be seen within 60 minutes),
4. Routine (patient to be seen within 4 hours) and
5. Dead on Arrival (certification of death to be issued within 2 hours of patient being seen)

Whereas the CTAS, ATS, MTS and ESI are all five-level scales and do not include a Dead on Arrival level. The CTAS for instance includes these five levels:

1. Resuscitation (immediate),
2. Emergency (within 15 minutes),
3. Urgent (within 30 minutes),
4. Less Urgent (within 60 minutes), and
5. Non Urgent (within 120 minutes)

Note the CTAS does not have a *dead on arrival* category as part of the scale.

Triage Scales available in high-resource hospital settings

Emergency hospital-based triage systems are still relatively new. Over the last 30 years, four major triage systems have been extensively researched – all from high-income settings.[27] All four are five-level categorical measurement scales and include the ATS (previously known as the National Triage Scale) which was implemented in 1993 and revised and renamed ATS in 2000[12]; the CTAS developed in the late 1990s and revised in 2004 and again in 2008[28];

the MTS developed in 1994 and widely used across the United Kingdom and Europe[29], and the ESI developed in 1998 for the United States of America to replace the three-level triage scales widely used at the time.[30] These are all complex five-level triage scales that have demonstrated good performance results in the countries in which they were developed. The MTS, ESI and CTAS have been implemented outside of their home countries but have shown lower performance measures in Spain, Ireland and Taiwan.[9,31,32] We have not been able to identify any studies on the implementation of these triage scales in low-resource settings. This is not surprising as these triage scales are complex in structure and practically unfeasible in the context of low-resource settings. For example, the MTS consists of 52 flow charts that nurses need to understand and be able to apply to each patient scenario[13]; the ESI requires the resources that might be used by each patient, to be counted and assigned.[14] These triage scales require a high level of staff training and need highly qualified staff to use them reliably.

Other triage scales exist but have been studied much less. These include the Taiwan Triage Acuity Scale (TTAS) which was originally constructed as a four-level scale and later revised into a five-level scale in 2006 based on the CTAS[28,33]; the Rapid Emergency Triage and Treatment System (RETTS) developed in 2006 and revised in 2010, RETTS was implemented in Sweden, Norway and Denmark[34–37]; the Italian Triage Emergency Method (TEM)[38]; and the French Emergency Nurses Classification in Hospital Scale (FRENCH)[39].

Triage scales available in low-resource hospital settings

Over the last 30 years, few triage scales have been developed for use in low-resource settings.[40,41] Triage has been identified as part of an essential package of emergency care, and yet across low-resource settings very few hospitals use a formal triage scale.[1,42–44]
In 1999, the WHO developed the ETAT system to address high paediatric mortality. The ETAT aimed to improve triage for all sick children arriving at outpatient clinics in low-resource settings. Paediatricians believed that improved triage would significantly reduce morbidity and mortality in infants and young children.[45] The ETAT has subsequently been implemented widely across low-resource settings.[46,47] One major drawback, however, is the lack of adult and older child triage scales; ETAT only allows for triage of children under

five years of age. This has limited its used as a triage scale in hospital-based settings where a patient mix of adults and children is usual. The use of two different triage scales has the potential to cause confusion and errors.

In 1996 the Kampala Trauma Score (KTS) was developed as a severity scoring system for trauma registries in Uganda.[48] It was theorised that it might serve as a triage scale in low-resource settings too. However, a study by Kobusingye et al. showed its use as a triage scale was limited given the low sensitivity and specificity across two hospitals in Uganda.[49] In addition to this limitation, it is only designed for trauma patients and therefore has limited utility in EC settings where the trauma load is between 8-36% of emergency cases.[40,50]

The South African Triage Scale was developed in 2004 specifically for low-resource South African settings.[51] At the start of this thesis, it was the only triage scale that existed for the triage of both adults and children in a low-resource setting, and which showed good reliability and validity.

The South African Triage Scale

The SATS was previously known as the Cape Triage Score and was designed to be used by enrolled nurses (i.e. entry level nursing staff possessing a basic one-year Diploma). A scale that could reliably be carried out by enrolled nurses would be ideal, given the paucity of highly-trained healthcare workers.[21,52] The SATS uses a physiologically based scoring system, namely, the Triage Early Warning Score (TEWS) and it also uses a list of discriminators designed to triage patients into one of five colour coded priority groups for medical attention. The colour categories are as follows: i) red – immediate priority; ii) orange – very urgent priority; iii) yellow – urgent priority; iv) green – delayed priority (minor injuries/illness); and v) blue – dead (in some contexts the colour black is used). The SATS 2008 scale consists of three age specific charts (i.e. adult, child and infant; Figures 1-3, respectively) whereas the 2012 revision consists of two age specific charts (i.e. adult and paediatric; Figures 4 and 5, respectively).

ADULT TRIAGE SCORE								©South African Triage Group 2008
	3	2	1	0	1	2	3	
Mobility				Walking	With Help	Stretcher/ Immobile		Mobility
RR		less than 9		9-14	15-20	21-29	more than 29	RR
HR		less than 41	41-50	51-100	101-110	111-129	more than 129	HR
SBP	less than 71	71–80	81-100	101-199		more than 199		SBP
Temp		Cold OR Under 35		35-38.4		Hot OR Over 38.4		Temp
AVPU		Confused		Alert	Reacts to Voice	Reacts to Pain	Unresponsive	AVPU
Trauma				No	Yes			Trauma
over 12 years / taller than 150cm								

Colour	RED	ORANGE	YELLOW	GREEN	BLUE
TEWS	7 or more	5-6	3-4	0-2	DEAD
Target time to treat	Immediate	less than 10 mins	less than 60 mins	less than 240 mins	DEAD
Mechanism of injury		High energy transfer			
Presentation		Shortness of breath - acute		ALL OTHER PATIENTS	
		Coughing blood			
		Chest pain			
		Haemorrhage - uncontrolled	Haemorrhage - controlled		
	Seizure - current	Seizure - post ictal			
		Focal neurology - acute			
		Level of consciousness reduced			
		Psychosis / Aggression			
		Threatened limb			
		Dislocation - other joint	Dislocation - finger or toe		
		Fracture - compound	Fracture - closed		
	Burn – face / inhalation	Burn over 20%	Burn - other		
		Burn - electrical			
		Burn - circumferential			
		Burn - chemical			
		Poisoning / Overdose	Abdominal pain		
	Hypoglycaemia - glucose less than 55	Diabetic - glucose over 200 & ketonuria	Diabetic - glucose over 305 (no ketonuria)		
		Vomiting - fresh blood	Vomiting - persistent		
		Pregnancy & abdominal trauma or pain	Pregnancy & trauma		
			Pregnancy & PV bleed		
Pain		Severe	Moderate	Mild	
	Senior Healthcare Professional's Discretion				

Figure 1: South African Triage Scale, Adult 2008

CHILD TRIAGE SCORE								© South African Triage Group 2008
	3	2	1	0	1	2	3	
Mobility				Walking	With Help	Stretcher/ Immobile		Mobility
RR	less than 15	15-16		17-21	22-26	27 or more		RR
HR	less than 60	60-79		80-99	100-129	130 or more		HR
Temp		Cold OR Under 35		35-38.4		Hot OR Over 38.4		Temp
AVPU		Confused		Alert	Reacts to Voice	Reacts to Pain	Unresponsive	AVPU
Trauma				No	Yes			Trauma
3 to 12 years old / 96 to 150 cm tall								

COLOUR	RED	ORANGE	YELLOW	GREEN	BLUE
TEWS	7 or more	5-6	3-4	0-2	DEAD
Target time to treat	Immediate	less than 10 mins	less than 60 mins	less than 240 mins	DEAD
Mechanism of injury		High energy transfer			
Presentation	Drooling	Shortness of breath		ALL OTHER PATIENTS	
		Stridor			
		Wheeze			
		Haemorrhage - uncontrolled	Haemorrhage - controlled		
	Seizure - current	Seizure - post ictal			
		Focal neurology - acute			
		Level of consciousness reduced			
		Exhaustion			
		Purpura			
		Dislocation - other joint	Dislocation - finger or toe		
		Fracture - compound	Fracture - closed		
	Burn – face / inhalation	Burn over 10%	Burn - other		
		Burn - electrical			
		Burn - circumferential			
		Burn - chemical			
		Poisoning / Overdose	Abdominal pain		
	Hypoglycaemia – glucose less than 55	Diabetic - glucose over 200 & ketonuria	Diabetic - glucose over 305 (no ketonuria)		
		Dehydration	Vomiting - persistent		
		PR bleeding	Inappropriate history		
Pain		Severe	Moderate	Mild	
	Senior Healthcare Professional's Discretion				

Figure 2: South African Triage Scale, Child 2008

INFANT TRIAGE SCORE								© South African Triage Group 2008
	3	2	1	0	1	2	3	
Mobility				Normal for age		Stretcher/ Immobile		Mobility
RR	less than 20	20-25		26-39		40-49	50 or more	RR
HR	less than 70	70-79		80-130		131-159	160 or more	HR
Temp		Cold OR Under 35		35-38.4		Hot OR Over 38.4		Temp
AVPU				Alert	Reacts to Voice	Reacts to Pain	Unresponsive	AVPU
Trauma				No	Yes			Trauma
younger than 3 years / smaller than 95cm								

COLOUR	RED	ORANGE	YELLOW	GREEN	BLUE
TEWS	7 or more	5-6	3-4	0-2	DEAD
Target time to treat	Immediate	less than 10 mins	less than 60 mins	less than 240 mins	DEAD
Mechanism of injury		High energy transfer			
Presentation	Drooling	Shortness of breath		ALL OTHER PATIENTS	
	Stridor	Wheeze			
		Haemorrhage - uncontrolled	Haemorrhage - controlled		
	Seizure - current	Seizure - post ictal			
		Focal neurology - acute			
		Level of consciousness reduced			
		Floppy infant			
		Purpura			
		Dislocation - other joint	Dislocation - finger or toe		
		Fracture - compound	Fracture - closed		
			Unable to weight bear		
	Burn – face / inhalation	Burn over 10%	Burn - other		
		Burn - electrical			
		Burn - circumferential			
		Burn - chemical			
		Poisoning / Overdose	Abdominal pain		
	Hypoglycaemia - glucose less than 55				
		Dehydration	Vomiting - persistent		
			Not feeding		
			Not urinating		
		PR bleeding	Inappropriate history		
			Prolonged or uninterrupted crying		
Pain		Severe	Moderate	Mild	
	Senior Healthcare Professional's Discretion				

Figure 3: South African Triage Scale, Infant 2008

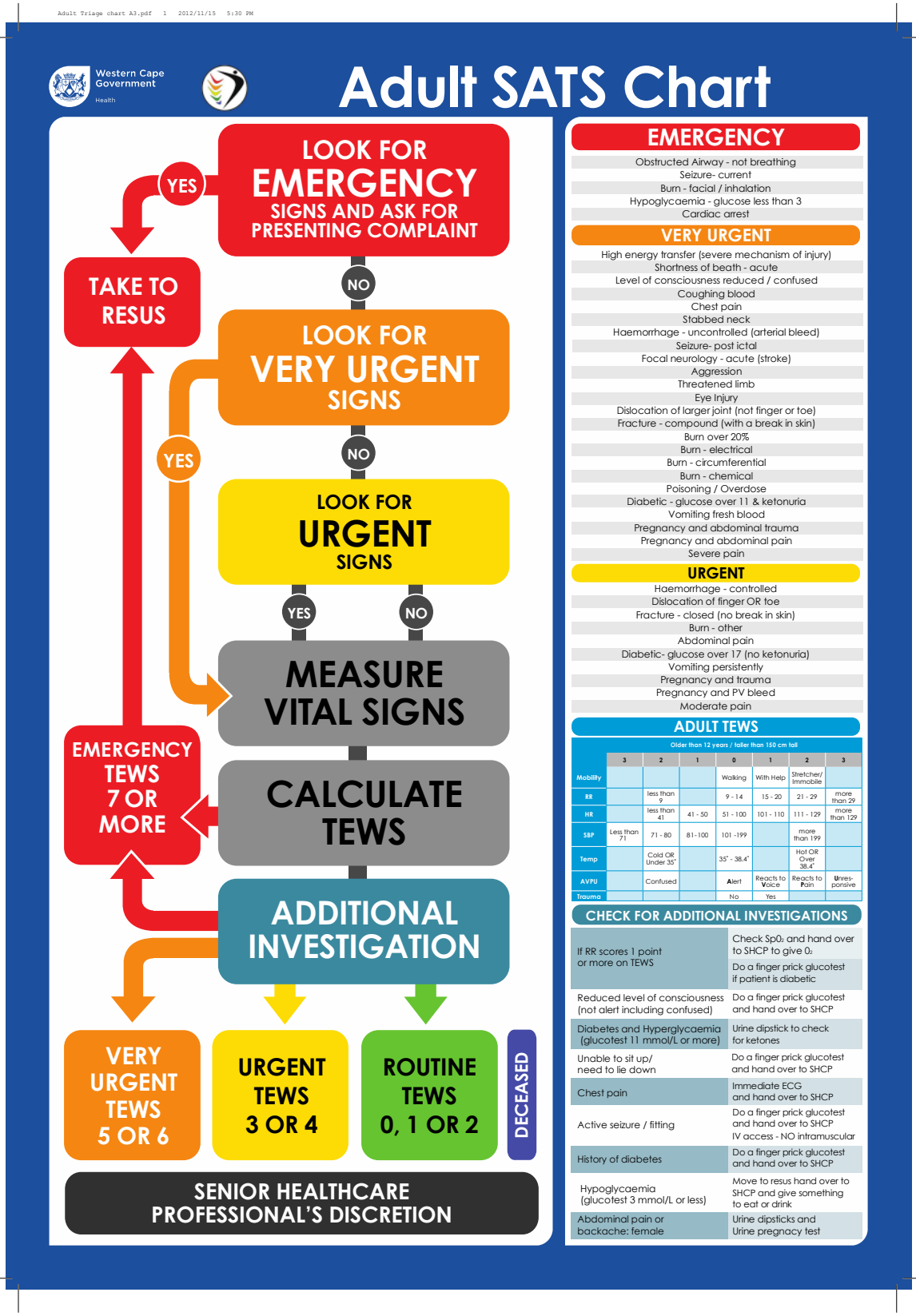


Figure 4: South African Triage Scale, Adult 2012

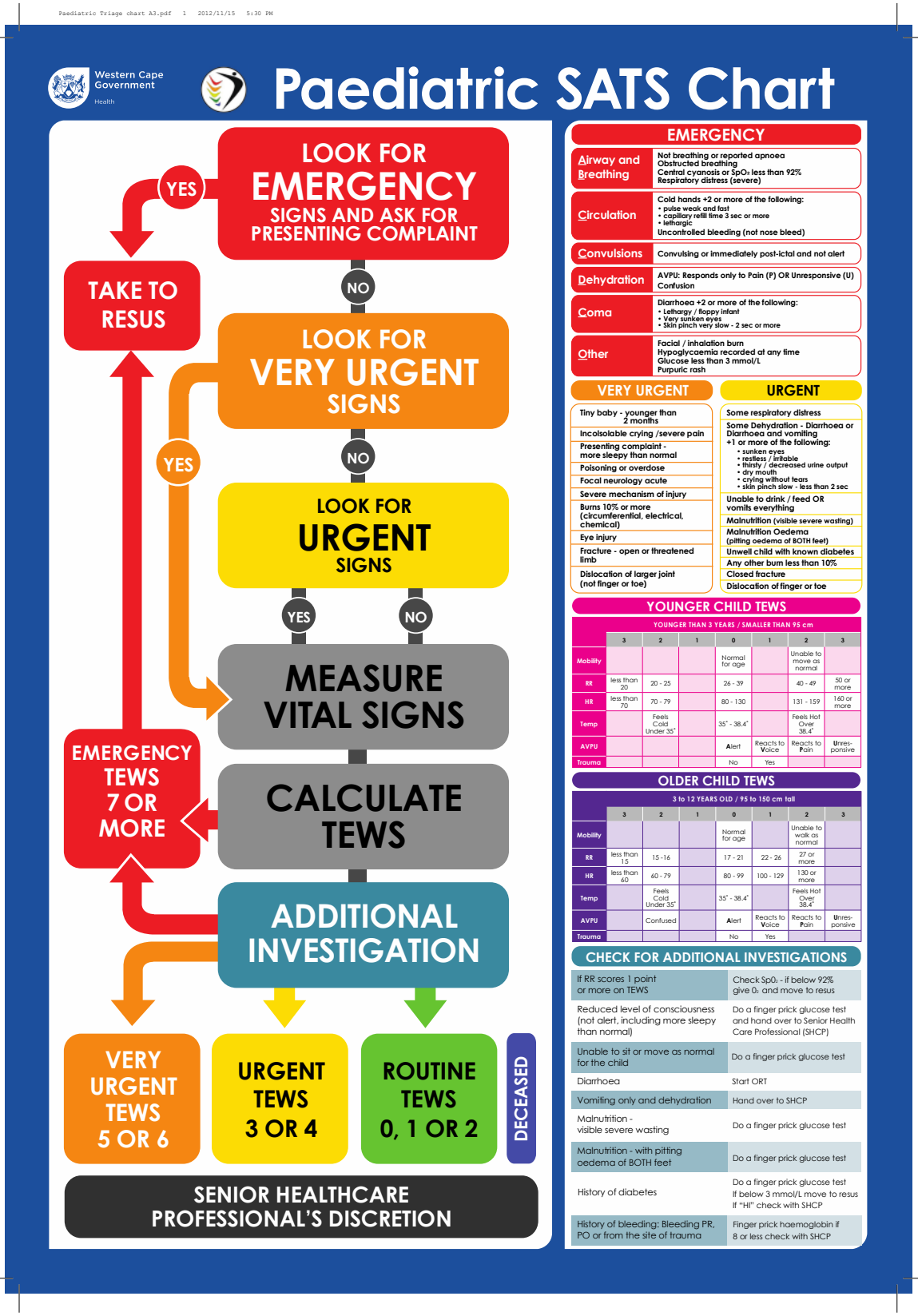


Figure 5: South African Triage Scale, Paediatric 2012

The SATS began to be implemented across South Africa in 2006, and was revised in 2008 and 2012.[40,50,51] It has been extensively validated across South Africa and other sub-Saharan settings but has yet to be evaluated elsewhere. In South Africa, it has demonstrated good reliability and validity when used by enrolled nurses[18–21,53,54]; it was for these reasons that MSF chose to implement the SATS in their ECs.

Reliability of the SATS

Reliability allows us to reflect on the amount of error inherent in a measurement. Many terms have been used to describe reliability in the literature including 'agreement', 'reproducibility' and 'consistency'.[55,56] Use of these different terms however can cause confusion, lead to erroneous conclusions and can limit the comparability of studies. The formal definition of reliability is 'subject variability' divided by 'subject variability plus measurement error'. This definition stems from Classical Test Theory, and should essentially help test whether the scale can differentiate between patients.[57] Reliability can be defined as the ability of a scale to reproduce the same results either when used by different raters or when used by the same rater over time. Inter-rater reliability refers to how well a scale yields the same result when used by different raters assessing the same subject; intra-rater reliability refers to how well a scale yields the same result when the same rater assesses the same subject at different points in time.[58]

Common measures used in reliability

Reliability is an important performance indicator for a triage scale, and many studies assess it separately from validity. Reliability in triage literature is commonly reported on using either the kappa coefficient or the intraclass correlation coefficient (ICC).[59,60] The kappa measurement can be either unweighted or weighted; most authors report on the weighted kappa. The unweighted kappa only corrects the observed percentage between raters for the effects of chance and does not take into account the degree of disagreement across the scale. Linear weighted and quadratically weighted kappa (QWK) coefficients were introduced to take account of the magnitude of any disagreement. Kappa statistics – weighted and unweighted – are associated with a number of limitations:

1) They are dependent on the number of categories in the ordinal data. This means five level triage scales have a higher weighted kappa estimate than a three level scale in the presence of the same agreement[59];
2) They are dependent on the distribution of cases[61]: the wider the distribution of cases, the lower the kappa statistic; and
3) They are insensitive to differences in agreement for different ordinal values.

These limitations can make it difficult to interpret kappa estimates and do not always allow unambiguous interpretations on the different classification models of Landis and Koch, Fleiss and Cicchetti and Sparrow.[52,62] Even though the weighted kappa takes into account the magnitude of disagreement, some have suggested that the QWK does not truly reflect clinical urgency. Wulp et al. have proposed a new type of kappa altogether, instead of the linear weighted or the QWK, which is commonly used. Wulp argues that the QWK with equal weightings could misrepresent actual clinical outcome. When using the QWK, a mis-triage of two categories counts as 75% exact agreement; in real life however, if a patient is triaged as non-urgent, when in fact they are an emergency case, the ramifications could be deadly (e.g., death or serious disability). The current QWK does not account for this and as such a new weighting system has been proposed. The new weighting system accommodates for one category difference and seriously penalises for a difference of more than one category.[58] This is all very well, but in the absence of a gold standard and variations in the use of kappa estimates that use different weighting systems, it is impossible to compare reliability results from different studies.

The ICC is the other statistic used in triage research. There are three types of ICC defined by Shrout and Fleiss:

- Type 1: assumes that a different set of raters is used for the assessment of each vignette;
- Type 2: assumes that the raters are a random sample from a population of raters; and
- Type 3: assumes a given number of raters are the only raters of interest for the reliability of these ratings.[63]

The ICC can be used to assess either:

1. the reliability of a single rater's ratings (among a group of *k* raters), or
2. the reliability of the mean rating of the *k* raters (i.e. what is the reliability of a rating that is calculated by averaging the ratings across raters).

Berk has identified 11 reasons why the ICC is superior to the kappa when assessing reliability. Some of these reasons include i) its ability to isolate factors affecting reliability, ii) its flexibility to analyse more than two raters and more than two responses, iii) its ability to include and exclude systematic differences between raters as part of the error term, iv) its ability to handle missing data, and iii) its ability to provide a unifying framework that links together different ways of measuring inter-rater agreement. [56,64]

The QWK has been shown to be equivalent to the ICC type 2 in terms of results obtained through the different calculations.[56] Still, there is no universally accepted coefficient for the measurement of reliability, and this renders the comparison of different study results difficult. For this reason, the Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were published. These guidelines set out specific measures that should be reported on in the context of reliability testing, and their main purpose is to allow comparison of results across related studies.[57]

Practical assessments of reliability in triage

Studies assessing the reliability of a triage system – those conducted in both low- and high-resource settings – have often used paper-based vignettes (e.g., short, written clinical scenarios descriptive of a patient's presenting complaint, including additional simple investigations and basic vital signs). Some studies in high resource settings have used parallel research nurses to assess reliability.[65,66] This, however, is not feasible in low-resource settings. Paper-based vignettes, although limited in their ability to convey audio-visual cues, are accepted by most triage researchers as being an acceptable way of presenting the relevant case details of a patient and allowing that patient's respective acuity level to be determined.[13,38,67,68] Most studies use local EC cases as a basis for the vignettes. Generalisability of vignettes across settings poses a significant limitation. Twomey et al.

proposed the use of a set of standard adult reference vignettes for use in the assessment of reliability and validity.[69] However, no such vignettes exist in the area of paediatrics.

The SATS had been shown to be a reliable triage scale in South Africa and Botswana at the start of this thesis[59,70], but its reliability in other contexts had not been assessed.

Validity of the SATS

Validity in the context of a triage scale, measures a scale's ability to make accurate inferences about a patient's true acuity level.[56] Three common types of validity exist, namely, content, criterion and construct validity. Criterion validity can be divided into concurrent and predictive validity.

1. Content validity – is not based on the scores of a scale but only on the judgement of experts regarding content of the items.
2a. Concurrent validity – most often used when trying to replace one tool with another one.
2b. Predictive validity – used when developing instruments that allow us to get answers earlier than current instruments allow.
3. Construct validity – used when trying to measure a hypothetical construct rather than something that can readily be observed.[56]

In triage validity research, given the lack of an agreeable gold standard, two methods are commonly accepted:

1. Use of expert opinion: Paper-based vignettes are most commonly used. However, some studies have used real-time, simultaneous side-by-side triage with an expert. With simulated paper-based cases, vignettes are assessed by 'triage experts' and assigned a rating (i.e., the expert gold standard). Thereafter, triage officers independently triage the vignettes under classroom conditions. The officers' ratings are then compared to the expert gold standard, and measures of accuracy (i.e., over-triage and under-triage) are calculated.[21,71–73]

2. Use of EC outcomes as surrogate markers for patient acuity: hospital admission, mortality, length of hospital stay, and resource utilisation are commonly used as predictive markers for true patient acuity. The assumption is that patients with a high triage acuity are more likely to be admitted to hospital, more likely to die, more likely to have a longer hospital stay and more likely to require greater resources than patients with lower triage acuities.[18,20,33,74,75]

It can be argued that both these methods are a form of construct validity as they both refer to a hypothetical construct. However, there is some confusion on this: Twomey et al. referred to the use of outcome markers as a form of predictive validity whereas Wulp et al. referred to them as a form of construct validity.[22,58] Streiner has mentioned this confusion and suggested that we stay away from trying to label the type of validity but instead clearly describe each construct used when dealing with validity studies. This helps future researchers to understand how they went about validity of the tool rather than what name they called it.[56]

Measures of validity

Validity can be measured using the two methods as described above, however, interpretation of results are not always straight-forward.

When using paper-based vignettes, researchers compare the results of their raters to that of an expert opinion (i.e., a proxy gold standard). From these results, they commonly work out over- and under-triage. Many use the American College of Surgeons - Committee on Trauma (ACS-COT) guidelines of acceptable over- and under-triage levels. ACS-COT states that 30-50% over-triage and 5-10% under-triage are acceptable rates.[21] This was revised in 2014 to 25-35% over-triage and 5% under-triage.[76] ACS-COT is based on trauma pre-hospital patients only and designed for high-resource settings. The feasibility of these benchmarks was proven to be unreasonable with the current resources in Pennsylvania, America.[77] Twomey et al. have argued that it does not represent appropriate measures for over- and under-triage in a low-resource setting.[21] The implications of such a high over-triage rate is significant for low resource settings. Such a high number of over-triaged patients results in a significant strain on already limited resources with little or no clinical benefit for patients. Some validity studies

have moved away from using ASC-COT as a reference range and instead use Odds Ratio of admission across acuity levels.[58] We explore this further in Chapter 7.

When using outcome measures it becomes difficult to represent it as over- and under-triage. Some studies have tried this by grouping triage categories together using the outcome markers of death, hospitalisation or discharge as the proxy gold standard. The expectation is that the outcome for high acuity patients should either be death or admission whereas the outcome for low-acuity patients should be discharge. Studies differ in how they stratify high-acuity and low-acuity patients (e.g. Rosedale et al. stratified Emergency (Red) and Very urgent (Orange) patients into the high acuity category with routine (Green) patients as low acuity).[18] They ignored the urgent (Yellow) patients in their analysis. In their validation of the 2012 SATS paediatric revision, Twomey et al. stratified Emergency, Very Urgent and Urgent as high acuity with Routine representing low-acuity.[20] These stratifications can dramatically affect the results of a study. The need for a standard or novel way for assessing validation is needed when using over- and under-triage with surrogate markers.

Other researchers have used the correlation of hospitalisation across triage acuities as a factor to validate triage scales.[11,33] However, no guidelines exist on how strong this correlation has to be before validity can be confirmed.

SATS was validated in South Africa using both paper-based cases and surrogate marker methods.[20,21] Twomey et al. has suggested the use of reference vignettes to improve the ability to test validity in low-resource settings.[69] These would be vignettes that have gone through a consensus process and would be the reference gold standard for future studies. This would be an easy, cost-effective way to assess validity going forward. At the beginning of this thesis the reference vignettes only assessed adult cases and no standardised paediatric cases existed.[69]

Summary

In this literature review, we explored the various triage scales available both in high- and low-resource settings. We found that triage scales designed for use in high-resource settings were not appropriate for use in low-resource settings given their complex nature and the high skill level associated with their use. In low-resource settings, we found very few triage scales with only one addressing both adults and children. We found that the SATS was reliable and valid

in South Africa but found that information across a wide range of low-resource settings was missing. Assessing validity was a challenge in low-resource settings and there was a lack of standardised paediatric vignettes.

As MSF was implementing the SATS across the different contexts in which it operates, the fundamental question that arose was; is the SATS a reliable and valid scale for use in low-resource settings outside sub-Saharan Africa?

Chapter 3: Implementation and accuracy of the South African Triage Scale in a low-resource setting

Reference

Dalwai MK, Tayler-Smith K, Trelles M, Jemmy JP, Maikéré J, Twomey M, Wakeel M, Iqbal M, Zackariah R. Implementation of a triage score system in an emergency room in Timergara, Pakistan. PHA. DOI: 10.5588/pha.12.0083

Declaration from Author

The following co-authors contributed to the paper:

Mrs. Katie Tayler-Smith, Dr. Miguel Trelles, Dr. Jean-Paul Jemmy, Dr. Jacob Maikéré, Dr. Michele Twomey, Dr. Muhamed Wakeel, Dr. Mohamed Iqbal and Dr. Rony Zackariah

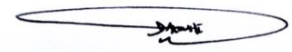
In the case of Chapter 3, the nature and extent of contribution by the authors was as follows:

Nature of contribution

MKD, KTS and MT contributed to study conception and design; MKD, MT, JPJ, JM, MW and MI facilitated the acquisition of data; MKD, MT, RZ and MI helped with the analysis and interpretation of the data; the initial version was drafted by MKD and KTS. All authors contributed to the redrafts and revised it critically for important intellectual content. All authors gave final approval on the version submitted and the revisions thereof.

Extent of contribution

MKD: 80%; KTS, MT, JPJ and RZ together: 10%; JM, MT, MW and MI together: 10%



Signed: Mohammed. K. Dalwai 27 Dec 2017

Declaration from Co-Authors

The undersigned hereby certify that:

1. The above declaration correctly reflects the nature and extent of the candidate's contribution to this work, and the nature of the contribution of each of the co-authors.
2. They meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
3. They take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
4. There are no other authors of the publication according these criteria;
5. Potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and
6. The original data are stored in a two factor password protected cloud account and will be held for at least five years.

Prof. Lee Wallis 10 Jan 2018

Dr Michele Twomey 10 Jan 2018

Mrs Katie Tayler-Smith 10 Jan 2018

Introduction to the topic

Few triage scales exist for low-resource settings. As discussed in Chapter 2, the SATS was developed for use in ECs with limited resources. The SATS has been widely implemented in sub-Saharan settings but at the time of writing nothing was available in the literature about its implementation elsewhere.

Motivation for conducting the study

MSF has been working in Timergara since 2009 in collaboration with the Ministry of Health to help strengthen the EC. At the outset, no formal triage system existed, and with a very high emergency patient case-load and overstretched staff, critically ill patients were often missed or seen only when their clinical condition had deteriorated.

Aim

This study aimed to describe the implementation of the SATS in a low-resource setting outside sub-Saharan Africa.

Objectives

- Describe the implementation of the SATS in Timergara hospital
- Report on the accuracy of the SATS post implementation

Main findings

- Implementation of the SATS in a low-resource, high conflict setting requires buy-in from all stakeholders involved (beyond just patients and healthcare workers)
- The SATS can be used accurately in a low-resource setting

Copy of published paper

Public Health Action
International Union Against Tuberculosis and Lung Disease
Health solutions for the poor
VOL 3 NO 1 PUBLISHED 21 MARCH 2013

SHORT COMMUNICATION

Implementation of a triage score system in an emergency room in Timergara, Pakistan

M. K. Dalwai,[1,2] K. Tayler-Smith,[3] M. Trelles,[4] J-P. Jemmy,[4] J. Maikéré,[1] M. Twomey,[2] M. Wakeel,[5] M. Iqbal,[1] R. Zachariah[3]

http://dx.doi.org/10.5588/pha.12.0083

Following implementation of the South African Triage Scale (SATS) system in the emergency department (ED) at the District Headquarter Hospital in Timergara, Pakistan, we 1) describe the implementation process, and 2) report on how accurately emergency staff used the system. Of the 370 triage forms evaluated, 320 (86%) were completed without errors, resulting in the correct triage priority being assigned. Fifty completed forms displayed errors, but only 16 (4%) resulted in an incorrect triage priority being assigned. This experience shows that the SATS can be implemented successfully and used accurately by nurses in an ED in Pakistan.

Emergency triage is the systematic process of determining patients' priority for treatment based on the severity of their condition. The principle aim of triage is to ensure that patients receive the most appropriate level and quality of care relative to their clinical status and need.[1]

Since 2009, Médecins Sans Frontières (MSF) has been working in Pakistan in collaboration with the Ministry of Health (MoH) to strengthen emergency department (ED) care at the District Headquarter (DHQ) Hospital in Timergara in the North-Western Province of Khyber Pakthunkhwa. At the outset, no formal triage system existed, and with a very high emergency patient caseload and overstretched staff, critically ill patients were often missed or seen only when their clinical condition had deteriorated. Anecdotal evidence suggested that ED mortality rates were very high. The need for a triage system was identified with reports from other resource-poor settings showing that the process of triage can improve patient flow and reduce patient waiting times and mortality.[2,3]

As one of the only validated triage tools that exists for the triage of adults specifically in resource-poor settings, we chose to pilot the South African Triage Scale (SATS)[4,5] in Pakistan. The SATS was devised for use in both pre- and in-hospital emergency units throughout South Africa to address the increasing health burden on emergency care services in the country, in particular the high numbers of severe emergency cases, medical staff shortages and limited resources.[5] Implementing the SATS in Pakistan was justified by the fact that Pakistan and South Africa are both developing countries facing similar challenges in the delivery of emergency medical care, and both are faced with high caseloads of trauma patients (66 trauma presentations per 1000 patients in South Africa and 41/1000 in Pakistan).[6,7] The SATS system has been validated in South Africa and has been implemented in various other settings such as Malawi,[8] Botswana[9] and Ireland.[10] However, its use in an Asian setting has not yet been formally examined.

Based on our experience in Pakistan, we 1) describe how the SATS was implemented, 2) report on how accurately emergency staff were able to use the system, and 3) discuss the lessons learnt.

METHODS

Design

A cross-sectional study involving an audit of routine clinical data.

Setting and population

Timergara DHQ Hospital is situated in Timergara town, in the rural and relatively insecure district of Lower Dir, Pakistan. The hospital serves a catchment population of approximately 1.8 million. The ED has a capacity of 18 beds and a monthly caseload of about 4000 patients. The main caseload is medical emergencies, typically cardiac and diabetes-related illnesses, with a proportion of trauma cases (mostly due to road traffic accidents).

The study involved the cross-checking of a series of triage forms completed for patients triaged at the ED between 27 June and 3 July 2011.

SATS and its implementation

The SATS uses a physiologically based scoring system, the Triage Early Warning Score (TEWS), and a list of discriminators designed to triage patients into one of five colour-coded priority groups for medical attention.[5] The colour categories are as follows: 1) red, immediate priority; 2) orange, very urgent priority; 3) yellow, urgent priority; 4) green, delayed priority (minor injuries/illness); and 5) black, dead.

The SATS was implemented in the Timergara ED in June 2011. Table 1 indicates the measures and activities undertaken 1) before the introduction of the SATS and 2) during the month following its introduction.

Once the SATS was introduced, triage was routinely undertaken by two triage nurses, supervised by the local doctor. Post triage, patients were directed

AFFILIATIONS
1 Médecins Sans Frontières (MSF), Islamabad, Pakistan
2 Emergency Medicine, University of Cape Town, Cape Town, South Africa
3 Medical Department, Operational Research, Operational Centre Brussels, MSF–Luxembourg, Luxembourg
4 MSF, Brussels Operational Centre, Brussels, Belgium
5 Ministry of Health, Timergara, Lower Dir, Pakistan

CORRESPONDENCE
Katie Tayler-Smith
Operational Centre Brussels
Medical Department
Operational Research
Médecins Sans Frontières
rue Dupré 94
1090 Brussels, Belgium
Tel: (+32) 352 335 133
Fax: (+32) 352 332 515
e-mail: katie_harries@yahoo.co.uk

ACKNOWLEDGEMENTS
The authors are grateful to the Pakistani Ministry of Health for their collaboration; they are particularly grateful to the staff in the field for their hard work. Conflict of interest: none declared.

KEY WORDS
emergency triage; South African Triage Scale; Pakistan

Received 6 November 2012
Accepted 10 December 2012

PHA 2013; 3(1): 43–45

TABLE 1 Measures and activities undertaken before and after the SATS was introduced in Timergara Emergency Department, Pakistan

1 Before introduction of the SATS

Approval*

- Extensive discussions with the following hospital bodies and authorities to obtain permission to introduce the SATS and carry out the related ED care delivery changes (April–June 2011)
  - Hospital unions
  - Hospital superintendent
  - District health authorities (Ministry of Health)

Information dissemination

- Community awareness raising and sensitisation through discussions with religious leaders about how entry into the ED would change, why it would change and what the ED triage process would involve (May 2011)
- Dissemination of written information in the local language (leaflets and billboards) around the hospital for both staff and patients explaining the planned changes to the ED and the triage process (May 2011)
- Briefing of hospital staff about the proposed ED changes and triage, including security guards and cleaners, as these are the first point of contact for many patients coming to the hospital (May 2011)

Infrastructure

- Erection of a dedicated triage room outside the ED; patients were only able to enter the ED from the triage room, ensuring that the triage post was the first point of contact for patients presenting at the ED; previously patients could enter the ED via multiple points

Training

- Training in the SATS of the local doctor (2 × 2h sessions with one practical session), conducted by the medical doctor in charge, who had background experience of working with the tool and who received collaborative support from the chief implementation officer of the SATS from the Western Cape in South Africa (May 2011)
- One-hour nurse training sessions in the SATS conducted by the local doctor (May 2011)

2 During the month following the introduction of the SATS

- Measures put in place to ensure patient privacy, particularly for women
  - Outside the triage room, two waiting areas established, one for men and one for women
  - Guards permitting just two patients and two care givers into the triage room at a time
  - Curtains dividing the triage room into single patient areas
- Ongoing support by the medical doctor in charge for the first 2 weeks of triage
- Supervision of staff nurses carrying out the triage by the local doctor

*Due to complex political influences within the hospital, it was important to seek permission from the different authorities in the correct hierarchical sequence for a favourable agreement to be reached. These permissions were later captured in a written Memorandum of Understanding.
SATS = South African Triage Scale; ED = emergency department.

to the appropriate care posts depending on their designated colour code.

Data collection and analysis

One month after the introduction of the SATS, during which various process issues, such as ensuring efficient patient flow into the ED triage room, were able to be properly established, an evaluation of how accurately staff were using the system was undertaken. Over the course of 1 week, all patient triage forms available at the different triage posts were collected for evaluation. An evaluation period of 1 week was deemed sufficient for generating a large enough sample of triage forms. The medical doctor in charge, with no direct involvement in the day-to-day triage activities, collected and checked the forms. Nurses involved in completing the patient triage forms and the local doctor responsible for supervising them were blinded to the evaluation.

The triage forms were checked to determine whether the TEWS and identified discriminator conditions matched the final allocated triage colour. A 15% error rate is accepted in the SATS protocol, and this was thus used as the acceptable error threshold. Levels of under- and over-triage were also assessed and compared to the American College of Surgeons Committee on Trauma (ACSCOT) guidelines which accept threshold indicators for over-triage of up to 50% and up to 10% for under-triage.[11]

Ethical considerations

Due to the routine nature of the study data, the study was exempted from review by the Pakistani National Bioethics Committee. The study also fulfilled the criteria for analysis of routine data by the MSF Ethics Review Board.

RESULTS

Of 381 triage forms collected for this evaluation, 11 (3%) had triage information missing and were therefore excluded. Table 2 shows the accuracy of completion of the 370 forms evaluated. Overall, 320 (86%) had been completed without errors, resulting in the correct triage priority being assigned (the degree of correctness was highest for patients in the most urgent categories of red and orange). Of the 50 forms completed with errors, only 16 (4%) resulted in the incorrect triage priority being allocated: in 12 (3%) cases the patient was under-triaged and in the remaining 4 (1%) cases the patient was over-triaged.

A basic sensitivity analysis indicated that even if the 11 excluded triage forms had indicated an incorrect triage priority, the overall level of error would still have only been 7%, i.e., well below the accepted threshold of 15%.

DISCUSSION

This is the first time that the implementation of the SATS system for the triage of emergency cases has been studied in a resource-poor setting in Asia. The experience shows that the system is viable and can be implemented accurately in a district hospital in a rural district of Pakistan. Successful implementation of the SATS system in this particular setting was likely facilitated by a number of factors.

First, the tool itself was relatively easy to use. Our findings indicate that nurses were able to use the tool with a high degree of accuracy after just one short 1-hour training session each.

TABLE 2 Accuracy of completion of the SATS forms by staff at Timergara District Hospital, Pakistan (N = 381)*

	Accuracy of completion†			
Allocated triage colour	Correct *n* (%)	Incorrect but no change in triage colour *n* (%)	Incorrect and change in triage colour *n* (%)	Total *n* (%)
Green	77 (85)	8 (9)	6 (7)	91 (25)
Yellow	105 (81)	17 (13)	8 (6)	130 (35)
Orange	100 (91)	8 (7)	2 (2)	110 (30)
Red	36 (97)	1 (3)	0	37 (10)
Black	2 (100)	0	0	2 (0.5)
All	320 (86)	34 (9)	16 (4)	370 (100)

*11 triage forms had information missing and were therefore not included.
†A 15% error rate is accepted in the SATS protocol; this was thus used as the acceptable error threshold.
SATS = South African Triage Scale.

Second, understanding the decisional and management systems within the MoH and ensuring appropriate collaboration with key players were believed to be integral to the acceptance and implementation of the SATS. Engaging with hospital staff was also deemed important to foster their cooperation. Staff members were made aware of the fact that patients were dying unnecessarily in the ED due to delays in diagnosis and treatment, and the concept of a triage system was presented as a mechanism for better managing patient caseloads and improving patient care. This was intended to enhance ownership and responsibility.

Third, gaining community acceptance was considered important, especially for this very traditional setting, where segregation of the sexes is the cultural norm. By being sensitive to these cultural norms and incorporating them into the way the SATS was implemented, we hoped that the system would be less likely to be met by dissatisfaction from patients and their care givers.

Fourth, by creating open communication channels with key hospital staff and briefing other hospital staff such as security guards and cleaners about the changes, we felt that it would be possible to establish a more effective platform for transformation.

The implementation of the SATS in Timergara's ED did not require significant resources or investment. One of the main requirements was the availability of a staff member able to train other staff in the use of SATS. If capacity to train nurses in the use of the SATS was developed, we believe that the use of this triage tool would be feasible across routine care delivery systems in Pakistan and in other similar settings.

Finally, with triage deemed to be of low priority in the developing world, little research to date has been invested in this method of handling ED patients. Nonetheless, as emergency medicine becomes a more common feature of health care systems in developing countries, strategies need to be employed to ensure that limited, overstretched resources are used as effectively as possible, and that patient care and outcomes are optimised. Triage is one mechanism that can be used to reach these goals.[2,3] In conclusion, our experience shows that the SATS can be implemented successfully and used accurately by nurses in an ED in Pakistan.

References

1 Rogers I R, Evans L, Jelinek G A, et al. Using clinical indicators in emergency medicine: documenting performance improvements to justify increased resource allocation. J Accid Emerg Med 1999; 16: 319–321.
2 Molyneux E M, Ahmad S, Robertson A. Improved triage and emergency care for children reduces inpatient mortality in a resource-constrained setting. Bull World Health Organ 2006; 84: 314–319.
3 Bruijns S, Wallis L, Burch V. Effect of introduction of nurse triage on waiting times in a South African emergency department. Emerg Med J 2008; 25: 395–397.
4 Bruijns S R, Wallis L A, Burch V C. A prospective evaluation of the Cape Triage Score in the emergency department of an urban public hospital in South Africa. Emerg Med J 2008; 25: 398–402.
5 Gottschalk S B, Wood D, DeVries S, et al.; on behalf of the Cape Triage Group. The Cape Triage Score: a new triage system South Africa: proposal from the Cape Triage Group. Emerg Med J 2006; 23: 149–153.
6 Wallis L A, Twomey M. Workload and case mix in Cape Town emergency departments. South Afr Med J 2007; 97: 1276–1280.
7 Nasrullah M, Xiang H. The epidemic of injuries in Pakistan—a neglected problem. J Pak Med Assoc 2008; 58: 420–421.
8 Harrison H-L, Raghunath N, Twomey M. Emergency triage, assessment and treatment at a district hospital in Malawi. Emerg Med J 2012; 29: 924–925.
9 Emmanuel A, Ismail A, Kellett J. Assessing the need for hospital admission by the cape triage discriminator presentations and the simple clinical score. Emerg Med J 2010; 27: 852–855.
10 Twomey M, Mullan C, Torrey S, Wallis L, Kestler A. The Princess Marina Hospital accident and emergency triage scale provides highly reliable triage acuity ratings. Emerg Med J 2012; 29: 650–653.
11 American College of Surgeons Committee on Trauma. Resources for optimal care of the injured patient. Chicago, IL, USA: American College of Surgeons, 1998.

A la suite de la mise en œuvre du système d'Echelle de Triage d'Afrique du Sud (SATS) dans le Département des Urgences de l'Hôpital de District Head Quarter (DHQ) à Timergara, Pakistan, 1) nous décrivons ce processus de mise en œuvre et 2) nous signalons dans quelle mesure le système a été utilisé de manière appropriée. Sur les 370 formulaires de triage évalués, 320 (86%) ont été complétés sans erreur, avec comme conséquence une assignation correcte de la priorité de triage. On a noté des erreurs dans 50 formulaires complétés, mais 16 seulement (4%) ont entrainé l'assignation à une priorité incorrecte de triage. Cette expérience montre que le SATS peut être mis en œuvre avec succès et utilisé de manière précise par les infirmières dans un département des urgences au Pakistan.

Tras la introducción del sistema de una escala de selección (South African Triage Scale [SATS]) en el servicio de urgencias del Hospital Central del Distrito de Timergara en Paquistán, el presente artículo tuvo como objeto 1) describir el mecanismo de ejecución del proyecto y 2) rendir un informe sobre la exactitud con que el personal de urgencias utilizó el sistema. De los 370 formularios de selección analizados, 320 (86%) estaban completados sin equivocaciones, con lo cual se había asignado una prioridad correcta de atención. En 50 formularios se observaron errores, pero solo 16 de ellos (4%) tuvieron como consecuencia la asignación de una prioridad de selección inadecuada. Esta experiencia pone en evidencia la eficacia de la aplicación del SATS y confirma la precisión de su utilización por parte del personal de enfermería en un servicio de urgencias en Paquistán.

Public Health Action (PHA) The voice for operational research.
Published by The Union (www.theunion.org), PHA provides a platform to fulfil its mission, 'Health solutions for the poor'. PHA publishes high-quality scientific research that provides new knowledge to improve the accessibility, equity, quality and efficiency of health systems and services.

e-ISSN 2220-8372
Editor-in-Chief: Donald A Enarson, MD, Canada
Contact: pha@theunion.org
PHA website: http://www.theunion.org/index.php/en/journals/pha
Article submission: http://mc.manuscriptcentral.com/pha

Supplemental Discussion

In this paper, we describe the implementation process by which triage and a triage tool were introduced in a low-resource setting. Previous studies of triage in such settings have not described implementation in any detail[54,70,78]; most focus on reliability and validity. While these are essential, implementation is an important element in the success of triage. Successful implementation of an intervention in a healthcare system is dependent upon active change management. ECs in low-resource settings have limited resources and experience constant patient pressure; change in this type of environment is difficult to manage. There are many different theories of change, including systemic change, organisational change and change management.[79] Learning from these theories of change, we present a framework that worked well in Northern Pakistan.

The elements presented below have been borrowed from educational reform and theory of change.[80,81] We present and expand on three aspects that we found contributed to a successful implementation (Figure 6):

- Broad stakeholder ownership,
- Contextualisation, and
- Continuous learning and communication.

We believe that these elements can help others when implementing triage in various settings.

Figure 6: Elements for the successful implementation of triage

Broad stakeholder ownership

For the implementation of a system-based initiative to be successful, all stakeholders need to feel a sense of ownership of the process.[80] They not only need to be involved in the decision making process, but each stakeholder needs to also perceive the benefits of the system change.

Managing multiple stakeholder interests is common in ECs across low-resource settings.[82] This holds true for triage too, with various stakeholders contributing to its success or otherwise. We identified each stakeholder, and mapped out the positive and negative consequences that they might perceive and face because of the change. Taking time to try to view the change through their eyes helped us to develop a suitable approach. Being open to constructive criticism and empowering stakeholders was critical to our engagement.

Below we describe our engagement with each of the different stakeholder we identified. This is not meant to be an exhaustive list of all stakeholders for every setting, but we found it was comprehensive for Timergara Hospital.

Doctors and nurses

In many settings a power dynamic exists between doctors and nurses.[1] Triage in low-resource settings is commonly nurse driven.[54,83,84] The decisional power of nurses in the flow of patients can be a point of contention. The implementation team had to be open to listening and involving all healthcare workers in decisions. We presented the plan as a systemic change for the department. We avoided referring to it as a problem we were trying to fix. Once staff saw it as a system change they were less focussed on attributing blame for a problem. It allowed them to be open to change and to contribute ideas to the improvement. We had both doctors and nurses co-create ideas around the colour coded areas to manage the various acuity levels. We stressed the importance of teamwork and allowed the triage nurses to be the drivers of this initiative.

We built a new triage room outside the hospital and closed off other entrances to incorporate complaints that healthcare workers were having about patients coming from all sides. This created a funnel to allow rapid triage and redirection of patients. Post implementation, we had a local doctor support the triage nurses. He would work through any conflicts between doctors and nurses then and there, consistently feeding back to all parties about lessons learned and small adjustments made to the system. Weekly feedback sessions to all staff allowed everyone to be kept in the loop with any changes.

The view by staff that the triage process would introduce extra work, given that all patients would have to have their vital signs measured, was an area that had to be addressed. Pre-implementation of the triage process, doctors would ordinarily 'eyeball' patients and decide where they should go. This sometimes led to serious cases being missed and unnecessary overload in the resuscitation room (Resus). We explained that the new flow should decrease congestion by redirecting and separating the emergency patients from the low-acuity patients. This would allow more efficient use of our limited resources and staff. We ran games and competitions with staff during training to demonstrate how fast triage could work once staff were familiar with the scale. Centralising and providing dedicated equipment for the triage room helped improve the speed of triage. This alleviated staff fears related to the extra workload.

Non-medical staff

In our settings, we found that non-medical staff were often the first point of contact for many patients, with patients approaching the security and cleaning team more often than healthcare workers. We suspected that these non-medical staff are commonly overlooked stakeholders in low-resource settings. Indeed, few studies mention them in the implementation process.[74,85] Non-medical staff often never receive training on systems in the EC. Many did not understand why or how the flow of patients worked in a particular way but were expected to direct patients.

We conducted basic 30-minute talks on what triage is and why it helps patients. We found that training non-medical staff in triage created a sense of ownership of the triage system. Additionally, these staff felt a sense of empowerment because of their ability to assist patients. Post-implementation, they became strong supporters of the system.

Community and religious leaders

Patients play a key role in the acceptance of any system, but understanding the community from which the patient comes, can facilitate patient acceptance. The Khyber Paktunkhwa area of Northern Pakistan is a conservative area. Religion plays a central role in the community. Religious leaders play a guiding role in all affairs of the community. A large portion of information attained by the community comes from religious leaders. We identified religious leaders as key opinion leaders (KOL) in the community. MSF regularly met these leaders to discuss different issues and in these meetings, we put Triage on the agenda. We explained the process and advantages of triage. We were very open about the fact that this was a new untested system that we wanted to implement, backed by research showing good results in other parts of the world. Respect for cultural norms was key for many leaders. Many wanted to improve the hospital for their community. Being open, honest and having respect for the community's cultural norms helped us gain support from the KOLs.

Political leaders and management

Timergara Hospital was a district level hospital and was under local political oversight. The political party decides the hospital manager. Some high party officials worked as nursing staff in the hospital. This placed more power with some nurses than the hospital manager. The resulting dynamic affected the power structure within the hospital. Only after understanding this dynamic could we understand who to lobby and convince in terms of the merits of triage.

Contextualisation

There are few triage scales designed for low-resource settings. The SATS was designed for the South African context but at the time of this study had not been implemented and assessed widely outside of South Africa. In Botswana, it had been adapted based on the needs faced locally, with good results. Building on the existing literature and contextualising the SATS would thus seem a feasible route.

After acquiring buy-in from internal staff, we discussed previous attempts at system change. One strong sentiment that arose was the rejection of foreign systems. Many healthcare workers felt that these sorts of systems were often forced upon them without proper consultation and without the opportunity for feedback. Previous systems failed as they were perceived as foreign in nature with little room for contextualisation. To address this issue, we set about setting up a series of workshops to contextualise the SATS for Timergara Hospital. Taking on board staff suggestions, we took the following steps:

- Translated the triage algorithms into the local language,
- Incorporated common words used by patients when presenting to the EC,
- Conducted SATS training in the local language (a local doctor was trained in the SATS, and he then in turn trained all the nurses),
- Created posters explaining triage outside the triage room in the local language,
- Developed paper leaflets about triage that we distributed to local mosques, and

- Created separate colour pathways and improved the dressing room. This suggestion helped the flow of yellow patients that presented with fractures or lacerations.

Post-implementation, we listened to issues which arose in the triage room and we engaged with community leaders to address issues early on. One issue that arose was the way we triaged both male and females in the same room. The triage room had space for two patients and their respective care givers at any one time and in peak hours we would have two sets of triage nurses to improve patient flow. This meant that at any one time there could be a female and a male in the same room. We did not initially anticipate that this would be problematic given that patients are not required to undress during triage. However, when taking blood pressure or showing a wound, patients would be asked to expose either the arm or wound. This exposure in front of other patients of the opposite gender was frowned upon. After talking to the KOL, we realised that this was a point of unease for both the leaders and the patients. We explained the new setup of the triage room outside the hospital and asked for their insight into the new flow, incorporating small changes where possible. Before any solution was implemented we were careful to consider the solutions in accordance with medical ethics. Increasing harm or decreasing efficiency of care based on gender was unacceptable. We resolved the problem by separating the triage room into two by means of a curtain that provided privacy for each patient. We also separated the waiting areas for women and men to decrease mixing, which assimilated to the cultural norms.

These adjustments helped create a sense of locality around the SATS for the staff. Creating a scale that was not forced upon them but owned and adapted by them for them.

Continuous learning and feedback

Communication has long been acknowledged as a fundamental element in the change process.[2] Over the three-month period post-implementation, we made sustained efforts to maintain communication with all stakeholders. Any new nurses would be trained to use the SATS by the doctor in charge. Doctors and nurses were encouraged to communicate on any issues, and these issues were addressed immediately. Weekly feedback sessions between the doctor in charge and the staff ensured up-to-date information. Religious leaders were

regularly engaged with about any further issues. We would listen to any patient feedback via either the religious leaders, non-medical staff or patients themselves. We put structures in place to allow regular clinical audits of triage. We also continued to share all results and study publications with the team in Timergara and incorporated local staff into the study as co-authors.

Learning and feedback was made a continual and routine process and was never deemed to be a one-off event. Incorporating it into regular activities such as the daily meeting or monthly meetings with stakeholders facilitated this on-going cycle.

Chapter Conclusion

While little has been written about triage implementation in low-resource settings, we believe that the way in which triage is implemented is key in determining the success of the system. Important elements include taking time to engage with the necessary stakeholders in an appropriate manner, allowing for contextualisation of the SATS to improve ownership by local people, and maintaining communication channels during and post-implementation.

After the successful implementation of the SATS in Northern Pakistan we set out to develop an easier way to test reliability and validity in paediatric low-resource settings.

Chapter 4: Development of paediatric reference vignettes to assess triage scales in low-resource settings

Reference

Dalwai MK, Tayler-Smith K, Twomey M, Wallis L. Developing a reference standard for assessing paediatric triage scales in resource poor settings. AFJEM. DOI: 10.1016/j.afjem.2015.08.002

Declaration from author

The following co-authors contributed to the paper:
Mrs. Katie Tayler-Smith, Dr. Michele Twomey, Prof. Lee Wallis

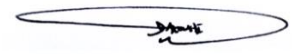
In the case of Chapter 4, the nature and extent of contribution by the authors was as follows:

Nature of contribution

MKD, MT and LW contributed to study conception and design; MKD and KTS facilitated the acquisition of data; MKD, MT, KTS and LW helped with the analysis and interpretation of the data; the initial version was drafted by MKD. KTS, MT and LW contributed to the redrafts and revised it critically for important intellectual content; all authors gave final approval on the version submitted and the revisions thereof.

Extent of contribution

MKD: 80%; KTS: 10%; MT and LW together: 10%



Signed: Mohammed. K. Dalwai 23 Dec 2017

Declaration from Co-Authors

The undersigned hereby certify that:

1. The above declaration correctly reflects the nature and extent of the candidate's contribution to this work, and the nature of the contribution of each of the co-authors.
2. They meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
3. They take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
4. There are no other authors of the publication according these criteria;
5. Potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and
6. The original data are stored in a two factor password protected cloud account and will be held for at least five years

Prof. Lee Wallis 10 Jan 2018

Dr Michele Twomey 10 Jan 2018

Mrs Katie Tayler-Smith 10 Jan 2018

Introduction to the topic

Assessing the effectiveness of triage scales in low-resource settings is not without its challenges. After implementation of the SATS in Chapter 3, we needed an easy way to assess reliability. There are a number of methods for assessing reliability and validity that we will be expanding on in Chapter 6 and Chapter 7, respectively. To assess reliability, one commonly used method is to ask nurses to triage paper-based vignettes. Some studies then have the paper-based vignettes triaged by experts and the acuity levels assigned by the experts taken as the gold standard to test for validity.

Motivation for conducting the study

In low-resource settings, it is not always possible to collect and create paper-based vignettes, given limited resources and expertise. Currently there are no freely available sets of pre-validated paper-based paediatric reference vignettes for use in low-resource settings. Twomey et al. produced and made available a set of pre-validated adult vignettes; to date however, there has been no set for paediatric settings.[69] In Chapter 3, we described the successful implementation of the SATS. As the SATS is more widely used across diverse low-resource settings than any other triage scale, it would be useful to create an easy way to assess reliability and validity in each setting.

Aim

This study aims to create a set of paediatric reference vignettes that can be used to test the reliability and validity of triage scales.

Objectives

Build a set of consensus-based reference vignettes.

Main findings

- Context-specific reference vignettes can be developed to provide a cheap, effective, and feasible means by which to evaluate paediatric triage systems in low-resource settings.
- This set of reference vignettes should only be used in settings that are similar in disease profile to South Africa.

Footnote: The article was published in AfJEM as the authors felt it was most applicable. AfJEM readership includes healthcare workers working in low-resource settings. Many are implementing or have implemented triage scales. We felt that they would benefit the most from a set of reference vignettes for low-resource settings.

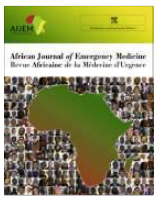
Copy of published Paper

African Journal of Emergency Medicine (2015) **5**, 181–184

African Federation for Emergency Medicine

African Journal of Emergency Medicine

www.afjem.com
www.sciencedirect.com



ORIGINAL RESEARCH ARTICLES

Developing a reference standard for assessing paediatric triage scales in resource poor settings

Élaborer une norme de référence pour évaluer les échelles de triage pédiatrique dans des contextes caractérisés par un manque de ressources

Mohammed Dalwai [a,*], **Katie Tayler-Smith** [b], **Michele Twomey** [a], **Lee Wallis** [a]

[a] *Division of Emergency Medicine, University of Cape Town, Bellville, Cape Town 7535, South Africa*
[b] *Médecins Sans Frontières, Operation Centre Brussels, Operational Research Unit, MSF-Luxembourg, 68 rue de Gasperich, Luxembourg 1617, Luxembourg*

Received 29 May 2015; revised 3 August 2015; accepted 24 August 2015; available online 19 October 2015

Introduction: One of the main challenges for emergency healthcare services in low to middle income countries (LMICs) is limited capacity to deal with heavy emergency caseloads. The process of triage is one mechanism for mitigating this challenge.
Methods: In a two-round consensus building process (the Delphi process), a panel of emergency centre (EC) experts were asked to independently triage 50 clinical vignettes using one of four acuity levels: emergency (patient to be seen immediately), very urgent (patient to be seen within 10 min), urgent (patient to be seen within 60 min), or routine (patient to be seen within four hours). The vignettes were based on real paediatric EC cases in South Africa. Vignettes that reached a minimum of 80% group consensus for acuity ratings on either round one or two were included in the final set of reference vignettes.
Results: Of the 50 vignettes presented to 11 EC experts, in the first round, 80% group consensus on acuity ratings was obtained for 10 (20%) of the vignettes. In the second round, 80% consensus was reached for 30 of the 40 remaining vignettes. Thus, overall, 40 (80%) of the vignettes reached a minimum group consensus of 80% (emergency $n = 4$; very urgent $n = 8$; urgent $n = 12$; routine $n = 16$).
Conclusion: This study demonstrates how context-specific reference vignettes can be developed to provide a cheap, effective, and feasible means by which to evaluate paediatric triage systems in LMICs.

Introduction: L'une des principales difficultés associées aux services de santé d'urgence dans les pays à faible et moyen revenu (PFMR) est leur capacité limitée à faire face à une lourde charge d'urgences médicales. Le processus de triage est l'unique mécanisme permettant d'atténuer cette difficulté.
Méthodes: Au cours d'un processus de recherche de consensus en deux étapes (la méthode Delphi), il a été demandé à un panel d'experts issus de centres d'urgence (CU) de trier indépendamment 50 vignettes cliniques en sélectionnant un niveau d'acuité parmi les 4 niveaux proposés: urgence (le patient doit être examiné immédiatement), très urgent (le patient doit être examiné dans les 10 min), urgent (le patient doit être examiné dans les 60 min) ou routine (le patient doit être examiné dans les 4 heures). Les vignettes étaient basées sur de véritables cas d'urgences pédiatriques en Afrique du Sud. Les vignettes résultant sur un consensus de groupe de 80% minimum quant aux évaluations de l'acuité à la première ou à la deuxième étape ont été incluses à l'ensemble final de vignettes de référence.
Résultats: Sur les 50 vignettes présentées à 11 experts issus de CU, au cours de la première étape, un consensus de groupe de 80% quant aux évaluations de l'acuité a été obtenu pour 10 (20%) des vignettes. Dans la seconde étape, un consensus de 80% a été obtenu pour 30 des 40 vignettes restantes. Ainsi, au total, 40 (80%) vignettes ont atteint un consensus de groupe minimum de 80% (urgence n = 4; très urgent n = 8; urgent n = 12; routine n = 16).
Conclusion: Cette étude a montré comment des vignettes de référence spécifiques à un environnement donné pouvaient constituer un moyen peu coûteux, efficace et faisable d'évaluer les systèmes de triage pédiatrique dans les PFMR.

African relevance

- Triage is poorly researched in Africa, especially paediatric triage.
- Poor record keeping makes validating triage scales very difficult.
- Paper based vignettes can be used in low resource settings.

* **Correspondence to Mohammed Dalwai**. mkdalwai@gmail.com

Peer review under responsibility of African Federation for Emergency Medicine.

Production and hosting by Elsevier

Introduction

One of the main challenges for emergency health care services in low to middle income countries (LMICs) is their limited capacity to deal with heavy emergency caseloads. The process of triage is one mechanism for mitigating this challenge.[1]

Triage aims to determine a patient's urgency for medical care (defined as their acuity level) in order to separate critically ill patients, who need immediate lifesaving interventions, from patients who need medical attention but can safely wait to be seen.[2] Triage is recognised as being one of the core requirements for the provision of effective emergency care, and has been shown to reduce patient morbidity and mortality.[3]

http://dx.doi.org/10.1016/j.afjem.2015.08.002

In LMIC settings, however, triage remains under-used and under-researched, particularly in the area of paediatric emergency care.

The triage of adults and children relies on different triage scales in order to take account of physiological differences between the two. Very few paediatric triage scales exist for the triage of children in LMIC settings.[4] Until recently, the most widely recommended scale was the Emergency Triage Assessment and Treatment (ETAT) system developed in 1998 by the World Health Organisation. However, this scale is only applicable for use in children under five.[5] The South African Triage Scale (SATS) – developed in 2004 by the Cape Triage Group – is the only other triage scale designed specifically for LMIC settings, and one of its advantages over the ETAT, is that it includes scales for the triage of infants and children up to the age of 12 years.[6] The main issue with both the paediatric versions of the SATS and the ETAT is that they are not formally validated in various contexts of use.

Validating a triage system in many contexts remains a challenge due to lack of a gold standard.[7] To circumvent this, various studies have assessed validity using surrogate outcome markers such as mortality rates, resource utilisation, and length of hospital stay as proxies for true acuity level.[8,9] In LMICs, however, reliance on such surrogate markers is difficult due to varying levels of care, lack of basic resources, and poor record keeping. As an alternative, Twomey et al. have recommended using the modified Delphi method to develop an objective reference standard against which to evaluate a triage scale.[10] The Delphi method is a consensus-building technique, which, in the context of triage validation, can be used to develop a set of reference vignettes (short written case reports based on real emergency centre (EC) cases). This methodology has been used by Twomey et al. to assess the validity of the adult version of the SATS.[10]

In this study, we aim to demonstrate how the modified Delphi method can be applied to develop a representative sample of context specific vignettes for assessing paediatric triage systems.

Methods

The study involved applying a modified Delphi method to a series of clinical vignettes based on real paediatric EC cases. Fifty paediatric vignettes were generated by randomly selecting real paediatric EC cases aged 0–12 years presenting at ECs across the Western Cape from 3 October to 30 November 2011. The source of these cases was an electronic database. Each vignette included information on patient gender, age, presenting complaint, mode of arrival to the EC, and vital signs. All information included in the triage paperwork was included in the vignettes, including information from additional investigations such as blood glucose and haemoglobin (see Appendix 1 for examples of the vignettes – data Supplement). The 50 vignettes comprised 10 surgical/trauma cases and 40 medical cases.

The Delphi group of study participants was comprised of individuals deemed by the authors to be EC triage experts on account of them either having published research on triage or having worked in ECs in LMICs, and who gave consent to participate in the study. The Delphi group was asked to independently complete a two-round consensus building process, each round lasting a month. In round one, each participant was emailed and asked to triage the 50 vignettes based on their clinical experience and the triage tool that they were most familiar with using. They were requested to assign one of the following four acuity levels to each vignette: emergency (patient to be seen immediately), very urgent (patient to be seen within 10 min), urgent (patient to be seen within 60 min) or routine (patient to be seen within 4 h).

The vignettes were made available online to facilitate easy access at a time that was convenient to each participant and an online survey tool was used to collect the triage ratings that they assigned to each vignette. On completion of round one, any vignettes achieving at least 80% consensus among the participants were kept aside. Vignettes that did not reach this level of consensus were sent back to the Delphi group for round two, where the acuity level assigned to each vignette by the majority of participants during round one was indicated and participants were given the opportunity to either change the original acuity rating that they had assigned or leave it as it was. At the end of round two, triage ratings were summarised for all vignettes, and only those that reached a minimum of 80% consensus on either round one or two were included in the final set of reference vignettes.

Informed consent was obtained from all experts participating in the study. The Western Cape Paediatric Triage database, which was used to develop the vignettes, contains no patient names or identifying information. Ethics approval was obtained from the University of Cape Town Human Research Ethics Committee.

Results

A total of 59 triage experts were contacted to participate in the study. Of these, 14 took part in the first round of the Delphi process, and 11 of the 14 completed the second round. These 11 participants made up the final Delphi group.

In the first round of the Delphi process, a minimum of 80% group consensus on acuity ratings was obtained for 10 (20%) of the 50 vignettes. Of the 40 vignettes for which 80% group consensus was not reached, 19 (48%) reached between 60% and 79% consensus, 16 (40%) reached 50–59% consensus and five (13%) reached less than 49% consensus. Discrepancies for these 40 vignettes were as follows: 13 (33%) were found between 'routine' and 'urgent' acuity levels, 11 (28%) between 'urgent' and 'very urgent' acuity levels, 6 (15%) between 'very urgent' and 'emergency' acuity levels and ten (25%) had discrepancies at multiple acuity levels.

In the second round, a minimum of 80% consensus was reached for 30 (75%) of the 40 vignettes that had failed to reach 80% consensus in round one. The degree to which panel members changed their assigned acuity levels to reach a 80% consensus between round one and round two for these 30 vignettes was as follows: for nine (30%) of the vignettes only one member changed their acuity level, for a further nine (30%) vignettes, two members changed their acuity levels, for five (17%) vignettes three members changed their acuity levels, for three (10%) vignettes four members changed their acuity level, and for the remaining four (13%) vignettes five members changed their acuity level.

Overall, 40 (80%) of the 50 vignettes reached a minimum group consensus of 80% (emergency $n = 4$; very urgent $n = 8$; urgent $n = 12$; routine $n = 16$).

See Appendix 1 (data supplement) for the final 40 vignettes reaching 80% consensus and Appendix 2 (data supplement), for the 10 vignettes failing to reach consensus. Of note, five of the vignettes that failed to reach consensus were for respiratory presentations.

Discussion

This study demonstrates how the modified Delphi method can be used to develop a validated set of context specific reference vignettes for assessing paediatric triage scales. This methodology has previously been applied by Twomey et al. to generate a set of adult reference vignettes but has not been formally used to develop paediatric vignettes.[10]

Validation of triage scales (including paediatric scales) is a major challenge, particularly in resource limited settings. As such, in most LMIC settings where triage systems are being used, these scales have not been formally evaluated for that context. Reference vignettes based on real EC cases and formally validated by a group of experts provide a potentially cheap, effective, and more feasible means by which to evaluate such triage scales.

The merits of these validated reference vignettes are that (i) they are based on the random selection of real EC cases (and are thus representative of true EC case presentations seen in a particular context), and (ii) the ascribed acuity ratings for each vignette are based on a two-round consensus building process using the expert opinions of emergency medicine specialists either working or having experience in different LMIC settings.

Possible limitations of the reference vignettes are linked to the composition of the Delphi panel and the Delphi method itself. First, fifty-nine experts were invited to take part in this study, of which 14 agreed to participate, and of these only 11 completed both rounds of the study. Despite this low initial response rate and some attrition during the second round of the study, the final Delphi panel was relatively well-represented comprising of approximately equal numbers of emergency medicine doctors, paediatricians, and emergency nurses, with half practicing in LMICs and half practicing in developed countries. Second, the dramatic increase in the proportion of vignettes attaining a minimum of 80% consensus between round one and two (20–80%) may appear to imply that a substantial proportion of the panel members changed their mind on acuity ratings from round one to round two. In fact, this was often not the case. For a significant number of those vignettes where 80% consensus was not reached, the actual level of consensus was often not far below the 80% threshold in round one. Third, where panel members did change the acuity rating that they assigned to a vignette between rounds, we did not explore the basis for this change and cannot therefore speculate on the relative logic underlying a panel member's decision to change their mind. Further exploration of this would provide useful information about the relative validity of this consensus building technique for developing reference vignettes. Fourth, important minority issues may have been overlooked by trying to obtain consensus.[10]

The Delphi methodology was first introduced by the Research and Development (RAND) Corporation in 1946 to develop information in fields where little data exist.[10] It deals with complex problems by acknowledging human judgment as legitimate and useful as a method of gathering expert opinion into a single useful statement(s). The strengths include rapid results, cheap costs, avoidance of self-censorship common in group meetings, and the ability of experts from all around the world to participate. Some of the limitations of this method include high dropout rates of experts and difficulty to ensure engagement over a period of time, poor expert selection, time demands related to coordinating the process, and a decreased transparency of decisions comparative to face to face meetings.[11]

Despite the limitations, we believe that the Delphi Method is a feasible way for developing an objective reference standard against which to validate paediatric triage scales in LMIC settings. What is important, however, is that the reference vignettes are context specific (i.e. are based on EC cases specific to the setting in which the triage system is being validated), and thus, an epidemiologic pattern of disease and trauma is reflected. In the current study, the reference vignettes were based on EC cases seen in South Africa. We believe it would be reasonable to use these vignettes to assess a triage system in other very similar LMIC settings, but in dissimilar settings, context specific reference vignettes would need to be developed.

Several other findings in our study also have implications for further study and for the future development of reference vignettes using the Delphi methodology. First, it was observed that lack of consensus in the first round was most often linked to discrepancies between 'routine' and 'urgent' acuity ratings and 'urgent' and 'very urgent' acuity ratings. Difficulties in reaching consensus between 'urgent' and 'very urgent' acuity ratings have been reported previously by Twomey et al. for adult EC cases (reference). This extended to 'routine' and 'urgent' acuity levels for paediatric presentations in the current study. Exploring the reasons underlying Delphi members' decisions to assign specific acuity ratings would help to better understand the basis of these discrepancies and to decide how best to accommodate this issue when developing a set of paediatric reference vignettes. Finally, half of the vignettes that failed to reach consensus were for respiratory presentations, suggesting possible difficulties in triaging respiratory signs and symptoms in children. As such, when a set of paediatric reference vignettes are generated using the Delphi method, there may be a tendency for respiratory presentations to be under-represented. If this is found to be the case (on account of many of the vignettes that fail to reach consensus being respiratory in nature), this should be acknowledged when using the reference vignettes to assess a triage scale.

In conclusion, this study demonstrates how context specific reference vignettes can be developed to provide a cheap, effective and feasible means by which to evaluate paediatric triage systems in LMICs. Formal studies reporting on the use of these reference vignettes to evaluate the paediatric version of the SATS are now needed.

Conflict of interest

The authors declare no conflict of interest. The study was funded by the Medical Research Council of South Africa under the MRC Clinician Researcher Program.

Author contribution

M.D. and M.T. conceived the original idea. M.D., M.T. and L.W. designed the experiments and helped with acquisition of data. K.T.S. and M.D. carried out analysis and interpretation of data. M.D. and K.T.S. prepared the manuscript. K.T. S., L.W. and M.T. revised the article and approved final content for submission.

Acknowledgements

The authors would like to thank all the triage experts who took part in the Delphi for their time. It is only with your contribution that we can move paediatric triage research forward.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.afjem.2015.08.002.

References

1. Razzak JA, Kellermann AL. Emergency medical care in developing countries: is it worthwhile? *Bull World Health Organ* 2002;**80**(11):900–5.
2. Dalwai MK, Twomey M, Maikere J, Said S, Wakeel M. Reliability and accuracy of the South African Triage Scale when used by nurses in the Emergency Department of Timergara Hospital, Pakistan. *S Afr Med J* 2014;**104**(5):372–5.
3. Robison JA, Ahmad ZP, Nosek CA, Durand C, Namathanga A, Milazi R, et al. Decreased pediatric hospital mortality after an intervention to improve emergency care in Lilongwe, Malawi. *Pediatrics* 2012;**130**(3):e676–82.
4. Twomey M, Cheema B, Buys H. Vital signs for children at triage: a multicentre validation of the revised South African Triage Scale (SATS) for children. *S Afr Med J* 2013;**103**(5):3–8.
5. Tamburlini G, Di Mario S, Maggi RS, Vilarim JN, Gove S. Evaluation of guidelines for emergency triage assessment and treatment in developing countries. *Arch Dis Child* 1999;**81**(6):478–82.
6. Gottschalk SB, Wood D, DeVries S, Wallis LA, Bruijns S. The Cape Triage Score: a new triage system South Africa. Proposal from the Cape Triage Group. *Emerg Med J* 2006;**23**(2):149–53.
7. Twomey M, Wallis LA, Myers JE. Limitations in validating emergency department triage scales. *Emerg Med J* 2007;**24**(7):477–9.
8. Olofsson P, Gellerstedt M, Carlström ED. Manchester Triage in Sweden – interrater reliability and accuracy. *Int Emerg Nurs* 2009 Jul;**17**(3):143–8.
9. Wuerz RC, Milne LW, Eitel DR, Travers D, Gilboy N. Reliability and validity of a new five-level triage instrument. *Acad Emerg Med* 2000;**7**(3):236–42.
10. Twomey M, Wallis L, Myers J. Evaluating the construct of triage acuity against a set of reference vignettes developed via modified Delphi method. *Emerg Med J. Published Online First* 2013;**24**. http://dx.doi.org/10.1136/emermed-2013-202352.
11. Sackman H. Delphi Assessment: Expert Opinion, Forecasting and Group Process. United States Air Force Project RAND. 1974; p 130. http://rand.org/pubs/reports/2006/R1283.pdf.

Supplemental Discussion

In the methods section of the paper above, we describe the Delphi process. In recent years, the Delphi method has been used frequently in emergency medicine to build adult triage vignettes, develop criteria for pre-hospital trauma care, and construct criteria for major incident trauma algorithms.[69,86–88]

The Delphi methodology used in the paper above produced validated paper-based vignettes. We have touched on some of the limitations in the article above but will elaborate on lessons learnt after the creation and subsequent attempts to use the above vignettes.

Validated triage vignettes can be used in two main ways for triage research:

1. Assessing Reliability – Triage raters were asked to rate each vignette in a classroom setting. Inter-rater reliability can be calculated using ICC or QWK. Intra-rater reliability can be assessed using a subset of the same vignettes after a pre-defined time period. We explain this further in Chapter 6.
2. Assessing Validity – The ratings of the triage raters obtained were compared to the 'true ratings' (as defined by the Delphi process) of each vignette and over- and under-triage were calculated. We elaborate on all methods of validity in Chapter 7.

Validated vignettes provide researchers with an opportunity to assess both reliability and validity at one time, making the process both cost- and time-efficient. Adult vignettes have already been created; this study aimed to add to the pool, a set of paediatric vignettes (Appendix 6).

Using vignettes to assess reliability

When assessing reliability, many studies – both in low- and high-resource settings – use paper-based vignettes. Some studies in high-resource settings have used a research nurse to simultaneously assess the same patient directly after the triage nurse as a means to test reliability.[65,66] This is, however, expensive and impractical in low-resource settings. Paper-based vignettes, although limited in their ability to convey audio-visual cues, are deemed by

most triage researchers to provide an acceptable indication of the patient's acuity.[13,38,67,68]

After development and implementation of the above paediatric clinical vignettes, a clear limitation became apparent. We found that their use was limited to similar settings only. In different settings, results could be inaccurate because the vignettes might not represent the actual case presentations seen that setting. Our suggestion therefore was that researchers only use the set of vignettes if the disease profiles, epidemiological patterns and clinical presentations seen in their setting, were similar to the those seen in the setting from which the vignettes were derived. This of course restricts the settings in which these vignettes can be used. In Chapter 6 we describe the reliability of the SATS in Sierra Leone. Unfortunately, given the high burden of malaria and a different disease profile of patients, we could not use the set of vignettes that we developed. Given this limitation, it was be sensible to create vignettes based on local presentations when assessing reliability.

Using vignettes to assess validity

When using vignettes solely for the purposes of reliability, studies almost never validate them. When using vignettes for validation, each vignette needs to have a 'true acuity' – that is, a defined triage level. The defined triage level is then used as a proxy gold standard against which a rater can be compared. If the raters triage the vignette the same as the gold standard, then the triage scale is said to have high validity. A degree of discrepancy is allowed with most studies quoting the ACS-COT indicators of up to 10% under-triage and up to 50% over-triage which has been revised to 25-35% over-triage and up to 5% under-triage in 2014.[76] To validate vignettes, studies usually use experts to create a 'gold standard'.[19,89] The composition of these experts/panels is a limitation that can affect the final triage acuity assigned to each vignette. Experts from high resource settings, or those with limited experience in low-resource settings, might offer different acuity levels than those in low-resource settings. Using a Delphi method with its blinded responses and consensus building process facilitates the creation of a more robust set of vignettes. We recommend that when using paper-based vignettes as a tool for validation, a Delphi process should be followed to develop the gold standard.

Generalisability is a limitation for validation, as it is for reliability. Having locally relevant cases would decrease any bias that could be introduced. We suggest creating a broad range of disease specific paediatric vignettes from various settings using the Delphi methodology. Ensuring the panel consists of experts from the local settings or experts that have extensive experience in low-resource settings would be highly recommended. As different disease sets become available, researchers could choose the set that best fits their setting.

An additional limitation when using vignettes for validation is the number of acuity levels. Triage scales range from three to five levels. Vignettes created for a three- or four-level system are not suitable for studies assessing scales with five levels. This further limits the ability to use vignettes for validation across various scales in low-resource settings. This is discussed further in Chapter 7.

Chapter Conclusion

Initially, we set out to create an easy and cheap way to assess reliability in low-resource settings. This is the first attempt at developing paediatric clinical vignettes for such settings. The Delphi process proved to be a useful tool but has certain limitations. The creation of vignettes that can be applied across various settings remains a challenge. Further disease specific sets are needed to improve on this for varied settings.

After describing the implementation of the SATS in Chapter 3 and having developed a set of paediatric vignettes, we needed to look at formal testing of the SATS in a low-resource setting.

Chapter 5: Reliability and accuracy of the SATS in a low-resource setting

Reference

Dalwai MK, Twomey M, Maikere J, Said S, Wakeel M, Jemmy JP, Valles P, Tayler-Smith K, Wallis L, Zachariah R. Reliability and accuracy of the South African Triage Scale when used by nurses in the emergency department of Timergara Hospital, Pakistan. SAMJ. DOI: 10.7196/samj.7604

Declaration from author

The following co-authors contributed to the paper:
Dr. Michele Twomey, Dr. Jacob Maikere, Mr. Shujaat Said, Dr. Muhammed Wakeel, Dr. Jean-Paul Jemmy, Dr. Pola Valles, Mrs. Katie Tayler-Smith, Prof. Lee Wallis and Dr. Rony Zachariah

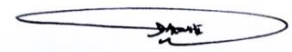
In the case of Chapter 5, the nature and extent of contribution by the authors was as follows:

Nature of contribution

MKD, MT and LW contributed to study conception and design; MKD, JM, SS, MW, JPJ and PV facilitated the acquisition of data; MKD, LW, KTS, PV, RZ and MT helped with the analysis and interpretation of the data; the initial version was drafted by MKD, KTS and MT with all authors contributing to the redrafts and revising it critically for important intellectual content; all authors gave final approval on the version submitted and the revisions thereof.

Extent of contribution

MKD: 80%; MT, LW, KTS together: 10%; JM, SS, MW, JPJ together: 5%; PV and RZ together 5%



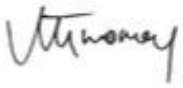
Signed: Mohammed. K. Dalwai 02 Sept 2017

Declaration from Co-Authors

The undersigned hereby certify that:

1. The above declaration correctly reflects the nature and extent of the candidate's contribution to this work, and the nature of the contribution of each of the co-authors.
2. They meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
3. They take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
4. There are no other author of the publication according these criteria;
5. Potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and
6. The original data are stored in a two factor password protected cloud account and will be held for at least five years

Prof. Lee Wallis 22 Dec 2017



Dr Michele Twomey 22 Dec 2017

Mrs Katie Tayler-Smith 22 Dec 2017

Introduction to the topic

With all triage scales, reliability and validity are essential to test in the various hospitals of deployment. There are multiple ways of assessing both reliability and validity that will be expanded upon in Chapter 6 and 7, respectively.

Motivation for conducting the study

In Chapter 5 we described the successful implementation of the SATS in a low-resource setting. The study called for the further assessment of inter- and intra-rater reliability. The SATS was designed for use in the context of South Africa; few studies have examined its reliability outside of South Africa. Examination of the latter could help ascertain the feasibility of the SATS as a triage tool in settings outside of South Africa.

Aim

This study aims to investigate the inter- and intra-rater reliability and accuracy of the adult version of the SATS in a low-resource setting.

Objectives

- To test the inter- and intra-rater reliability of the SATS when used by emergency centre staff in a non-South African setting.
- To determine the accuracy of the SATS using expert reference standard methodology in a non-South African setting.

Main findings

- Overall, the SATS demonstrated substantial inter-rater reliability and 87% intra-rater reliability in a rural hospital setting in Northern Pakistan.
- The SATS was accurate in predicting true acuity for patients triaged as ‘green’, ‘yellow’ and ‘orange’, but not for patients triaged as ‘red’.
- Assessing validity using paper-based vignettes had some limitations.

Copy of published paper

Reliability and accuracy of the South African Triage Scale when used by nurses in the emergency department of Timergara Hospital, Pakistan

M K Dalwai,[1,2] MB ChB; **M Twomey**,[2] BSc, PhD; **J Maikere**,[1] MB ChB, FCS, PhD; **S Said**,[1] BSc (Nursing); **M Wakeel**,[3] MB ChB; **J-P Jemmy**,[4] MB ChB; **P Valles**,[4] MB ChB; **K Tayler-Smith**,[4] **L Wallis**,[2] MB ChB, FCEM, PhD; **R Zachariah**,[4] MB ChB

[1] *Médecins Sans Frontières, Pakistan*
[2] *Division of Emergency Medicine, Faculty of Health Sciences, University of Cape Town, South Africa*
[3] *Ministry of Health, Islamabad, Pakistan*
[4] *Médecins Sans Frontières, Medical Department (Operational Research), Operational Centre, Brussels, Belgium, MSF-Luxembourg (LuxOR), Luxembourg*

Corresponding author: *M K Dalwai (mkdalwai@gmail.com)*

Background. Triage is one of the core requirements for the provision of effective emergency care and has been shown to reduce patient mortality. However, in low- and middle-income countries this strategy is underused, under-resourced and poorly researched.
Objective. To assess the inter- and intra-rater reliability and accuracy of nurse triage ratings when using the South African Triage Scale (SATS) in an emergency department (ED) in Timergara, Pakistan.
Methods. Fifteen ED nurses assigned triage ratings to a set of 42 reference vignettes (written case reports of ED patients) under classroom conditions. Inter-rater reliability was assessed by comparing these triage ratings; intra-rater reliability was assessed by asking the nurses to re-triage 10 random vignettes from the original set of 42 vignettes and comparing these duplicate ratings. Accuracy of the nurse ratings was measured against the reference standard.
Results. Inter-rater reliability was substantial (intraclass correlation coefficient 0.77; 95% confidence interval (CI) 0.69 - 0.85). The intra-rater agreement was also high with 87% exact agreement (95% CI 67 - 100) and 100% agreement allowing for a one-level discrepancy in triage ratings. Overall, the SATS had high specificity (97%) and moderate sensitivity (70%). Across all acuity levels the proportion of over-triage did not exceed the acceptable threshold of 30 - 50%. Under-triage was acceptable for all except emergency cases (66%).
Conclusion. ED nurses in Pakistan can reliably use the SATS to assign triage acuity ratings. While the tool is accurate for 'very urgent' and 'routine' cases, importantly, it may under-triage 'emergency' cases requiring immediate attention. Approaches that will improve accuracy and validity are discussed.

S Afr Med J 2014;104(5):372-375. DOI:10.7196/SAMJ.7604

With the increase in urbanisation and violent conflicts, together with the growing burden of chronic non-communicable diseases in many low- and middle-income countries (LMICs), there is an increased burden on emergency healthcare services.[1] In many LMICs, one of the main challenges facing emergency services is the capacity to deal with high patient loads.[2] The process of 'triage' is one way of addressing this challenge, since it optimises the allocation and use of existing resources.

Triage is the process of sorting critically ill patients who need immediate lifesaving interventions from patients who need medical attention but can safely wait to be seen.[3] Triage aims to determine a patient's 'acuity level' – i.e. how urgently they require medical attention. Triage is recognised as being one of the core requirements for the provision of effective emergency care and has been shown to reduce patient mortality.[4] However, in LMICs this strategy is underused, under-resourced and poorly researched.

The South African Triage Scale (SATS) was developed in 2004 for pre- and in-hospital emergency units throughout South Africa (SA).[5] It was specifically designed to be used by nursing assistants and as such was intended to serve as a coping measure to address medical staff shortages and limited resources – challenges that are commonplace in SA, as in other LMICs.[6]

In 2011, Médecins Sans Frontières (MSF), an international medical humanitarian organisation, implemented the SATS in Timergara Hospital (TH) in the rural district of Lower Dir in the province of Khyber Pakhtunkhwa (KPK), Pakistan. MSF had been working at this hospital alongside the Pakistan Ministry of Health to improve emergency healthcare for the population. Against a backdrop of limited resources, overstretched staff and the absence of a standardised triage system, MSF implemented the SATS with good preliminary results.[7]

The SATS has been assessed extensively in SA and implemented in several LMIC settings.[8,9] However, a more formal assessment of the SATS in a LMIC setting outside sub-Saharan Africa has not yet been undertaken.

The two most common measures for assessing a triage scale are reliability and validity. Reliability is the extent to which the triage scale yields the same result on repeated assessments of the same patient. Inter-rater reliability determines whether there is variability between different staff rating the same patient, while intra-rater reliability assesses the variability for one member of staff re-triaging the same patient. Validity has been defined as indicating how closely an acuity rating assigned using the triage scale is to the true acuity of that patient.[10] However, limitations exist when trying to validate triage scales in any setting, owing to lack of a gold standard. As

such, validity has been assessed using surrogate markers such as hospital admission, discharge and resource utilisation.[11] In LMICs, however, the use of these surrogate markers is difficult owing to poor record keeping, varying levels of clinical skills and limited resources. Previous studies in LMICs have instead attempted to assess the validity of a triage scale by comparing the triage ratings assigned by emergency department (ED) staff for a series of simulated cases against those obtained from an expert panel based on the panel's expert opinion.[12] For the purpose of this study, we will refer to this methodology and use a set of 42 reference vignettes as a reference standard against which accuracy is measured.[13]

This study therefore aimed to determine the reliability (inter- and intra-rater) and accuracy of the adult version of the SATS when used by ED nurses in TH, Pakistan.

Methods

Study design

This was a cross-sectional study using a set of 42 reference vignettes (short, written, clinical case reports of ED patients) as a proxy for live ED cases.

Setting

TH is situated in the predominantly rural district of Lower Dir in the KPK province of Pakistan. It is the only district hospital in Lower Dir, serving an estimated population of 1.8 million. The ED has an estimated annual caseload of ~48 000 patients, comprising both adults and children. The caseload is largely made up of medical emergencies (typically respiratory infection, cardiac disease and gastrointestinal illness) and trauma (most often road traffic accidents).

SATS and its use in the TH ED

The SATS uses a physiologically based composite scoring system, the Triage Early Warning Score, together with a list of discriminators, with which to triage patients into one of five colour-coded groups according to their degree of urgency for medical attention. The colour categories are as follows: (*i*) red, 'emergency' (to be seen immediately); (*ii*) orange, 'very urgent' (to be seen within 10 min); (*iii*) yellow, 'urgent' (to be seen within 60 min); (*iv*) green, 'routine' (to be seen within 240 min, i.e. minor injuries/ illness); and (*v*) black, 'dead'.

The SATS was introduced in the TH ED in June 2011. All ED staff received a 1-hour structured training course, which was carried out by the expatriate ED doctor. It involved explaining patient flow in the ED together with each step of the triage algorithm and the composite physiological score where each vital sign is not seen in isolation but rather as a composite part of an early warning score. Each discriminator was explained using common local ED examples.

Using the SATS, triage was routinely undertaken by two triage nurses during each work shift. Once triaged, 'red' and 'orange' patients were seen by the MSF team (a national doctor, three nurses and an expatriate doctor) in the resuscitation room, while 'yellow' and 'green' patients were seen by the national casualty medical officers in a room adjacent to the ED. At the time of the study, 23 nurses were on the ED rota and carrying out triage.

Study population

The study included all nurses at TH who fulfilled the following inclusion criteria: (*i*) those who had received training in the SATS and had at least 1 month's experience performing patient triage using this tool; and (*ii*) those who agreed to participate in the study. As the study attempted to recruit all nurses fulfilling the above criteria, it was not necessary to calculate the required sample size.

Data collection

Under classroom conditions, nurses participating in the study were required to assign one of four priority categories to the set of 42 reference vignettes according to the SATS acuity levels of 'emergency', 'very urgent', 'urgent' and 'routine'. The vignettes had been collected and validated in a previous study and were based on real ED cases from a secondary hospital in SA.[13] The type and spectrum of patient presentations captured in these vignettes closely mirrored the sort of cases presenting at the TH ED. The vignettes included information on patient gender, age, presenting complaint, mode of arrival to the ED, and vital signs. Some vignettes also included information from additional investigations such as blood glucose test and haemoglobin, as done at the time of triage. For the purpose of this study, the vignettes were translated from English into Urdu, the national language of Pakistan. This was carried out by a professional translator and ratified by a local bilingual doctor to ensure the correct medical terminology.

Reliability

Inter-rater reliability was measured by comparing the different nurse triage ratings for the 42 vignettes, while intra-rater reliability was measured by asking the nurses to re-triage 10 random vignettes from the original set of 42 vignettes and comparing these duplicate ratings.

Accuracy

The accuracy of nurse triage ratings for the 42 vignettes was measured by comparing their ratings with the acuity ratings assigned to the same set of vignettes by an international expert panel. The panel of 18 experts, made up of emergency medicine physicians and emergency nurses from developing and developed countries, were chosen from countries where triage scales were already established and validated or being established and validated. They had already independently reviewed the vignettes used in the current study, and via a modified Delphi technique, obtained consensus on 'true' acuity level for each vignette. They assigned an acuity level based on their expert opinion rather than through the application of the SATS. The acuity levels that they assigned had to fall into one of four categories to mirror the SATS categories of 'emergency', 'very urgent', 'urgent' and 'routine'.

Data analysis

In accordance with the Guidelines for Reporting Reliability and Agreement Studies (GRRAS), inter-rater reliability was assessed using the unweighted, linearly weighted and quadratically weighted κ (QWK) statistic, as well as the intraclass correlation coefficient (ICC).[14] The QWK is commonly used when reporting on reliability studies because it takes into account the degree of disagreement. A weighted κ uses maximum weights at two opposite ends of the scale and is therefore identical to the ICC.[10] Whereas the unweighted and linear weighted κ is not commonly used in triage literature, it has been reported in this case to follow the GRRAS for easy comparisons between other studies.[14] Point estimate values for QWK and ICC were graded using the Landis and Koch classification system as follows: 0.0 - 0.20 – slight agreement; 0.21 - 0.40 – fair agreement; 0.41 - 0.60 – moderate agreement; 0.61 - 0.80 – substantial agreement; and 0.81 - 1.00 – almost perfect agreement.[10] Intra-rater reliability was assessed by calculating the percentage of exact agreement and also the percentage of agreement allowing for one level of discrepancy in the triage ratings.

The accuracy of the nurse triage ratings was assessed by calculating the sensitivity, specificity, and associated over-/under-triage relative to the experts' triage ratings. Over- and under-triage were interpreted using an accepted range for average under-triage of not more than 5 - 10% and an

associated average over-triage rate of 30 - 50%; these are the ranges considered acceptable by the American College of Surgeons Committee on Trauma.[7] Data were analysed using STATA (version 9.2).[15]

Ethics approval

Ethics approval was obtained from the MSF Ethics Review Board, Geneva, Switzerland, and the Human Research Ethics Committee, University of Cape Town, as well as the Pakistan Bioethics Review Board. Informed consent was obtained from all nurses participating in the study.

Results

Characteristics of the study population

Of a total of 23 nurses carrying out triage, 20 met the study inclusion criteria and were invited to participate in the study. Fifteen of these nurses agreed to participate, while five declined due to scheduling conflicts and transport issues. The convenience sample therefore represented 75% of all eligible triage nurses.

Reliability of nurse triage ratings

A total of 780 ratings were obtained for analysis, consisting of 15 nurses assigning ratings for 42 vignettes (n=630) and the same 15 nurses assigning ratings for the 10 duplicate vignettes (n=150). Table 1 summarises the different reliability measures calculated to assess inter- and intra-rater reliability. Inter-rater reliability, as measured by the ICC and QWK, was substantial. Similarly, the level of exact intra-rater agreement among the nurses in our study was almost perfect (87%; 95% confidence interval (CI) 67 - 100), and there was 100% agreement when allowing for a one-level discrepancy in triage ratings.

Accuracy of nurse triage ratings

Table 2 summarises the accuracy of the nurse acuity ratings using the SATS, compared with the expert panel ratings of the vignettes. Overall, the SATS demonstrated a high level of specificity (97%) and a moderate level of sensitivity (70%). Broken down by acuity level, the SATS showed the highest sensitivity (93%) for 'very urgent' cases. However, the level of sensitivity for 'emergency' cases was exceptionally low (34%). Across all acuity levels, over-triage rates did not exceed the acceptable threshold of 30 - 50%. Similarly, for 'very urgent', 'urgent' and 'routine' cases, under-triage rates were below the acceptable threshold (5 - 10%). However, for emergency cases, the rate of under-triage was exceptionally high (66%), although almost all of these mis-triaged cases were only under-triaged by one acuity level, being rated as 'very urgent'.

Discussion

This is the first study to assess the reliability and accuracy of nurse triage ratings using the SATS in a resource-poor Asian setting.[7] Nurse ratings using this triage scale demonstrated good inter- and intra-rater reliability and acceptable accuracy for 'very urgent' and 'routine' cases. However, nearly two-thirds of 'emergency' cases were under-triaged as 'very urgent', which warrants attention.

Supported by study findings from Botswana and SA,[6,8] our study demonstrates that after minimal formal training, the SATS can be applied reliably by nursing staff in an ED in Pakistan. However, there are concerns about the accuracy of these ratings. In our study, the degree of accuracy of the nurse triage ratings using the SATS was acceptable for 'very urgent' and 'routine' cases, but not for 'urgent' and 'emergency' cases. In particular, a high proportion of emergency cases were under-triaged, which mirrors the findings from a study in SA evaluating the validity of the SATS.[13] The under-triage of 'emergency' cases may be reflected inaccurately on account of several study biases which we discuss below. Alternatively, it may be that this is really the case. If so, this could either be because nursing staff are applying the SATS inaccurately, or because the SATS is poorly constructed to accurately identify true emergency cases. We suspect

Table 1. Different measures calculated to assess inter- and intra-rater reliability of ED nurse triage ratings using the SATS at Timergara Hospital, Pakistan

Reliability measure	Point estimate (95% CI)	Level of agreement*
Inter-rater reliability		
Intra-class correlation coefficient	0.77 (0.69 - 0.85)	Substantial
κ statistic		
Unweighted	0.55 (0.51 - 0.60)	Moderate
Linearly weighted	0.65 (0.61 - 0.71)	Substantial
Quadratically weighted	0.77 (0.69 - 0.84)	Substantial
Intra-rater reliability		
Exact agreement, % (95% CI)	87 (67 - 100)	-
Agreement with one SATS category discrepancy, %	100	-

ED = emergency department; SATS = South African Triage Scale; CI = confidence interval.
*According to the Landis and Koch criteria.[10]

Table 2. Comparison of TH ED nurse ratings using the SATS with the expert panel's ratings of the vignettes

Expert panel triage category	Vignettes, n	Triage ratings, n	Nurse ratings, % (N=630)				SATS performance v. expert panel (reference standard)			
			Emergency	Very urgent	Urgent	Routine	Sensitivity, % (95% CI)	Specificity, % (95% CI)	Over-triage, % (95% CI)	Under-triage, % (95% CI)
Emergency	9	135	34*	0.4	1	0	34 (30 - 38)	99 (99 - 100)	0	66 (62 - 70)
Very urgent	17	255	64	93*	34	6	93 (91 - 95)	97 (95 - 98)	0.4 (0 - 1)	7 (5 - 9)
Urgent	10	150	2	4	59*	17	59 (55 - 63)	94 (93 - 96)	35 (33 - 37)	7 (5 - 9)
Routine	6	90	0	3	7	78*	78 (75 - 81)	97 (96 - 98)	22 (19 - 25)	0
Mean							70 (66 - 74)	97 (92 - 100)	15 (4 - 25)	22 (9 - 34)

TH = Timergara Hospital; ED = emergency department; SATS = South African Triage Scale; CI = confidence interval.
* Nurse ratings matching the expert panel's rating (reference standard) across each acuity level.

that staff inaccuracy is not to blame, as regular audits of the SATS in Pakistan together with the findings from a previous study have shown a high level of staff accuracy.[13]

If the construct of the SATS itself is responsible for the under-triage of 'emergency' cases, this needs further investigation. The clinical implications of under-triage of 'emergency' cases in our setting are negligible as almost all of the 'under-triaged' emergency cases were rated as 'very urgent', and in the context of TH all 'emergency' and 'very urgent' patients are seen by the same cadre of healthcare workers in the same area and within the same timeframe. Although we do not have data to substantiate this, a 10-minute delay linked to misclassification of 'emergency' to 'very urgent' cases is unlikely to have clinical implications. Nonetheless, in a setting where there are clear distinctions between the ways in which 'emergency' and 'very urgent' patients are managed, under-triage in this way needs to be avoided, as it may be associated with poorer outcomes (i.e. a higher risk of mortality, worsening morbidity and additional medical complications). This makes the case for ensuring that any assessment of the SATS is context specific.

Study limitations

A number of study limitations and various methodological issues related to assessing the validity of a triage tool have been brought to our attention by this study.

First, while there is no universally accepted time period recommended between assessments for inter- and intra-rater reliability, 2 - 14 days has been suggested.[10] Owing to ED staff time constraints, we conducted the intra-rater assessments immediately after the inter-rater assessment; this may have led to a recall bias in the response ratings.

Second, although the vignettes were paper based, in the absence of non-verbal patient cues and contextual information, raters' triage decisions may have been affected. That said, a previous study comparing the use of paper-based cases with live ED patients as a way of assessing the inter-rater reliability of a triage tool showed an acceptable level of agreement between the two methods.[16] The main benefits of using paper-based vignettes over real ED cases in LMIC settings is that they provide a cost-effective, time-saving, non-invasive and culturally acceptable way of undertaking this type of study.

Third, the written vignettes were based on ED cases seen in SA, not in the TH ED in Pakistan. In the study by Twomey *et al.*,[13] a set of vignettes ratified by a modified Delphi technique are proposed as a set of reference standard vignettes. Using these vignettes in Pakistan was deemed appropriate due to the following: (*i*) SA and Pakistan are both LMIC settings; (*ii*) the two settings have similar rates of trauma (66 trauma presentations per 1 000 patients in SA and 41/1 000 in Pakistan);[17,18] and (*iii*) the reference vignettes depict similar case presentations. However, the epidemiological pattern of disease is different. In future studies like this, it would seem important to develop specific reference vignettes based on ED cases seen in the actual study setting. This would ensure the use of a better reference standard of comparison adapted to the study context.

Fourth, when comparing nurse acuity ratings using the SATS to acuity ratings assigned by the expert panel, we cannot be sure whether an identified discrepancy between the two was: (*i*) because the nursing staff were not applying the SATS accurately; or (*ii*) because the SATS had poor construct validity – in other words did not measure what it purports to. As indicated earlier, we suspect that staff inaccuracy did not account for many of the observed discrepancies in this study. However, in future studies assessing the validity of a triage tool, it would be more appropriate to compare the ratings by several SATS experts (using the SATS) to the expert panel ratings (reference standard). This would help to control for the issue of staff error.

Finally, as in other studies, our reference standard was an expert panel that assigned acuity ratings to a series of paper vignettes according to their expert opinion. Almost all of these experts were based in high-income rather than LMIC settings and as such their opinion of 'true' patient acuity level may not have fully reflected the reality as in LMIC settings like Pakistan – they may have tended to over-rate patient acuity, especially at the higher end of the triage spectrum. In conjunction with this, it has been reported that nurses tend to under-rate patient acuity when using paper-based vignettes over live cases.[16] In our particular study, these two factors may have contributed to the under-triage of emergency cases that was reported.

Conclusion

Our study shows that the SATS can be used reliably by nurses in an ED in Pakistan. Our results suggest that the SATS is accurate for very urgent and routine cases but, importantly, may 'under-triage' 'emergency' cases. Although this is unlikely to influence patient outcomes in TH, there may be serious implications in other settings and it therefore merits specific investigation and correction.

Acknowledgements. We are grateful to the Pakistani Ministry of Health for their collaboration, and we are particularly grateful to the staff in the field for their hard work. The MSF project in TH, Pakistan, is funded by MSF-Operational Centre Brussels.

References

1. Van Rooyen M, Venugopal R, Greenough PG. International humanitarian assistance: Where do emergency physicians belong? Emerg Med Clin North Am 2005;23(1):115-131. [http://dx.doi.org/10.1016/j.emc.2004.09.006]
2. Hodkinson PW, Wallis LA. Emergency medicine in the developing world: A Delphi study. Acad Emerg Med 2010;17(7):765-774. [http://dx.doi.org/10.1111/j.1553-2712.2010.00791.x]
3. Horne S, Vassallo J, Read J, Ball S. UK triage – an improved tool for an evolving threat. Injury 2013;44(1):23-28. [http://dx.doi.org/10.1016/j.injury.2011.10.005]
4. Robison JA, Ahmad ZP, Nosek CA, et al. Decreased pediatric hospital mortality after an intervention to improve emergency care in Lilongwe, Malawi. Pediatrics 2012;130(3):e676-e682. [http://dx.doi.org/10.1542/peds.2012-0026]
5. Emergency Medicine Society of South Africa. The South African Triage Scale (SATS). http://emssa.org.za/sats/ (accessed 11 March 2014).
6. Twomey M, Wallis LA, Thompson ML, Myers JE. The South African Triage Scale (adult version) provides valid acuity ratings when used by doctors and enrolled nursing assistants. African Journal of Emergency Medicine 2012;2(1):3-12. [http://dx.doi.org/10.1016/j.afjem.2011.08.014]
7. Dalwai M, Tayler-Smith K. Implementation of a triage score system in an emergency room in Timergara, Pakistan. Public Health Action 2013;3(1):43-45. [http://dx.doi.org/10.5588/pha.12.0083]
8. Twomey M, Mullan PC, Torrey SB, Wallis L, Kestler A. The Princess Marina Hospital accident and emergency triage scale provides highly reliable triage acuity ratings. Emerg Med J 2012;29(8):650-653. [http://dx.doi.org/10.1136/emermed-2011-200503]
9. Harrison H-L, Raghunath N, Twomey M. Emergency triage, assessment and treatment at a district hospital in Malawi. Emerg Med J 2012;29(11):924-925. [http://dx.doi.org/10.1136/emermed-2011-200472]
10. Streiner D, Norman G. Health Measurement Scales. A Practical Guide to Their Development and Use. 4th ed. New York: Oxford University Press, 2008.
11. Twomey M, Wallis LA, Myers JE. Limitations in validating emergency department triage scales. Emerg Med J 2007;24(7):477-479. [http://dx.doi.org/10.1136/emj.2007.046383]
12. Olofsson P, Gellerstedt M, Carlström ED. Manchester Triage in Sweden – interrater reliability and accuracy. Int Emerg Nurs 2009;17(3):143-148. [http://dx.doi.org/10.1016/j.ienj.2008.11.008]
13. Twomey M, Wallis L, Myers J. Evaluating the construct of triage acuity against a set of reference vignettes developed via modified Delphi method. Emerg Med J 2013. [http://dx.doi.org/10.1136/emermed-2013-202352]
14. Kottner J, Audigé L, Brorson S, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. J Clin Epidemiol 2011;64(1):96-106. [http://dx.doi.org/10.1016/j.jclinepi.2010.03.002]
15. StataCorp. 2005. Stata Statistical Software: Release 9. College Station, TX: StataCorp LP.
16. Worster A, Sardo A, Eva K, Fernandes CMB, Upadhye S. Triage tool inter-rater reliability: A comparison of live versus paper case scenarios. J Emerg Nurs 2007;33(4):319-323. [http://dx.doi.org/10.1016/j.jen.2006.12.016]
17. Wallis LA, Twomey M. Workload and casemix in Cape Town emergency departments. S Afr Med J 2007;97(12):1276-1280.
18. Nasrullah M, Xiang H. The epidemic of injuries in Pakistan – a neglected problem. J Pak Med Assoc 2008;58(8):420-421.

Accepted 7 January 2014.

Supplemental Discussion

This study used a standard set of patient vignettes to assess reliability and accuracy of the SATS for the first time in a context outside of South Africa. Twomey et al. created the reference set of vignettes in order to standardise and simplify reliability testing of the SATS in low-resource settings.[69] Prior studies have used local patient presentations as a basis for vignette composition.[21,73]

Two key learning areas came to light in relation to the use of these reference vignettes. These are discussed in the paper, but expanded upon further in this chapter.

1. The reference vignettes used in this study were developed in South Africa and were based on commonly seen cases in South African ECs.[69] As such, they did not fully represent all the common EC cases seen in an EC in Northern Pakistan. For example, alcohol intoxication is frequently observed in many South African EC presentations. However, given the cultural setting of Northern Pakistan, alcohol intoxication is rarely seen and as such would have been an unfamiliar case presentation for the triage nurse, and thus potentially more difficult to interpret. We believe that in future studies of this nature it would be more beneficial to develop a set of vignettes based on local cases presenting to the EC over the previous months.

2. Context-specific language and presenting complaints used to create the vignettes did not correspond to local case presentations. Even with specific translation, we found that particular meanings could be understood differently based on cultural interpretation of vocabulary when describing presenting complaints and their inherited definition. Moving forward we would suggest that vignettes are developed using terminology that is commonly used in each specific cultural setting.

Chapter Conclusion

The SATS demonstrated good reliability in the EC setting of Northern Pakistan. However, in the process of assessing the reliability and accuracy of the SATS in this setting, a much better understanding of the difficulties and limitations of trying to assess these particular measures was facilitated. This in itself helped us to devise more robust ways to assess the reliability of the SATS in Afghanistan, Sierra Leone and Haiti, thereafter (Chapter 6). One clear limitation is that of using standard reference vignettes versus vignettes based on local cases. We believe that given the limitations above, the use of local cases should improve the assessment of reliability. Our recommendation is therefore that local cases be used as a basis to create vignettes, with careful attention to the use of local vernacular when describing common presentations. Previous studies have used a similar method of 'expert opinion' when addressing accuracy, which in some studies they have referred to as validity.[72,73] In Chapter 7, we discuss this in further detail and suggest a modified approach when using outcome markers to assess validity.

In the next two chapters we assess reliability and validity of the SATS across multiple settings.

Chapter 6: Inter- and intra-rater reliability of SATS in low-resource settings

Reference

Dalwai MK, Tayler-Smith K, Twomey M, Nasim M, Popal AQ, Haqdost WH, Gayraud O, Cherestal S, Wallis L, Valles P. The inter and intra-rater reliability of the South African triage scale in low resource settings of Haiti and Afghanistan. **In Press**

Declaration from author

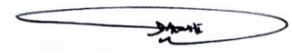
The following co-authors contributed to the paper:
Mrs. Katie Tayler-Smith, Dr. Michele Twomey, Dr. Abdul Qayum Popal, Dr Masood Nasim, Mr. Waliul Haq Haqdost, Dr. Olivia Gayraud, Dr Sophia Cherestal, Prof. Lee Wallis and Dr. Pola Valles
In the case of Chapter 6, the nature and extent of contribution by the authors was as follows:

Nature of contribution

MKD, MT, LW and PV contributed to study conception and design; MKD, MN, AQP, WHH, OG, SC and PV facilitated the acquisition of data; MKD, LW, KTS, PV, OG and AQP helped with the analysis and interpretation of the data; the initial version was drafted by MKD, KTS and PV with all authors contributing to the redrafts and revising it critically for important intellectual content; all authors gave final approval on the version submitted and the revisions thereof.

Extent of contribution

MKD: 80%; MT, LW together: 5%; MN, AQP, WHH, OG, SC together: 10%; PV and KTS together 5%



Signed: Mohammed. K. Dalwai 19 Oct 2017

Declaration from Co-Authors

The undersigned hereby certify that:

1. The above declaration correctly reflects the nature and extent of the candidate's contribution to this work, and the nature of the contribution of each of the co-authors.
2. They meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
3. They take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
4. There are no other author of the publication according these criteria;
5. Potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and
6. The original data are stored in a two-factor password protected cloud account and will be held for at least five years

Prof. Lee Wallis 02 Sept 2017

Dr Michele Twomey 02 Sept 2017

Mrs Katie Tayler-Smith 02 Sept 2017

Introduction to the topic

The reliability of the SATS has yet to be tested across a range of low-resource settings. In triage, inter-rater reliability reflects the ability of a triage tool to yield the same result when used by different clinicians on the same patient. This is an important measure of the effectiveness of a triage scale. So too, its ability to yield the same result when used by the same clinician for the same patient over different time periods (intra-rater reliability).

Motivation for conducting the study

The SATS demonstrated good preliminary results when assessed in Northern Pakistan (Chapter 5) but the study called for further assessment of inter- and intra-rating reliability. We set out to measure the reliability of the SATS across multiple low-resource settings, in order to determine its feasibility in such settings.

Aim

The aim of this study was to investigate the inter- and intra-rater reliability of the adult version of the SATS across low-resource settings.

Objectives

- To test the inter- and intra-rater reliability of the SATS when used by emergency centre nurses in Afghanistan, Haiti and Sierra Leone.

Main Findings

- The SATS is moderately reliable across multiple low-resource settings.

Copy of published paper

Original article

OPEN ACCESS

Inter-rater and intrarater reliability of the South African Triage Scale in low-resource settings of Haiti and Afghanistan

Mohammed Dalwai,[1,2] Katie Tayler-Smith,[3] Michèle Twomey,[1] Masood Nasim,[4] Abdul Qayum Popal,[4] Waliul Haq Haqdost,[5] Olivia Gayraud,[6] Sophia Cheréstal,[6] Lee Wallis, Pola Valles[2]

[1]Division of Emergency Medicine, University of Cape Town, Cape Town, South Africa
[2]Medical Department, Médecins Sans Frontières, Operational Centre Brussels, Brussels, Belgium
[3]Operational Research Unit Luxembourg, Médecins Sans Frontières, Luxembourg City, Luxembourg
[4]Medical Department, Médecins Sans Frontières, Kabul, Afghanistan
[5]Ministry of Health, Kabul, Afghanistan
[6]Medical Department, Médecins Sans Frontières, Port au Prince, Haiti

Correspondence to
Dr Mohammed Dalwai, Division of Emergency Medicine, University of Cape Town, Cape Town 7708, South Africa; mkdalwai@gmail.com

Received 18 July 2017
Revised 6 February 2018
Accepted 13 February 2018

To cite: Dalwai M, Tayler-Smith K, Twomey M, *et al*. *Emerg Med J* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/emermed-2017-207062

ABSTRACT

Objective The South African Triage Scale (SATS) has demonstrated good validity in the EDs of Médecins Sans Frontières (MSF)-supported sites in Afghanistan and Haiti; however, corresponding reliability in these settings has not yet been reported on. This study set out to assess the inter-rater and intrarater reliability of the SATS in four MSF-supported EDs in Afghanistan and Haiti (two trauma-only EDs and two mixed (including both medical and trauma cases) EDs).

Methods Under classroom conditions between December 2013 and February 2014, ED nurses at each site assigned triage ratings to a set of context-specific vignettes (written case reports of ED patients). Inter-rater reliability was assessed by comparing triage ratings among nurses; intrarater reliability was assessed by asking the nurses to retriage 10 random vignettes from the original set and comparing these duplicate ratings. Inter-rater reliability was calculated using the unweighted kappa, linearly weighted kappa and quadratically weighted kappa (QWK) statistics, and the intraclass correlation coefficient (ICC). Intrarater reliability was calculated according to the percentage of exact agreement and the percentage of agreement allowing for one level of discrepancy in triage ratings. The correlation between years of nursing experience and reliability of the SATS was assessed based on comparison of ICCs and the respective 95% CIs.

Results A total of 67 nurses agreed to participate in the study: In Afghanistan there were 19 nurses from Kunduz Trauma Centre and nine from Ahmed Shah Baba; in Haiti, there were 20 nurses from Martissant Emergency Centre and 19 from Tabarre Surgical and Trauma Centre. Inter-rater agreement was moderate across all sites (ICC range: 0.50–0.60; QWK range: 0.50–0.59) apart from the trauma ED in Haiti where it was moderate to substantial (ICC: 0.58; QWK: 0.61). Intrarater agreement was similar across the four sites (68%–74% exact agreement); when allowing for a one-level discrepancy in triage ratings, intrarater reliability was near perfect across all sites (96%–99%). No significant correlation was found between years of nursing experience and reliability.

Conclusion The SATS has moderate reliability in different EDs in Afghanistan and Haiti. These findings, together with concurrent findings showing that the SATS has good validity in the same settings, provide evidence to suggest that SATS is suitable in trauma-only and mixed EDs in low-resource settings.

Key messages

What is already known on this subject

- There are few triage scales designed specifically for use in low/middle-income countries (LMIC); the South African Triage Scale (SATS) is one of them and has been shown to have good validity in such settings. The inter-rater reliability of SATS in South Africa has been reported as moderate to substantial, with intrarater reliability ranging from 80% to 86%. Its performance across a spectrum of different LMIC settings, mainly non-sub-Saharan African and trauma-only settings, has not been adequately evaluated.

What this study adds

- In this cross-sectional study using case vignettes, ED nurses in Afghanistan and Haiti assigned triage ratings using the SATS. Inter-rater reliability was moderate and intrarater reliability for exact agreement ranged from 68% to 74%. Added to evidence showing good validity of this scale, this suggests the SATS could be suitable for low-resource settings.

INTRODUCTION

Triage has a central role in emergency care systems: prioritising patients based on acuity improves effective use of resources, and ultimately patient outcomes.[1] A number of different scales exist for in-hospital use, but most of these have been developed for and evaluated in high-resource settings.[2 3] Context-appropriate triage tools for low/middle-income countries (LMIC) are very uncommon.[4] Among the few tools that have been contextually modified, validated and implemented in various settings is the South African Triage Scale (SATS), which was developed for in-hospital EDs.[5] The SATS has been assessed extensively in South Africa and implemented in several settings,[6–8] but further assessment of its performance in low-resource settings, particularly non-sub-Saharan settings, is still needed.[4 9]

For a triage scale to be effective, it needs to demonstrate good validity (ie, an acuity rating assigned using the scale must closely reflect a patient's true acuity) and a high degree of reliability

Table 1 Characteristics of the study sites in Afghanistan and Haiti

	Ahmad Shah Baba	Kunduz Trauma Centre	Martissant Emergency Centre	Tabarre Surgical and Trauma Centre
Country	Afghanistan	Afghanistan	Haiti	Haiti
Location	Kabul City, district 12	Kunduz Province	Port-au-Prince, Martissant district	Port-au-Prince, Tabarre district
Estimated catchment population	219 000	1 000 000	1 200 000	1 000 000
Level of MSF support	Partnership with Ministry of Health	MSF only	MSF only	MSF only
Services offered	OPD ED Maternity IPD: surgery, internal medicine, paediatric	OPD ED IPD trauma care: surgery, orthopaedic ICU, physiotherapy	ED only	OPD ED IPD visceral and trauma care: surgery, orthopaedic ICU, physiotherapy
Type of ED cases	Mixed	Trauma only	Mixed	Visceral surgery and trauma
ED caseload (per month)*	4715	1848	4919	793
Introduction of the SATS	2011	2011	2013	2012

*Mean cases seen per month during 2014.
ICU, intensive care unit; IPD, inpatient department; MSF, Médecins Sans Frontières; OPD, outpatient department; SATS, South African Triage Scale.

(ie, it must yield the same triage rating on repeated assessments of the same patient). For any given patient, tools should have high inter-rater (the degree of variability among different nurses) and intrarater (the variability of retriage ratings for one nurse) reliability.

Médecins Sans Frontières (MSF), an international medical humanitarian organisation, provides free medical care to vulnerable populations in many LMIC settings. It operates within constrained resources and serves populations with little healthcare access. Since 2011, MSF-Operational Centre Brussels has implemented the SATS in projects where it provides emergency care. The validity of the SATS was recently assessed in the EDs of MSF-supported sites in Afghanistan and Haiti[10] with good results, but corresponding reliability in these sites has yet to be reported on. This is the basis of the current study.

METHODS

Study design

This was a cross-sectional study using a set of ED vignettes (short written clinical case reports of actual ED patients) as a proxy for live patients, in which ED nurses assigned triage ratings using the SATS.

Study setting

The study was conducted at four active MSF project sites between December 2013 and February 2014: two hospitals in Afghanistan (Ahmad Shah Baba (ASB) and Kunduz Trauma Centre (KTC)) and two facilities in Haiti (Martissant Emergency Centre (MT) and Tabarre Surgical and Trauma Centre (TB)). Specific details on these four sites are summarised in table 1.

SATS and its use in the ED

Described in detail elsewhere,[10] the SATS is a four-tiered triage tool which depicts a patient's urgency for care using the following colour codes: priority 1: red—'emergency' (to be seen immediately); priority 2: orange—'very urgent' (to be seen within 10 min); priority 3: yellow—'urgent' (to be seen within 60 min); priority 4: green—'routine' (to be seen within 240 min). The SATS also allocates the colour blue (black was used in the study countries for cultural purposes) to 'dead on arrival' cases.

Study population

The study included all ED nurses at the four study sites who fulfilled the following inclusion criteria: (1) had received training in use of the SATS and (2) agreed to participate in the study. All nurses employed by MSF have a basic nursing degree and are registered with the country nursing authority.

Study protocol

Under classroom conditions, all nurses who agreed to participate in the study were asked to use the SATS to triage a set of vignettes and assign one of the following four categories to each vignette: 'emergency', 'very urgent', 'urgent' and 'routine'. Each set comprised between 28 and 30 vignettes generated from information extracted from randomly selected patient files of real ED cases who had presented at the study centres between June and December 2013. Each vignette included information on patient gender, age, presenting complaint, mode of arrival to the ED and vital signs. All clinical information in the triage paperwork was copied into the vignettes including information from additional investigations such as blood glucose and haemoglobin levels (see box 1 for an example of a vignette).

Professionals translated the vignettes from English into the relevant local languages. Local bilingual doctors ratified the translations to ensure correct medical terminology.

Under classroom conditions, all nurses who agreed to participate in the study assigned one of four SATS categories to the set of reference vignettes.

Data analysis

Inter-rater reliability was measured by comparing the triage ratings assigned for each of the vignettes by different nurses at each

Box 1 Example of a vignette used to assess the South African Triage Scale (SATS) in Afghanistan and Haiti, 2013

A 17-year-old boy presents with abdominal pain, loose motion and vomiting since this morning. He says he ate something last night that did not agree with his stomach and since this morning has not been feeling well. At triage, you find an alert boy with moderate abdominal pain. No signs of dehydration are present.

BP: 120/80; HR: 109; RR: 16; temperature: 36°C

Table 2 Nurses' response rate at each study site in Afghanistan and Haiti, 2013

Study site	Nurses invited to participate (n)	Nurses agreeing to participate (n)	Response rate (%)
KTC (Afghanistan)	21	19	90
ASB (Afghanistan)	9	9	100
MT (Haiti)	21	20	95
TB (Haiti)	20	19	95

ASB, Ahmad Shah Baba; KTC, Kunduz Trauma Centre; MT, Martissant Emergency Centre; TB, Tabarre Surgical and Trauma Centre.

study site. Intrarater reliability was measured by asking nurses to retriage 10 random vignettes from the original set 1–10 days later (depending on their availability), and comparing these duplicate ratings.

In accordance with the Guidelines for Reporting Reliability and Agreement Studies (GRRAS), inter-rater reliability was assessed using the unweighted kappa (UWK), linearly weighted kappa (LWK) and quadratically weighted kappa (QWK) statistics, as well as the intraclass correlation coefficient (ICC).[11] UWK and LWK point estimates were assessed and included as per GRRAS guidelines, but in keeping with triage literature we only interpreted QWK and ICC point estimates using the Landis and Koch classification system: 0.0–0.20—slight agreement; 0.21–0.40—fair agreement; 0.41–0.60—moderate agreement; 0.61–0.80—substantial agreement; 0.81–1.00—almost perfect agreement.[12] In triage reliability studies, UWK and LWK can be ignored. QWK and ICC yield almost identical results hence either one could be used based on ease of calculation.[12]

Intrarater reliability was assessed by calculating both the percentage of exact agreement and the percentage of agreement allowing for one level of discrepancy in triage ratings. 95% CIs were calculated for all measures.

In addition, we assessed whether there was any correlation between years of nursing experience (ie, years of being a qualified nurse) and the ICC based on comparison of the 95% CIs and use of bootstrapping.

RESULTS

Study population

Table 2 shows the sample size at each study site. The response rate ranged from 90% in KTC to 100% in ASB.

Reliability of nurse triage ratings

Table 3 summarises the different reliability measures calculated to assess inter-rater and intrarater reliability across the four study sites. Inter-rater agreement was moderate across all sites, apart from TB where it was moderate to substantial. Trauma-only facilities (KTC and TB) yielded very similar results (ICC: 0.60 and 0.58 and QWK: 0.59 and 0.61, respectively) whereas among the mixed settings (ASB and MT), there was a wider variability in results (ICC: 0.50 and 0.59 and QWK: 0.50 and 0.59, respectively).

Intrarater agreement was similar across the four sites, ranging from 68% exact agreement in ASB to 74% in MT. When allowing for a one-level discrepancy in triage ratings, intrarater reliability was near perfect across all sites ranging from 96% in TB and ASB to 99% in KTC.

Table 4 shows the correlation between years of nursing experience and ICC across the four sites. The mean years of nursing experience were similar across all sites ranging (6.3–7.1 years). The ICC for nurses with 5 or more years of nursing experience appeared to be higher than for those with less than 5 years of experience, but 95% CIs overlapped (even after applying a bootstrapping technique) indicating no statistical significant difference.

DISCUSSION

Our study shows that the SATS has moderate inter-rater and intrarater reliability when used by nurses in trauma-only and mixed ED settings in Afghanistan and Haiti. This is evidenced to suggest that the SATS could be suitable for use in low-resource settings. Further reliability studies in low-resource settings are needed to confirm these findings.

The main strengths of this study are its multisite nature, the high response rate of participants and the fact that the vignettes reflected real ED cases seen in each specific setting. In previous studies assessing the SATS in contexts outside of South Africa, the vignettes used were based on South African ED cases, not ED cases specific to the study setting.[8 9]

Limitations

There were a number of study limitations. First, using paper-based vignettes as a proxy for real ED cases has the inherent limitation of not mimicking real life.[9] Although conducting consecutive live triage assessments on a single patient at one point in time and at multiple points in time is not feasible or practical,[13] use of paper-based vignettes assessed under

Table 3 Inter-rater and intrarater reliability measures for the SATS in Afghanistan and Haiti, 2013

Inter-rater reliability measures	Ahmad Shah Baba	Kunduz Trauma Centre	Martissant Emergency Centre	Tabarre Surgical and Trauma Centre
Point estimates (95% CI)				
ICC	0.50 (0.37 to 0.66)	0.60 (0.48 to 0.74)	0.59 (0.46 to 0.73)	0.58 (0.44 to 0.73)
QWK	0.50 (0.37 to 0.66)	0.59 (0.22 to 0.77)	0.59 (0.32 to 0.78)	0.61 (0.32 to 0.79)
LWK	0.44 (0.16 to 0.69)	0.46 (0.21 to 0.66)	0.50 (0.27 to 0.70)	0.48 (0.23 to 0.68)
UWK	0.40 (0.26 to 0.53)	0.33 (0.23 to 0.41)	0.41 (0.31 to 0.48)	0.35 (0.23 to 0.47)
Level of agreement*				
ICC	Moderate	Moderate	Moderate	Moderate
QWK	Moderate	Moderate	Moderate	Substantial
Intrarater reliability measures, % (min-max)				
Mean exact agreement	68 (40–90)	71 (40–100)	74 (40–90)	73 (50–90)
Mean agreement with one-degree discrepancy	96 (80–100)	99 (90–100)	97 (60-100)	96 (70-100)

*According to the Landis and Koch criteria.[11]
ICC, intraclass correlation coefficient; LWK, linearly weighted kappa; QWK, quadratically weighted kappa; SATS, South African Triage Scale; UWK, unweighted kappa.

Table 4 Effect of nurse experience on inter-rater reliability of the SATS

Study site	Nurses agreeing to participate (n)	Mean years of experience	Nurses with <5 years of experience		Nurses with ≥5 years of experience	
			n	ICC	n	ICC
KTC (Afghanistan)	19	6.6	11	0.53 (0.38–0.68)	9	0.67 (0.54–0.79)
ASB (Afghanistan)	8*	7.1	5	0.46 (0.29–0.64)	3	0.55 (0.35–0.74)
MT (Haiti)	20	6.3	11	0.53 (0.38–0.68)	9	0.64 (0.50–0.78)
TB (Haiti)	19	6.3	6	0.70 (0.57–0.83)	13	0.57 (0.43–0.71)

*Information available on eight out of the nine nurses.
ASB, Ahmad Shah Baba; ICC, intraclass correlation coefficient; KTC, Kunduz Trauma Centre; MT, Martissant Emergency Centre; SATS, South African Triage Scale; TB, Tabarre Surgical and Trauma Centre.

classroom conditions may have influenced the relative degree of reliability that was observed. For example, the wording of the vignettes may have been interpreted differently by different nurses. That said, a previous study has shown that there is little difference between the inter-rater reliability measures generated using paper-based cases compared with live cases.[13] Second, translation of the vignettes from English into the local language may have slightly distorted some of the original information. We tried to limit this by recruiting professional translators with some medical background to carry out the translations in each setting, and having local medical staff back translate.

Originally developed for use in South Africa, the SATS has been assessed extensively in South Africa, and also in Botswana, Malawi and Pakistan with good results.[4 7 8 14] But the degree to which these findings are applicable to other LMIC settings—particularly those outside of sub-Saharan Africa and those that deal with trauma-only caseloads—has remained unclear. This is what prompted a recent study assessing the validity of the SATS in different EDs in Afghanistan and Haiti.[10] The results of this study were good, but reliability in these settings was still unknown. Reliability of triage across both high-resource and low-resource settings varies greatly. Two articles assessing reliability in South Africa report moderate to substantial reliability with QWK of 0.57 and 0.66, respectively.[14 15] In Ghana, the SATS showed moderate reliability with QWK of 0.59 and 0.60[16] while studies in Pakistan and Botswana reported substantial to near-perfect results with QWK of 0.77 and 0.87, respectively.[8 9] In high-resource settings, the Canadian Emergency Department Triage and Acuity Scale (CTAS), a 5-level triage scale, reported a chance corrected kappa of 0.80 and a weighted kappa of 0.77.[17 18] The Emergency Severity Index (ESI) has reported inter-rater reliability ranging from 0.76 to 0.8 with the Manchester Triage System (MTS) showing a weighted kappa from 0.62 to 0.82.[2 19 20] No studies were found in low-resource settings for either the CTAS or MTS. The ESI was implemented in Iran but according to Mirhaghi *et al* may not reveal optimal outcomes for LMICs.[21] Standardisation of reporting reliability is poor with some studies not identifying which weighted kappa statistic was used to calculate reliability, making comparisons between studies difficult.[11] The one-two-triage scale, the only other new scale developed in 2015 for low-resource settings, reported a kappa of 0.308 among nurses in Cambodia.[22]

The results of our study confirm that the SATS is valid in Haiti and Afghanistan and demonstrates moderate reliability. This latter finding is most certainly a reflection of the relative simplicity of the SATS, both in terms of its construct and application, and supports its value in resource-constrained settings where highly skilled staff are often in short supply.

Reliable use of the SATS did appear to be higher among nurses with 5 or more years of nursing experience, although our results were not statistically significant. The latter however may be related to our relatively small sample size, and thus low statistical power. This finding is similar to previous research by Göransson *et al* that found no significant difference between nursing experience and reliability of triage when using the Canadian Triage and Acuity Scale.[18]

In addition, there may be other factors that influence reliability and which confounded the relationship between years of experience and reliable use of the SATS, for example, how regularly the nurses were working in triage (all the nurses were working on a rotational basis and therefore were not permanently based in the ED).

It would be useful to explore these sorts of factors further in order to establish how they affect reliability and ultimately what could be done to optimise the reliable use of the SATS.

CONCLUSION

In conclusion, our study shows that the SATS is a moderately reliable tool for use in different EDs in Afghanistan and Haiti. These findings, together with concurrent findings showing that the SATS has good validity in the same settings, provide evidence to suggest that SATS is suitable in trauma-only and mixed EDs in low-resource settings.

Acknowledgements We thank all the staff who participated in the study both nationally and internationally for going above and beyond and always trying to improve healthcare for vulnerable populations. We also thank the Centre for Evidence-based Health Care at the University of Stellenbosch for help with the statistics.

Contributors MD, PV, MT, LW and KTS designed, analysed and interpreted the study and data. AQP, WHH and MN were the project leads in Afghanistan. OG and SC were the project leads in Haiti. All authors contributed to the revision of the final article.

Funding The degree from which this study emanated was funded by the South African Medical Research Council under the SAMRC Clinicians Researcher Development Scholarship PhD programme.

Competing interests None declared.

Patient consent Not required.

Ethics approval Ethics approval was obtained from the National Ethics Committees in Afghanistan and Haiti, from the MSF Ethics Review Board and from the University of Cape Town (UCT).

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Data sharing is available on request.

REFERENCES

1 Robison JA, Ahmad ZP, Nosek CA, *et al*. Decreased pediatric hospital mortality after an intervention to improve emergency care in Lilongwe, Malawi. *Pediatrics* 2012;130:e676–82.
2 Wuerz RC, Milne LW, Eitel DR, *et al*. Reliability and validity of a new five-level triage instrument. *Acad Emerg Med* 2000;7:236–42.
3 van der Wulp I, van Baar ME, Schrijvers AJ. Reliability and validity of the Manchester Triage System in a general emergency department patient population in the Netherlands: results of a simulation study. *Emerg Med J* 2008;25:431–4.
4 Dalwai MK, Tayler-Smith K, Trelles M, *et al*. Implementation of a triage score system in an emergency room in Timergara, Pakistan [Short communication]. *Public Health Action* 2013;3:43–5.
5 Gottschalk SB, Wood D, DeVries S, *et al*. The cape triage score: a new triage system South Africa. Proposal from the cape triage group. *Emerg Med J* 2006;23:149–53.
6 Twomey M, Wallis LA, Myers JE. Evaluating the construct of triage acuity against a set of reference vignettes developed via modified Delphi method. *Emerg Med J* 2014;31:562–6.
7 Harrison HL, Raghunath N, Twomey M. Emergency triage, assessment and treatment at a district hospital in Malawi. *Emerg Med J* 2012;29:924–5.
8 Twomey M, Mullan PC, Torrey SB, *et al*. The Princess Marina Hospital accident and emergency triage scale provides highly reliable triage acuity ratings. *Emerg Med J* 2012;29:650–3.
9 Dalwai MK, Twomey M, Maikere J, *et al*. Reliability and accuracy of the South African Triage Scale when used by nurses in the emergency department of Timergara Hospital, Pakistan. *S Afr Med J* 2014;104:372.
10 Dalwai M, Valles P, Twomey M, *et al*. Is the South African Triage Scale valid for use in Afghanistan, Haiti and Sierra Leone? *BMJ Glob Health* 2017;2:e000160.
11 Kottner J, Audigé L, Brorson S, *et al*. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *J Clin Epidemiol* 2011;64:96–106.
12 Streiner DL, Norman GR, Cairney J. *Health measurement scales: a practical guide to their development and use*. 5th ed. New York: Oxford University Press, 2015. doi.
13 Worster A, Sardo A, Eva K, *et al*. Triage tool inter-rater reliability: a comparison of live versus paper case scenarios. *J Emerg Nurs* 2007;33:319–23.
14 Twomey M, Wallis LA, Thompson ML, *et al*. The South African Triage Scale (adult version) provides reliable acuity ratings. *Int Emerg Nurs* 2012;20:142–50.
15 Twomey M, de Sá A, Wallis LA, *et al*. Inter-rater reliability of the South African Triage Scale: Assessing two different cadres of health care workers in a real time environment. *African J Emerg Med* 2011;1:113–8.
16 Gyedu A, Agbedinu K, Dalwai M, *et al*. Triage capabilities of medical trainees in Ghana using the South African triage scale: an opportunity to improve emergency care. *Pan Afr Med J* 2016;24.
17 Manos D, Petrie DA, Beveridge RC, *et al*. Inter-observer agreement using the Canadian Emergency Department Triage and Acuity Scale. *CJEM* 2002;4:16–22.
18 Beveridge R, Ducharme J, Janes L, *et al*. Reliability of the Canadian emergency department triage and acuity scale: interrater agreement. *Ann Emerg Med* 1999;34:155–9.
19 Parenti N, Reggiani ML, Iannone P, *et al*. A systematic review on the validity and reliability of an emergency department triage scale, the Manchester Triage System. *Int J Nurs Stud* 2014;51:1062–9.
20 Eitel DR, Travers DA, Rosenau AM, *et al*. The emergency severity index triage algorithm version 2 is reliable and valid. *Acad Emerg Med* 2003;10:1070–80.
21 Mirhaghi A, Kooshiar H, Esmaeili H, *et al*. Outcomes for emergency severity index triage implementation in the emergency department. *J Clin Diagn Res* 2015;9:OC04–OC07.
22 Khan A, Mahadevan SV, Dreyfuss A, *et al*. One-two-triage: validation and reliability of a novel triage system for low-resource settings. *Emerg Med J* 2016;33:709–15.

Inter-rater and intrarater reliability of the South African Triage Scale in low-resource settings of Haiti and Afghanistan

Mohammed Dalwai, Katie Tayler-Smith, Michèle Twomey, Masood Nasim, Abdul Qayum Popal, Waliul Haq Haqdost, Olivia Gayraud, Sophia Cheréstal, Lee Wallis and Pola Valles

Emerg Med J published online March 16, 2018

Supplemental Discussion

In the paper above, we report on the reliability of the adult version of the SATS in Afghanistan and Haiti only, excluding the reliability of Sierra Leone. Data collection for Chapters 6 and 7 were done simultaneously but the validity results was reported and published first. In the next chapter, Chapter 7, we describe the validity in Haiti, Afghanistan and Sierra Leone. In Chapter 7, validity was found to be acceptable in Haiti and Afghanistan, in the paediatric only setting of Sierra Leone it showed an unacceptable level of under-triage and was considered not valid, requiring further research. However, reliability alone is insufficient and considered alone without acceptable validity can lead to misinterpretation. Hence, given the poor validity we did not report on reliability of Sierra Leone in the paper above. There is very little published research on reliability of paediatric triage in low-resource settings and as such we see value in reporting the findings of our assessment in Sierra Leone and the lessons learnt in the section that follows. We caution that these results should not be used in supporting the use of SATS and is presented here only to share the lessons learnt from assessing reliability in a low-resource paediatric setting.

The study was conducted in the region of Gondoma, Sierra Leone. We summarise the characteristics of the study site below (Table 1). Reliability was assessed as described in the paper above. In 2013, under a classroom setting, nurses involved in triage were asked to rate vignettes based on local EC presentations. Inter-rater reliability was calculated and reported in line with the GRRAS.[57] Intra-rater reliability could not be assessed due to a Lassa fever outbreak and the nursing staff being unavailable for further participation in the study. We invited a total of 18 nurses to participate in the study. Of these, 16 agreed and completed the assessment, eliciting a response rate of 89%.

	Gondoma Referral Centre
Country	Sierra Leone
Specific location	Southern Province, 250km from the capital city
Catchment population	300,000
Level of MSF support	MSF paediatric only hospital
Services offered	EC, ICU, IPD, Maternity, Malnutrition, Lassa fever care
Type of EC cases	Paediatric only
Size of EC	6 beds
EC caseload	600
Introduction of the SATS	2012

Table 1: Characteristics of study site in Sierra Leone

EC: Emergency Centre; ICU: Intensive Care Unit; IPD: In-Patient Department; MSF: Médecins Sans Frontières; SATS: South African Triage Scale

Results: In Sierra Leone, in a paediatric only EC, the 2008 paediatric version of the SATS demonstrated substantial reliability. Point estimates for ICC, QWK, LWK and UWK, including level of agreement are presented below (Table 2).

Inter-rater reliability measures	Gondoma Referral Centre
Point estimates (95% CI)	
Intra-class correlation coefficient	0.67 (0.54-0.79)
Quadratically weighted	0.66 (0.39-0.84)
Linearly weighted	0.54 (0.30-0.73)
Un-weighted	0.43 (0.29-0.53)
Level of agreement*	
Intra-class correlation coefficient	Substantial
Quadratically weighted	Substantial

Table 2: Inter-rater reliability measures for the South African Triage Scale when used in Gondoma referral centre, Sierra Leone, 2013

CI: Confidence interval; *According to the Landis and Koch criteria[57]

In Chapter 4, we showed the developed set of paediatric vignettes using the Delphi methodology. We created this set of vignettes to try to provide a standardised and easier way to assess the reliability of triage scales in low-resource settings. This set of vignettes was developed using paediatric cases in South Africa. In Sierra Leone, our initial attempts to assess the reliability of the SATS used this set of vignettes. However, while assessing the validity of the SATS in the same setting, we realised that the types of cases presenting at Gondoma Referral Centre (GRC) in Sierra Leone were markedly different from those in the reference set of vignettes. Malaria is endemic in Sierra Leone but is not common within the SA context; as such, none of the paediatric reference vignettes dealt with a malarial case. For the vignettes to assess the local use of the scale, they need to represent at least some of the common presentations. To solve this problem, we created paper-based vignettes based on local cases from the EC at Gondoma. Lack of contextualisation in the paediatric reference vignettes was a limitation highlighted in Chapter 4, and prudence by other researchers should be noted in

relation to the use of standardised, non-contextualised sets of vignettes that are not disease specific.

In Chapter 5 we used a set of 42 sample vignettes developed by Twomey et al.[69] Given the time and cost constraints in Afghanistan, Sierra Leone and Haiti we had to reduce this number of vignettes to between 28-30

Chapter Conclusion

We set out to assess the reliability of the SATS in low-resource settings. Across our study sites, the SATS demonstrated moderate to substantial inter- and intra-rater reliability. A secondary finding relating to our methodology, was that in general, paper-based vignettes that are locally applicable seem to be a more appropriate methodology in low-resource settings. This is the largest study to date examining the reliability of the SATS in low-resource settings, and our findings corroborate the findings of studies conducted on SATS in Botswana, Ghana and South Africa.[54,59,90,91] The results of the research support the idea that the SATS is a reliable tool for hospitals in low-resource settings.

Chapter 7: Validity of the SATS in low-resource settings

Reference

Dalwai MK, Valles P, Twomey M, Nzomukunda Y, Jonjo P, Sasikumar M, Nasim M, Razaaq A, Gayraud O, Jecrois PR, Wallis L, Tayler-Smith K. Is the South African Scale valid for use in Afghanistan, Haiti and Sierra Leone? BMJ Global Health 2017;2:e000160.doi:10.1136/bmjgh-2016-000160

Declaration from Authors

The following co-authors contributed to the paper:
Dr. Pola Valles, Dr. Michele Twomey, Dr. Yvonne Nzomukunda, Mr. Prince Jonjo, Dr Manoj Sasikumar, Dr. Masood Nasim, Mr Abdul Razaaq, Dr. Olivia Gayraud, Prof. Lee Wallis and Mrs. Katie Tayler-Smith.

In the case of Chapter 7, the nature and extent of contribution by the authors was as follows:

Nature of contribution

MKD, MT, LW and KTS contributed to study conception and design; MKD, PV, YN, PJ, MS, MN, AR and OG facilitated the acquisition of data; MKD, LW, KTS, PV, OG, PJ and MS helped with the analysis and interpretation of the data; the initial version was drafted by MKD, KTS, LW and PV with all authors contributing to the redrafts and revising it critically for important intellectual content; all authors gave final approval on the version submitted and the revisions thereof.

Extent of contribution

MKD: 60%; MT, PV together: 10%;YN, PJ, MS, OG, MN, AR together: 20%; LW and KTS together 10%

Signed: Mohammed. K. Dalwai 19 Oct 2017

Declaration from Co-Authors

The undersigned hereby certify that:

1. The above declaration correctly reflects the nature and extent of the candidate's contribution to this work, and the nature of the contribution of each of the co-authors.
2. They meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
3. They take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
4. There are no other author of the publication according these criteria;
5. Potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and
6. The original data are stored at the following location and will be held for at least five years from date indicated below.

Prof. Lee Wallis 25 November 2017

Dr Michele Twomey 25 November 2017

Mrs Katie Tayler-Smith 02 November 2017

Introduction to the topic

Validity testing is essential in the development and implementation of triage scales as it allows us to ascertain whether the scale in question accurately prioritises patients according to their true urgency for care (i.e., acuity level). The main challenge around such testing is the absence of a gold standard, i.e., a robust method that accurately measures what triage is intended to do. Different authors have measured validity in a variety of ways.[22] In recent studies, validity has been measured in two different ways: surrogate markers and expert opinion.[20,89] Expert opinion correlates triage acuity assigned by staff to an expert opinion. Surrogate markers correlate triage acuity with an expected outcome such as admission, death, length of stay or resource utilisation. The limitation of this approach is that the use of surrogate markers relies on proper record keeping and a baseline standard of care; in low-resource settings both are often lacking. Using MSF resources, we have ensured standardised record keeping and clinical care across multiple sites, thus enabling us to use surrogate markers such as hospital admission, discharge, referral and death as a proxy for true patient acuity.

Motivation for conducting the study

In Chapter 6 we reported on the reliability of the SATS in various hospital settings across different LMICs. In addition to reliability, establishing the validity of the SATS is just as important. The SATS has undergone extensive evaluations in South Africa, but its performance across a spectrum of different LMIC settings, particularly non-sub-Saharan African and trauma-only settings, has not been adequately tested. Such information could inform us on the extent to which the SATS has a useful role in such settings.

Aim

The study aim was to investigate the validity of the SATS across multiple low-resource settings.

Objectives

- To determine the validity of the SATS using EC outcomes (e.g., discharge, hospital admission, death) as a proxy for true patient acuity.

Main Findings

- The SATS is a valid tool for use across trauma settings in low-resource settings.
- The SATS demonstrates mixed validity results for medical and mixed medical settings and as such context specific assessments would seem warranted.
- The SATS under-triaged a high proportion of paediatric cases, many of whom had severe anaemia which went unnoticed at the time of triage.

Research

BMJ Global Health

Is the South African Triage Scale valid for use in Afghanistan, Haiti and Sierra Leone?

Mohammed Dalwai,[1,2] Pola Valles,[3] Michele Twomey,[1] Yvonne Nzomukunda,[4] Prince Jonjo,[4] Manoj Sasikumar,[4] Masood Nasim,[5] Abdul Razaaq,[5] Olivia Gayraud,[6] Pierre Ronald Jecrois,[6] Lee Wallis,[1] Katie Tayler-Smith[2]

To cite: Dalwai M, Valles P, Twomey M, *et al*. Is the South African Triage Scale valid for use in Afghanistan, Haiti and Sierra Leone? *BMJ Glob Health* 2017;**2**:e000160. doi:10.1136/bmjgh-2016-000160

Received 23 August 2016
Revised 13 February 2017
Accepted 22 February 2017

[1]Division of Emergency Medicine, University of Cape Town, Cape Town, South Africa
[2]Operational Research Unit Luxembourg, Médecins Sans Frontières, Luxembourg
[3]Medical department, Médecins Sans Frontières, Operational Centre, Brussels, Belgium
[4]Medical department, Médecins Sans Frontières, Free Town, Sierra Leone
[5]Medical department, Médecins Sans Frontières, Kabul, Afghanistan
[6]Medical department, Médecins Sans Frontières, Port au Prince, Haiti

Correspondence to
Dr Mohammed Dalwai;
MKDALWAI@gmail.com

ABSTRACT

Objective To assess the validity of the South African Triage Scale (SATS) in four Médecins Sans Frontières (MSF)-supported emergency departments (ED, two trauma-only sites, one mixed site (both medical and trauma cases) and one paediatric-only site) in Afghanistan, Haiti and Sierra Leone.

Methods This was a retrospective cohort study conducted between June 2013 and June 2014. Validity was assessed by comparing patients' SATS ratings with their final ED outcome (ie, hospital admission, death or discharge).

Results In the two trauma settings, the SATS demonstrated good validity: it accurately predicted an increase in the likelihood of mortality and hospitalisation across incremental acuity levels ($p<0.001$) and ED outcomes for 'green' and 'red' patients matched the predicted ED outcomes in 84%–99% of cases. In the mixed ED, the SATS was able to predict an incremental increase in hospitalisation ($p<0.001$) across both trauma and non-trauma cases. In the paediatric-only settings, SATS was able to predict an incremental increase in hospitalisation in the non-trauma cases only ($p<0.001$). However, 87% (non-trauma) and 94% (trauma) of 'red' patients in the mixed-medical setting were overtriaged and 76% (non-trauma) and 100% (trauma) of 'green' patients in the paediatric settings were undertriaged.

Conclusion The SATS is a valid tool for trauma-only settings in low-resource countries. Its use in mixed settings seems justified, but context-specific assessments would seem prudent. Finally, in paediatric settings with endemic malaria, adding haemoglobin level to the SATS discriminator list may help to improve the undertriage of patients with malaria.

Key questions

What is already known about this subject?

- Emergency medical care (EMC) is recognised as being one of the core components of a horizontal approach to improving population health in low to middle-income countries (LMICs). Unfortunately, EMC is seldom seen as a priority in LMICs. One of the major challenges is the extreme mismatch between the demand for EMC and resource capacity. Triage may offer a simple and cost-effective mechanism for managing this problem.
- There are few triage scales designed specifically for use in LMICs; the South African Triage Scale (SATS) is one of them. The SATS has been assessed extensively in South Africa, but its performance across a spectrum of different LMIC settings, mainly non-sub-Saharan African and trauma-only settings, has not been adequately evaluated.

What are the new findings?

- This is the first study to look at the use of the SATS and its validity in Afghanistan, Sierra Leone and Haiti.

Recommendations for policy?

- The findings of our study have several practice and policy implications. Given that the SATS was found to be highly valid in two different trauma settings in Haiti and Afghanistan, the implication is that this tool would be of value in other low-resource trauma settings too.

INTRODUCTION

Emergency medical care (EMC) is recognised as being one of the core components of a horizontal approach to improving population health in low to middle-income countries (LMICs).[1] A substantial number of the conditions that contribute to the burden of disease in these countries could be alleviated through the provision of effective EMC. Unfortunately, however, EMC in LMIC settings is seldom considered a priority: EMC systems remain underdeveloped and under-resourced. One of the major challenges is around the extreme mismatch between the demand for EMC and resource capacity. Triage may offer a simple and cost-effective mechanism for managing this challenge.

Triage aims to determine a patient's urgency for care (defined as their acuity level) in order to separate critically ill patients, who need immediate life-saving interventions, from patients who need medical attention but can safely wait to be seen.[2] It is one of the core requirements for the provision of

effective EMC and has been shown to reduce morbidity and mortality.[3]

There are few triage scales designed specifically for use in LMICs; the South African Triage Scale (SATS) is one of these: a four-level triage scale that colour codes patients as follows: (1) red—emergency; (2) orange—very urgent; (3) yellow—urgent; or (4) green—routine.[4] The SATS has been assessed extensively in South Africa, but its performance across a spectrum of different LMIC settings, particularly non-sub-Saharan African and trauma-only settings, has not been adequately assessed.

To be of value, a triage tool must demonstrate good validity, that is, the acuity ratings assigned using the triage scale must closely reflect patients' true acuities.[5] Assessing the validity of a triage tool is inherently challenging due to the absence of a gold standard. Previous studies have either used surrogate markers such as hospital admission, mortality or resource utilisation as a proxy for true patient acuity[6–10] or they have validated triage ratings assigned by emergency department (ED) staff for a series of simulated cases against ratings obtained from an expert panel.[5 11] Both methods have their strengths and limitations. In particular, the use of some surrogate markers relies on reliable and accurate record keeping and standardisation of clinical care, which is often lacking in LMIC settings.

Médecins sans Frontières (MSF), an international medical humanitarian organisation, provides free medical care, including emergency care, to vulnerable populations in many LMICs. Since 2011, MSF-Operational Centre Brussels has been using the SATS in projects where it provides emergency care. Aside a project in Pakistan,[11] the validity of the SATS has not been assessed in any of these settings.

Unlike many public EDs in LMICs, MSF-supported EDs are more suited for testing and comparing the validity of a triage tool using ED outcomes as a proxy for true acuity. This is because a standardised package of ED care is implemented; on-site support, training and supervision are ensured; and robust and standardised data collection systems are in place. As such, we set out to assess the validity of the SATS in four MSF emergency care settings (two trauma sites, one mixed (both trauma and medical cases) site and one paediatric-only site) in Afghanistan, Haiti and Sierra Leone, using ED outcomes (discharge, hospitalisation and death) as our reference standard.

METHODS

Study design

This was a retrospective cohort study. Validity was assessed by comparing patients' recorded SATS acuity ratings with their final ED outcome (hospital admission, death or discharge). The rationale for using ED outcomes as a reference standard for this validation was based on the logic that (1) moving incrementally from low to high acuity levels ('green' to 'red'), the SATS would demonstrate an increasing trend in the likelihood of mortality and hospitalisation; (2) SATS triaged 'routine' (green) patients should not die or require hospitalisation; and (3) SATS triaged 'emergency' (red) patients would require hospitalisation or may die.

Ethics

Approval was obtained from the national ethics bodies in Afghanistan, Haiti and Sierra Leone, from the MSF Ethics Review Board and from the University of Cape Town.

Study setting and population

The study involved four MSF sites: two trauma centres in Kunduz, Afghanistan and Port-au-Prince, Haiti; a mixed ED in Martissant, Haiti; and a paediatric ED in Bo, Sierra Leone.

Kunduz Trauma Centre, Afghanistan

Kunduz province is located in North-Eastern Afghanistan, with a population of approximately one million. It has experienced a large amount of active conflict and sees high rates of trauma, mainly due to violence and road traffic accidents. The MSF Kunduz Trauma Centre was opened in August 2011 to fill a gap in the provision of trauma care in the area. At the time of the study, the centre had 92 inpatient beds and offered emergency, orthopaedic, surgical, physiotherapy and intensive care. The ED had a caseload of over 1800 patients with trauma per month. The SATS has been in use in the ED since the centre opened. In October 2015, the hospital was bombed and destroyed by US-led coalition forces killing patients and 14 MSF staff members, in the largest loss of life for MSF ever.

Tabarre surgical and trauma centre, Haiti

The MSF Tabarre trauma centre is located in an industrial area of Port au Prince—Haiti's capital—and serves a catchment population of about one million. The city has a high burden of trauma emergencies related, in particular, to high numbers of road accidents and growing urban violence. The centre is a modular hospital that started providing specialised care for trauma and acute surgical conditions in February 2012. At the time of the study, the hospital had 107 beds and offered emergency, surgical, orthopaedic and intensive care, including inpatient and physiotherapy services. It saw an average of 793 patients with trauma per month. The SATS was implemented from the time the centre opened.

Martissant clinic, Haiti

Situated in the busy urban area of Martissant in Port au Prince, this MSF clinic serves a catchment population of 1.2 million people. The facility has an ED with a short-stay ward and when needed has the ability to convert a section into a cholera treatment centre. It deals with a mixed medical caseload and treats approximately 4900 patients per month. The SATS was introduced in 2012.

Gondama Referral Centre, Bo, Sierra Leone

At the time of the study, Gondama Referral Centre (GRC) was a mother and child MSF hospital, admitting children aged 15 years and below, situated close to the town of Bo in the Southern Province of Sierra Leone, approximately 250 km from the capital city of Freetown. The hospital included a paediatric-only ED, an inpatient department with wards dedicated to malnutrition and Lassa fever and a separate maternity unit. At the time of the study, the hospital saw approximately 500 patients per month and served a catchment population of 300 000. The paediatric version of the SATS was implemented at GRC's ED during 2012.

SATS and its use in the EDs

The SATS is a four-tiered triage tool which has been extensively described elsewhere.[12] The colour categories show the urgency into priority 1: red—'emergency' (to be seen immediately); priority 2: orange—'very urgent' (to be seen within 10 min); priority 3: yellow—'urgent' (to be seen within 60 min) and priority 4: green—'routine' (to be seen within 240 min). The SATS also allocates the colour blue to 'dead on arrival cases' (MSF has used the colour black to denote dead on arrival cases due to cultural norms in certain countries).

Patient triage was carried out by registered nurses at each site, and training in triage and use of the SATS was covered in a 1-day training session.

Study protocol

The study included all patients who presented at the study sites between June 2013 and June 2014. Patients were excluded if they were declared dead on arrival (and therefore not triaged). Patients were also excluded if they were referred for care, or absconded from care without medical permission, because of not being able to use these two outcomes as logical proxies for acuity level.

Data collection

Data pertaining to the study were sourced from the electronic ED registers at each site and extracted into an Excel database. Variables included age, sex, triage acuity (green, yellow, orange or red), ED outcome (discharged, hospitalised or died) and type of case (trauma, non-trauma).

Analysis

Reference standards (ie, predicted ED outcomes) according to acuity were as follows. (1) Moving incrementally from a low to high acuity ('green' to 'red'): an increasing trend in the likelihood of mortality and hospitalisation. (2) 'Routine' (green) patients: discharge; no hospitalisations or deaths. (3) 'Emergency' (red) patients: hospitalisation or death; no discharges

Trends in mortality and hospitalisation over incremental acuity levels were assessed using the χ^2 test for trend. Additionally, a univariate analysis was performed to determine the relative risk of hospitalisation by acuity level. A log-binomial model was used as the first approach for this analysis, but as the model failed to reach convergence, a Poisson regression with robust variance estimator model was used instead[13] (the latter was not performed for mortality due to there being zero deaths for some acuity levels). For the last two reference standards, frequencies and proportions were calculated to assess the distribution of ED outcomes for 'green' and 'red' patients. Undertriage was calculated as the proportion of 'green' patients requiring hospitalisation or dying, and overtriage, the proportion of 'red' patients being discharged. Accepted thresholds for undertriage and overtriage were set at 10% and 50%, respectively, according to the American College of Surgeons Committee on Trauma.[14] In Martissant and GRC, the reference standards were considered in relation to trauma and non-trauma ED presentations. The level of significance was set at p=0.05 throughout. Data were analysed using the STATA/IC V. 12.0 software (Stata, Texas, USA).

RESULTS

During the study period, the total numbers of ED patients presenting at Kunduz, Tabarre, Martissant and GRC were 19474, 7706, 56 919 and 8190, respectively. Data on sex, outcome, acuity or type of case were missing for 3 patients at Kunduz and 15 patients at Martissant, and as such these patients were excluded from the analysis. Further patients were excluded if they were dead on arrival or if they were either referred for or absconded from care: 1845 (9%) at Kunduz, 286 (4%) at Tabarre, 2173 (4%) at Martissant and 301 (4%) at GRC.

Patient characteristics and their ED outcomes by acuity level

Table 1 shows the characteristics of the patients included in the analysis and table 2 shows their ED outcomes by acuity level.

Performance of the SATS according to the reference standards

In the trauma-only settings of Kunduz and Tabarre, the SATS performed well across all acuity levels: it accurately predicted an increase in the likelihood of mortality and hospitalisation moving from low to high acuity levels (p<0.001) (tables 2 and 3 and figure 1), and ED outcomes for 'green' and 'red' patients matched the predicted ED outcomes in 84%–99% of cases (table 4). Undertriage for 'green' patients and overtriage for 'red' patients did not exceed the acceptable thresholds of 10% and 50%, respectively.

In the mixed setting of Martissant, the SATS accurately predicted an increasing trend in mortality and hospitalisation moving from low to high acuity levels for both trauma and non-trauma cases, p<0.001 (tables 2 and 3 and figure 1. For 'green' patients, ED outcomes matched the predicted outcome in over 99% of cases (indicating virtually no undertriage), whereas predicted ED outcomes for 'red' patients were only matched in 12% of non-trauma cases and 6% of trauma cases

(table 4); most 'red' cases were discharged (table 3) indicating a situation of overtriage that exceeded the acceptable threshold of 50%. The most common ED presentations among 'red' patients who were discharged (n=1024) were fever without identified cause (22%), accidental trauma (13%), lower respiratory tract infection (11%), upper respiratory tract infection (8%) and asthma (6%).

In the paediatric setting of GRC, the SATS was able to accurately predict a significant increase in the likelihood of mortality moving from low to high acuity levels for patients with trauma and non-trauma patients, p<0.001 (table 2). The trend in hospitalisation moving from low to high acuity levels however was marginal (for patients with trauma, there was no significant trend and for non-trauma patients although statistically significant we would question the clinical significance) and largely distorted by the high proportion of 'green' patients' (the reference group) being hospitalised. The SATS performed well for 'red' patients with predicted ED outcomes being matched in 100% and 98% of patients with trauma and non-trauma patients, respectively (overtriage was therefore well below the acceptable threshold of 50%), but no 'green' patients with trauma and only 24% of 'green' non-trauma patients were discharged in accord with what was predicted (table 4); the majority of 'green' patients were hospitalised indicating a situation of undertriage that exceeded the acceptable threshold of 10% (table 2). Among 'green' patients who were hospitalised (n=547), more than half had confirmed severe malaria (33%) or confirmed uncomplicated malaria (25%).

Across the four sites, gender did not affect the predicted ED outcome (data not shown).

DISCUSSION

This is one of few studies assessing and comparing the validity of the SATS in different low-resource settings—especially non sub-Saharan African settings—and the first to assess validity in trauma settings. The SATS demonstrated high validity in the trauma settings of Kunduz and Tabarre but in the mixed setting of Martissant and the paediatric setting of Gondama, it demonstrated a tendency to overtriage at the 'red' end of the scale in Martissant and to undertriage at the 'green' end of the scale in Gondama.

The main strengths of the study were that (1) it was multisite; (2) it included large numbers of patients; (3) it involved the analysis of real-life ED cases representative of the setting, rather than generic paper-based cases; and (4) the study period covered one whole year, thus taking account of seasonal variations in types of ED cases and disease burden.

The main study limitations were around the use of ED outcomes as a proxy for true acuity. First, while ED outcomes are accepted as being a suitable proxy for true acuity,[6 8 10] given the absence of a gold standard, various factors may have confounded this relationship including inaccurate measurement and recording of patient acuity,

Table 1 Sociodemographic and clinical characteristics of patients included in the study at Kunduz, Tabarre, Martissant and Gondama EDs, June 2013–June 2014

Variable	Kunduz n (%)	Tabarre n (%)	Martissant n (%)	GRC n (%)
Total	17626	7420	54731	7889
Sex				
Female	3858 (22)	2267 (31)	24206 (44)	3593 (46)
Male	13768 (78)	5153 (69)	30525 (56)	4296 (54)
Median age, years (range)	34 (<1–120)	45 (<1–101)	44 (<1–101)*	3 (<1–15)
Case type				
Trauma	17626 (100)	7420 (100)	31924 (58)	94 (1)
Non-trauma	0	0	22807 (42)	7795 (99)
SATS classification				
Green	5373 (30)	1225 (17)	18649 (34)	719 (9)
Yellow	8291 (47)	4387 (59)	27637 (51)	1809 (23)
Orange	3455 (20)	1625 (22)	7290 (13)	2782 (35)
Red	507 (3)	183 (2)	1155 (2)	2579 (33)
ED outcome				
Discharged	15016 (85)	4499 (61)	54260 (99)	766 (10)
Admitted	2598 (15)	2917 (39)	449 (1)	7015 (89)
Died	12 (<1)	4 (<1)	22 (<1)	108 (1)

*Age unknown for one record.
ED, emergency department; GRC, Gondama Referral Centre; SATS, South African Triage Scale.

Table 2 ED outcomes by acuity level for patients at Kunduz, Tabarre, Martissant and Gondama EDs, June 2013–June 2014

		ED outcome		
	n	Discharged n (%)	Admitted n (%)	Died n (%)
Kunduz				
Green	5373	5344 (99)	29 (1)	0
Yellow	8291	7459 (90)	832 (10)	0
Orange	3455	2134 (62)	1321 (38)	0
Red	507	79 (16)	416 (82)	12 (2)
Tabarre				
Green	1225	1110 (91)	115 (9)	0
Yellow	4387	2911 (66)	1476 (34)	0
Orange	1625	469 (29)	1156 (71)	0
Red	183	9 (5)	170 (93)	4 (2)
Martissant				
Trauma cases	31 924			
Green	14 276	14 262 (100)	14 (0)	0
Yellow	15 728	15 695 (100)	33 (0)	0
Orange	1716	1699 (99)	17(1)	0
Red	204	192 (94)	4 (2)	8 (4)
Non-trauma cases	22 807			
Green	4373	4336 (99)	37 (1)	0
Yellow	11 909	11 807 (99)	102 (1)	0
Orange	5574	5437 (98)	134 (2)	3 (0)
Red	951	832 (87)	108 (11)	11 (1)
GRC				
Trauma cases	94			
Green	6	0 (0)	6 (100)	0
Yellow	18	2 (11)	16 (89)	0
Orange	42	2 (5)	40 (95)	0
Red	28	0 (0)	27 (96)	1 (4)
Non-trauma cases	7795			
Green	713	171 (24)	541 (76)	1 (0.1)
Yellow	1792	316 (18)	1469 (82)	7 (0.4)
Orange	2739	228 (8)	2498 (91)	13 (0.5)
Red	2551	47 (2)	2418 (95)	86 (3)

ED, emergency department; GRC, Gondama Referral Centre.

inaccurate outcome reporting and suboptimal ED care resulting in a more severe ED outcome than would otherwise have been predicted. As it was, we believe that these factors were minimised as far as possible through the provision of standardised ED care, adequate resources (human and material), on-site support, supervision and regular training of ED staff, and robust data collection systems. Second, the reference standards (ie, the predicted ED outcomes) assigned for 'green' and 'red' patients may have failed to take account of anomalies in discharge and hospitalisation criteria related to certain patient conditions, leading to an overestimation of undertriage and overtriage. Third, because we were unable to predict specific ED outcomes for 'yellow' and 'orange' patients, undertriage and overtriage could not be estimated for these middle acuity levels. Finally, in the absence of any reference standards indicating how steep an increasing trend in mortality or hospitalisation across incremental acuity levels should be, we cannot use this to try to quantify the relative validity of the SATS in any way. A positive trend, however, provides evidence that the SATS is able to discriminate well between successive acuity levels and supports its validity.

To date, very few studies have assessed the validity of the SATS in low-resource settings; among those that have, only one has assessed it in a non-sub-Saharan African

Table 3 Hospital admissions by acuity level for patients attending Kunduz, Tabarre, Martissant and Gondama emergency departments, June 2013–June 2014

Acuity level	n	Admitted n (%)	RR (95% CI)	p Value*
Kunduz				
Green	5373	29 (1)	1	<0.001
Yellow	8291	832 (10)	19 (13 to 27)	
Orange	3455	1321 (38)	71 (49 to 102)	
Red	507	416 (82)	152 (106 to 219)	
Tabarre				
Green	1225	115 (9)	1	<0.001
Yellow	4387	1476 (34)	4 (3 to 4)	
Orange	1625	1156 (71)	8 (6 to 9)	
Red	183	170 (93)	10 (8 to 12)	
Martissant				
Trauma cases	31 924			
Green	14 276	14 (0)	1	<0.001
Yellow	15 728	33 (0)	2 (1 to 4)	
Orange	1716	17(1)	10 (5 to 21)	
Red	204	4 (2)	20 (7 to 60)	
Non-trauma cases	22 807			
Green	4373	37 (1)	1	<0.001
Yellow	11 909	102 (1)	1 (0.7 to 1.5)	
Orange	5574	134 (2)	3 (2 to 4)	
Red	951	108 (11)	13 (9 to 19)	
GRC				
Trauma cases	94			0.64
Green	6	6 (100)	1	
Yellow	18	16 (89)	0.9 (0.8 to 1.0)	
Orange	42	40 (95)	1.0 (0.9 to 1.0)	
Red	28	27 (96)	1.0 (0.9 to 1.0)	
Non-trauma cases	7795			
Green	713	541 (76)	1	<0.001
Yellow	1792	1469 (82)	1.1 (1.0 to 1.1)	
Orange	2739	2498 (91)	1.2 (1.2 to 1.3)	
Red	2551	2418 (95)	1.2 (1.2 to 1.3)	

* χ^2 test for trend.
GRC, Gondama Referral Centre; RR, risk ratio.

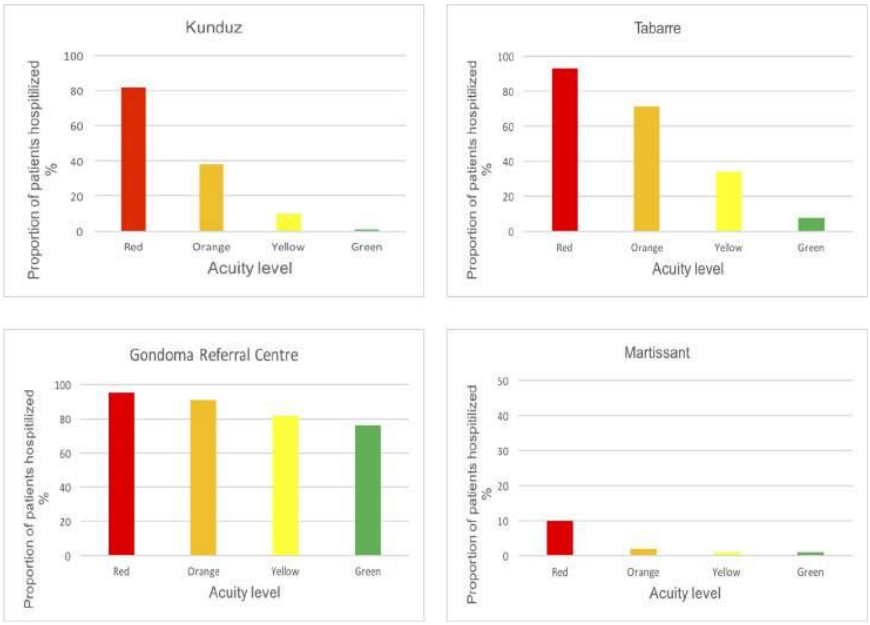


Figure 1 Proportion of hospital admissions by acuity level for patients attending Kunduz, Tabarre, Martissant and Gondama emergency departments, June 2013–June 2014.

setting[11] and none in any trauma-only settings. Furthermore, the study design used in most of these studies has generally involved comparing the triage ratings assigned by ED staff for a series of simulated paper-based cases with those obtained from an expert panel.[6 15] This methodology has several limitations: (1) non-verbal patient cues and contextual information can introduce a potential interpretation bias, (2) the simulated cases used do not always represent the patient populations and disease burdens seen in the study setting and (3) it is difficult to decipher whether any identified discrepancies between ED staff ratings and the expert panel ratings are in fact a reflection of poor validity of the SATS or because ED staff are not applying the triage tool accurately.[11]

There is only one other study, conducted in South Africa, that has assessed the validity of the SATS using surrogate markers as a proxy for true acuity.[6] In this study, hospital admission and discharge were used as the outcome markers for validating a revised version of the paediatric SATS. Validity indicators were determined based on the assumption that all discharges would comprise 'green' patients and that all hospital admissions would comprise 'yellow', 'orange' or 'red' patients. In our study, for 'green' and 'red' patients, we essentially applied this same assumption (ie, that among 'green' patients there would only be discharges and that among 'red' patients, there would be no discharges). However, for the acuity levels ofyellow and orange, we refrained from hypothesising that there would be no discharges; this, on account of the type of ED presentations that could potentially fall under these acuities. Instead, we assessed these middle acuities by applying the logic that moving incrementally from a low to high acuity, an increasing trend in the likelihood of mortality and hospitalisation would be observed. This same logic has been applied and demonstrated in other triage validity studies.[7 9 16 17] We feel that this trend analysis, in combination with the specific reference standards assigned for 'green' and 'red' patients at

Table 4 Proportion of 'green' and 'red' patients whose ED outcome matched the reference standard (ie, the predicted ED outcome) at Kunduz, Tabarre, Martissant and Gondama EDs, June 2013–June 2014

		Patients with expected ED outcome (% OT or UT)					
				Martissant		GRC	
SATS acuity level	Expected ED outcome	Kunduz	Tabarre	Trauma	Non-trauma	Trauma	Non-trauma
Green	Discharge (% UT)	99 (1)	91 (9)	100 (0)	99 (1)	0 (100)	24 (76)
Red	Hospitalisation or death (% OT)	84 (16)	95 (5)	6 (94)	12 (88)	100 (0)	98 (2)

ED, emergency department; GRC, Gondama Referral Centre; OT, overtriage; SATS, South African Triage Scale; UT, undertriage.

either end of the scale, provides a more rational way of evaluating the validity of a triage scale when using ED outcomes.

In the trauma-only settings of Kunduz and Tabarre, the SATS demonstrated excellent construct validity. In the mixed medical setting of Martissant, however, while the 'green' to 'orange' part of the SATS seemed to perform well, the 'red' end of the scale appeared to overtriage patients, that is, to overestimate their urgency for care. Given that this was not the case in the trauma settings of Kunduz and Tabarre, we had wondered whether this might be a phenomenon specific to non-trauma medical patients. This hypothesis seems unfounded however as high rates of overtriage among 'red' patients in Martissant were observed for both patients with trauma and non-trauma patients. An alternative explanation for this overtriage might instead relate to the set up at Martissant. Martissant is a clinic with only a small six-bed short-stay ward on site and therefore there is a strong drive to discharge patients as soon as possible. Anecdotal evidence suggests that a significant proportion of the discharged 'red' patients presented with conditions (accidental trauma, fever, respiratory conditions) which were justifiably 'emergency' in nature at the time of triage but which staff at the clinic were able to quickly stabilise so that patients could be safely discharged. It is also possible that in some cases there was unpredicted recovery of serious looking initial symptoms. This requires further investigation, but suggests that the construct of the SATS itself is not to blame for what appears to be a situation of overtriage. Further support for the validity of the SATS in this setting comes from the clear increasing trend in mortality and hospitalisation observed moving incrementally across acuity levels.

In Gondama, the 'green' end of the scale appeared to have a strong tendency to undertriage patients, that is, to underestimate their urgency for care. The ramifications of undertriage are more dangerous than those of overtriage, whereby sick patients do not receive the necessary medical attention as urgently as needed. At GRC, this tendency to undertriage patients was more common among patients with trauma compared with non-trauma patients. Given that patients with trauma only made up 1% of GRC's total presentations, one possible explanation for this difference may be related to clinicians erring on the side of caution more for trauma cases (on account of having less experience of managing them), leading to patients being admitted even if this is not always necessary. This hypothesis is indirectly supported by the fact that there were no deaths among any of these 'green' patients with trauma. Had deaths been observed among this group of patients, this would more strongly point towards the fact that the seriousness of the patient's condition had gone unrecognised, that is, they had indeed been undertriaged. At any rate, further investigation into the management of patients with trauma at GRC would be warranted to substantiate our speculation around this issue.

The bulk of the patients at GRC were non-trauma patients and more than half of these patients who were triaged as 'green' and then hospitalised had uncomplicated or severe malaria diagnosed after triage. Most of these patients had concurrent severe anaemia of <7 g/dL, and this was the primary reason for them being hospitalised. As the paediatric version of the SATS used in GRC does not include a measure for haemoglobin, severe anaemia no doubt went 'unnoticed' at the time of triage, explaining why these patients were undertriaged. To mitigate this, we would recommend adding haemoglobin cut-off values to the SATS discriminator list and reassessing the performance of the SATS. If this modification is able to negate the undertriage of all 'green' hospitalised patients with malaria, undertriage among green patients would theoretically reduce from 88% to 32% which is a significant improvement. However, it still leaves the undertriage rate above the acceptable threshold of 10% and this needs careful investigation.

The findings of our study have a number of practice and policy implications. First, given that the SATS was found to be highly valid in two different trauma settings in Haiti and Afghanistan, the implication is that this tool could be of value in other low-resource trauma-only settings too. Trauma is a leading cause of global mortality (injuries account for approximately 5.8 million deaths annually) with 90% of these deaths occurring in LMICs.[18] Minimising the morbidity and mortality related to injuries relies on timely and effective emergency care and the SATS could provide a simple and cost-effective mechanism by which to facilitate this.

Second, although there was some questionability about the performance of the SATS in the mixed setting of Martissant, we believe that this is largely related to the set up at the Martissant clinic itself, rather than the construct of the SATS. As such, implementation of the SATS in other low-resource mixed ED settings would seem justified, although context-specific assessments in such settings would still be prudent.

Third, in paediatric settings with endemic malaria, adding anaemia to the discriminator list could help ensure that patients with malaria are given greater priority.

Finally, when assessing the validity of a triage scale using ED outcomes as a proxy for true acuity, we would recommend using a trend analysis of mortality and hospitalisation across incremental acuity levels, in combination with specific reference standards for the extreme acuity levels of the scale.

CONCLUSION

The SATS is a valid triage tool for prioritisation of patients with trauma in low-resource settings. Its use in mixed ED settings seems justified, but context-specific assessments of its performance would nonetheless be prudent. In paediatric settings with endemic malaria, adding

haemoglobin levels to the SATS discriminator may help to improve the undertriage of patients with malaria.

Acknowledgements We would like to extend our condolences to the families of the staff and patients who lost their lives in Kunduz Hospital (one of our study sites). The hospital was bombed and destroyed by US coalition forces in October 2015. We would like to acknowledge and thank all the staff both nationally and internationally who helped with this study.

Contributors MD, PV, MT, LW and KT-S designed, analysed and interpreted the study and data. YN, PJ and MS were the project leads in Sierra Leone. MN and AR were the project leads in Afghanistan. OG and PRJ were the project leads in Haiti. All authors contributed to the revision of the final article.

Funding Mohammed Dalwai is supported by a grant from the Medical Research Council of South Africa, under the MRC Clinician Researcher Program

Competing interests No, there are no competing interests.

Patient consent Routine anonymised patient data was obtained from MSF sites.

Ethics approval Approval was obtained from the national ethics bodies in Afghanistan, Haiti and Sierra Leone, from the MSF Ethics Review Board and from the University of Cape Town.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Data sharing on request

REFERENCES

1. Razzak JA, Kellermann AL. Emergency medical care in developing countries: is it worthwhile? *Bull World Health Organ* 2002;80:900–5.
2. Harrison HL, Raghunath N, Twomey M. Emergency triage, assessment and treatment at a district hospital in Malawi. *Emerg Med J* 2012;29:924–5.
3. Molyneux E, AhmadS, Robertson A. Improved triage and emergency care for children reduces inpatient mortality in a resource-constrained setting. *BullWorld Health Organ* 2006;84:314–9.
4. Gottschalk SB, Wood D, DeVries S, *et al*. The cape triage score: a new triage system South Africa. Proposal from the cape triage group. *Emerg Med J* 2006;23:149–53.
5. Twomey M, Wallis LA, Myers JE. Limitations in validating emergency department triage scales. *Emerg Med J* 2007;24:477–9.
6. Twomey M, Cheema B, Buys H, *et al*. Vital signs for children at triage: a multicentre validation of the revised South African Triage Scale (SATS) for children. *S Afr Med J* 2013;103:304–8.
7. Wuerz RC, Milne LW, Eitel DR, *et al*. Reliability and validity of a new five-level triage instrument. *Acad Emerg Med* 2000;7:236–42.
8. Taboulet P, Moreira V, Haas L, *et al*. Triage with the french emergency nurses classification in hospital scale: reliability and validity. *Eur J Emerg Med* 2009;16:61–7.
9. van der Wulp I, Schrijvers AJ, van Stel HF. Predicting admission and mortality with the emergency severity index and the manchester triage system: a retrospective observational study. *Emerg Med J* 2009;26:506–9.
10. Parenti N, Ferrara L, Bacchi Reggiani ML, *et al*. Reliability and validity of two four-level emergency triage systems. *Eur J Emerg Med* 2009;16:115–20.
11. Dalwai MK, Twomey M, Maikere J, *et al*. Reliability and accuracy of the South African Triage Scale when used by nurses in the emergency department of Timergara hospital, Pakistan. *S Afr Med J* 2014;104:372.
12. Dalwai MK, Tayler-Smith K, Trelles M, *et al*. Implementation of a triage score system in an emergency room in Timergara, Pakistan. *Public Health Action* 2013;3:43–5.
13. Spiegelman D, Hertzmark E. Easy SAS calculations for risk or prevalence ratios and differences. *Am J Epidemiol* 2005;162:199–200.
14. Sasser SM, Hunt RC, Faul M, *et al*. Centers for Disease Control and Prevention (CDC). Guidelines for field triage of injured patients: recommendations of the National Expert Panel on Field Triage, 2011. *MMWR Recomm Rep* 2012;61:1–20.
15. Olofsson P, Gellerstedt M, Carlström ED. Manchester Triage in Sweden - interrater reliability and accuracy. *Int Emerg Nurs* 2009;17:143–8.
16. Tanabe P, Gimbel R, Yarnold PR, *et al*. Reliability and validity of scores on the emergency severity index version 3. *Acad Emerg Med* 2004;11:59–65.
17. Wuerz RC, Travers D, Gilboy N, *et al*. Implementation and refinement of the emergency severity index. *Acad Emerg Med* 2001;8:170–6.
18. Sakran JV, Greer SE, Werlin E, *et al*. Care of the injured worldwide: trauma still the neglected disease of modern society. *Scand J Trauma Resusc Emerg Med* 2012;20:64–5.

Is the South African Triage Scale valid for use in Afghanistan, Haiti and Sierra Leone?

Mohammed Dalwai, Pola Valles, Michele Twomey, Yvonne Nzomukunda, Prince Jonjo, Manoj Sasikumar, Masood Nasim, Abdul Razaaq, Olivia Gayraud, Pierre Ronald Jecrois, Lee Wallis and Katie Tayler-Smith

BMJ Glob Health 2017 2:
doi: 10.1136/bmjgh-2016-000160

Supplemental Discussion

Three types of validity are commonly mentioned in the literature: criterion, content and construct.[58] In their book however, Streiner and colleagues argue that the terms construct, criterion and content validity should try to be avoided as they cause confusion and unnecessary debate. The authors postulate that all forms of validity encompass the idea of construct validity and that researchers assessing validity should therefore just make sure that they describe which construct is used in their studies.[56]

As mentioned in Chapter 2, in triage validity research, given the lack of an agreeable gold standard, two methods are commonly accepted:
1. Use of expert opinion. Paper-based vignettes are most commonly used however some studies have used real-time simultaneous side-by-side triage with an expert. With simulated paper-based cases, the vignettes are assessed by 'triage experts' and assigned a rating (i.e., the expert gold standard). Thereafter, triage officers independently triage the vignettes under classroom conditions. The officers' ratings are then compared to the expert gold standard, and measures of accuracy (i.e., over-triage and under-triage) are calculated.

2. Use of EC outcomes as surrogate markers for patient acuity: hospital admission, mortality, length of hospital stay and resource utilisation are commonly used as predictive markers for true patient acuity. The assumption is that patients with a high triage acuity are more likely to be admitted to hospital, more likely to die, more likely to have a longer hospital stay and more likely to require greater resources than patients with lower triage acuities.

None of the available methods is deemed to be superior but each method has its strengths and weaknesses. In LMICs, paper-based vignettes are more commonly used because they are cost-effective, time-efficient, and simple to use and assess.[69,89,92,93] This methodology has some inherent limitations however. First, in most studies, paper-based vignettes are used rather than real-time patient cases. As such, triage nurses are not privy to the non-verbal cues that they would ordinarily receive from real patients, and in addition they are unable to probe further with questions to the patient. In a study by Rutschmann et al., use of a digital triage

simulator of vignettes (that included patient images and allowed triage nurses to ask questions) improved triage accuracy compared to triage accuracy using paper-based vignettes alone.[94] The authors suggested that vignettes could be further enhanced by adding short videos to each. This may mitigate some of the limitations around using paper based vignettes, but does not address the fact that vignette composition and construction are reliant on the study authors and as such are unstandardised. Twomey et al. have suggested using the Delphi methodology to create standardised vignettes for use across LMICs, but this methodology also has its limitations as we have discussed in Chapter 5.

Another limitation with the vignette methodology, is around the interpretation of the results and understanding whether discrepancies between nurse triage ratings and the expert panel ratings arise because there is an inherent problem with the triage scale itself or because the nurses are incorrectly triaging patients. We have discussed this problem in more detail in Chapter 5.

An additional limitation when using vignettes for validation is the number of acuity levels. Triage scales range from three to five levels. Vignettes created for a three- or four-level system are not able to be used to validate scales with five levels. This further limits the ability to use vignettes for validation across various scales in low-resource settings.

The strengths and limitations of using surrogate markers in validity studies have been discussed in detail in the paper above. We believe that many of the associated limitations were mitigated as far as possible by virtue of the way in which MSF-supported facilities are equipped and run. An additional limitation however, is around the lack of standardisation that exists with regards the grouping of triage acuity levels from one validity study to another. Some studies categorise 'red' and 'orange' patients as high-priority, while 'green' patients are categorised as low-priority, completely ignoring ‘yellow’ patients;[18,84] other studies group 'red', ‘orange’ and ‘yellow’ patients as high-priority, and 'green' patients as low-priority.[20] These different groupings prevent comparability of study results and directly affect measures of over- and under-triage. They also make assumptions about the EC outcomes of ‘yellow’ and ‘orange’ patients that are questionable.

We have suggested a grouping system that tries to apply more rational assumptions about predicted EC outcomes linked to triage acuity level. Our system classifies 'red' patients as high-priority and 'green patients' as low-priority; it applies no rigid EC outcome predictions to 'yellow' and 'orange' patients who are neither low- nor high-priority but somewhere in between. To assess these middle triage acuities, we apply the logic that when moving from a low- to middle- to high-triage acuity, one should observe an increasing trend in hospital admission, death, resource utilisation or length of stay. Not only does the latter allow for the less predictable nature of middle acuity levels, but it can be used for triage scales that have three, four or five levels. What we are unable to elaborate on is how steep an increasing trend in any one of these EC outcomes should be with respect to the relative validity of a triage scale. To date, there are no published data on this; further development of this concept thus seems warranted. Having a standard framework that sets out the relative increases in EC outcomes across acuities linked to degrees of validity, would be similar to the triage weighted kappa proposed by Wulp et al. and applied when comparing triage officer rating with expert panel member ratings.

With increasing digitisation of health records and the increased use of mobile technology for collecting data in LMICs, we believe that the use of surrogate markers will become a simpler and more reliable way to test validity going forwards.

This study demonstrates the importance of making validity testing an on-going practice, especially in the various settings where the SATS is deployed. Defining a scale as merely valid or invalid is insufficient; the validity of a scale needs to be described relative to the specific context and population that it is being used in/for (the assumption being that validity results can reasonably be extrapolated to similar settings or populations). Our findings at GRC also highlight that the different disease burdens of a specific region/country might require the incorporation of additional discriminators in a scale. Thereafter, validity testing must be carried out again.

Another point of note is that in all the countries we studied, the 2010 version of the SATS was implemented. There is now a newer 2012 version available for implementation and when this

new version is implemented, it will be essential to repeat the validation studies and review the results.

Chapter Conclusion

For the context of LMICs, we have introduced a new and seemingly more logical approach to assessing validity of triage scales using surrogate markers as a proxy for true patient acuity.

Chapter 8: Discussion

Triage is an essential element in emergency care systems for low -settings. This thesis presents a series of papers looking at the reliability and validity of the SATS in an MSF context, across a variety of low-resource settings. The discussion that follows highlights major themes from each chapter and presents new literature that has been published during the course of the thesis.

In Chapter 3 we addressed implementation of SATS in a low-resource setting, highlighting lessons learned. The chapter considered the broader context in which triage scales should be implemented and focused on key stakeholder groups. There is little research around implementation of triage in low-resource settings.[78] However, a recent study in Rwanda found that when implementing the ETAT plus, getting broader stakeholder support was an important factor, consistent with our findings.[46] In this thesis we found an important group to be the non-medical stakeholders, including cleaners and security guards. This has led to the inclusion of cleaners and security guard as a key group to train in MSF's internal guidelines for triage training.

Independent of this thesis, the SATS has been implemented across multiple settings, including additional sites in Botswana, Haiti, Sierra Leone and Somaliland.[78,84,91,95] While the implementation was not robustly studied, successful implementation is further evidence that SATS is a suitable tool for emergency care in low-resource settings.

In Chapter 4, the thesis sought to provide a set of paediatric reference vignettes that were previously identified as a gap in the literature. We describe the creation of the reference vignettes via the Delphi method, but warned against the utility of these vignettes in dissimilar contexts. Standard reference vignettes are not without their limitations, as we describe in Chapter 5; however, in low-resource settings they would improve the ease of assessing validity of triage scales given the poor record keeping and non-standardised care identified by Twomey et al.[69] For the reference vignettes to be widely utilised, cases would need to be similar to a variety of settings. Currently, the reference paediatric vignettes that we

created, and the adult vignettes produced by Twomey et al., do not represent a broad enough range of presentations to be used widely across low-resource settings. The major limitation is that the vignettes are based on South African ECs and only cover disease profiles found in this context. If a wide range of disease specific vignettes were collected and validated, this would allow researchers to identify common diseases in their location and use only the appropriate vignettes. In keeping with the findings of this chapter and Chapter 5, we used vignettes based on locally presenting patients for further testing of reliability in Chapter 6.

Chapter 5 sought to use adult reference vignettes as a way of validation. This was the first time it was attempted outside of South Africa, and although we could assess reliability from the reference vignettes, validity testing had its limitations. There is no clear gold standard when trying to validate triage scales, and Akira et al. have argued that using criterion validity is a better way to assess validity than surrogate markers, which are a form of construct validity.[96] They argue that the aim of triage validation is to ensure the scale can identify true urgency of the patient and that the use of criterion validity (expert opinion using paper-based vignettes) is superior for this. As discussed in Chapter 2 based on Streiner et al., we have moved away from this type of classification.[56] In Chapter 5 we identified an inherent difficulty when interpreting results while using paper-based vignettes in validation: that it is difficult to assess whether the triage scale itself is not valid, or whether nurses could not correctly apply the scale. Using the results of our study we recommended the use of surrogate markers as a way of validation. This is in line with results from systematic reviews on triage scales in high-resource settings, which show that outcome markers are the most common method used as a validation technique.[27,31,96] We based this recommendation on the realities on the ground which included good record keeping available in MSF supported structures, and a standard package of care. We recognise that this might not be the same for all low-resource settings.

In Chapter 6 we assessed reliability in multiple low-resource settings. Learning from the lessons in Chapters 4 and 5, reliability was assessed using locally created paper-based vignettes across Haiti and Afghanistan. Using research nurses to triage the same patient in parallel (as some studies do) is impractical and expensive in low-resource settings.[65] We believe paper-based vignettes – based on local presentations – serve as the most appropriate

way to assess reliability in such settings. Across the four sites, ICC was 0.50, 0.58, 0.59 and 0.60, representing moderate agreement. A systematic review by Farrohknia et al. of high-resource settings showed a varying range of kappa from 0.2-0.9.[6] Our results show three out of the four sites having near identical and consistent reliability despite the diverse settings. This could suggest the ease of use of the SATS and the potential for varying human resources to reliably apply it. Our reliability findings are similar to subsequent SATS research in Ghana, which also found an inter-rater reliability of 0.59-0.60.[90] The results bring up an important question: when considering reliability, when is a triage scale reliable enough?

Landis and Koch, Cicchetti and Sparrow and Fleiss have all proposed criteria to measure the magnitude of kappa (Table 3).[56,62,63]

<table>
<tr><th>Kappa</th><th>Landis and Koch</th><th>Cicchetti and Sparrow</th><th>Fleiss</th></tr>
<tr><td>≤ 0</td><td>Poor</td><td rowspan="3">Poor</td><td rowspan="3">Poor</td></tr>
<tr><td>0.00-0.20</td><td>Slight</td></tr>
<tr><td>0.21-0.40</td><td>Fair</td></tr>
<tr><td>0.41-0.60</td><td>Moderate</td><td>Fair</td><td rowspan="2">Fair to Good</td></tr>
<tr><td>0.61-0.75</td><td rowspan="2">Substantial</td><td rowspan="3">Excellent</td></tr>
<tr><td>0.75-0.80</td><td rowspan="2">Excellent</td></tr>
<tr><td>0.81-1.00</td><td>Almost perfect</td></tr>
</table>

Table 3: Various published criteria for kappa[55] in triage studies

It has been suggested that any value less than 0.60 should not be bothered with.[56] Yet in low-resource settings, reliability is poorly studied and higher performance is not commonly achieved.[89,97,98] Reliability allows us to get the same results from different raters or the same rater at a different time. Triage should be the same irrespective of who is administering the scale. However, in low-resource settings it is difficult to practically say where our cut-off should be. Given the resource constraints and limited human resource skill available in low-resource settings, is moderate reliability good enough? Would having similar results between raters some of the time be better than not having triage at all? Literature on ETAT has not reported an overall kappa value and includes kappa for either only emergency cases or a certain set of signs.[45,99] One-two Triage (OTT) has been deemed a reliable tool for low-resource settings yet has reported a kappa of only 0.308.[89] SATS so far is the only adult and

paediatric triage scale that has been deployed widely in low-resource settings and shown consistent reliability: it has shown reliability of 0.76 for emergency physicians and 0.66 for nurses in South Africa.[59] In this thesis, the majority of results were around 0.60, which is similar to findings in Ghana.[90] Sierra Leone, Botswana and Pakistan have all reported high SATS reliability with values of 0.877, 0.87 and 0.77 respectfully.[70,71,95] This could suggest that SATS is a reliable tool but the nurses administrating it need more training where the scale has values less than 0.60.

Reliability is not a static value or an inherent value of a scale and is directly linked to the interaction of the raters and the population to which one wants to apply the measure. This means that the coefficient only has meaning to a specific population.[56] As we continue to use the SATS in various settings, it would be important to consistently assess reliability and validity of the scale for each specific population. In low-resource settings it would mean finding easier ways to do this, either with standardised vignettes, robust data collection tools or even electronic mobile data collection.

In Chapter 7 we applied the lessons from Chapter 5 and assessed validity using outcome measures. We applied this across Haiti, Afghanistan and Sierra Leone. We saw a correlation across all sites that led higher acuity patients to have higher rates of admission. Our findings confirm that SATS is valid in trauma only settings and could be used in mixed medical and trauma settings. Our findings, however, suggest that in paediatric only settings further revision and research is needed to validate SATS. The results found in the paediatric setting of Sierra Leone could be due to the use of the 2008 version of the SATS. This version was revised in 2012 to improve the use of the SATS specifically in paediatric settings and incorporated key elements of ETAT.[20] It would seem prudent going forward to implement the newer version of SATS and conduct validity assessments using the new scale. The interpretation of validity when using surrogate markers also needs to be standardised going forward. We have suggested a novel way to do this taking into account over- and under-triage and with a system that can be applied for three-, four- or five-level scales.[83]

New revisions and other triage scales for low-resource settings

At the beginning of this thesis the SATS was the only validated triage scale for adults and children in low-resource settings. MSF has been using the 2008 revision of the SATS and this version has been the basis of this thesis. In 2012, SATS released a new update which includes a revision of the paediatric scale which combines elements of ETAT and SATS. MSF has yet to implement the 2012 version of SATS in any of the hospitals in which it works.

As discussed in the literature review, only two other scales were found specifically for low-resource settings. ETAT has not released an adult or even older child version, and is still only available for children under five years of age. This limits it from being used as a triage scale in low-resource settings. The KTS showed poor sensitivity and was seen to be limited in its use as a triage scale.[49] Further studies in Malawi showed that that the KTS was not a strong enough predictor of outcomes to merit its use as a triage scale.[97] Another study suggested it should only be used after a physician has seen a patient as a re-triage method given the poor ability to predict hospital admission.[100] This use case is impractical given the short supply of physicians in low-resource settings.[101] There have been some attempts to use the MTS and ESI in low-resource settings with mixed results. The MTS was used in Brazil and showed an ability to predict severity of patients.[98] The ESI has been implemented in Iran but the authors concluded that it might not have optimal outcomes for low-resource settings.[102]

New triage scales for low-resource settings

Two new triage scales of note were found since the beginning of this thesis. Sick Children Require Emergency Evaluation Now (SCREEN) was designed in South Africa for paediatric clinic settings.[103] It focussed on a set of questions to be carried out by lay queue marshals to identify sick children as they came into the clinic. This scale was shown to have good reliability and to decrease waiting time for sick children in the South African setting.[104] Again this new scale was focused only on children, and in a Primary Care clinic setting, and as

such is limited as a triage scale in a hospital setting. It has currently only been validated in South Africa and would require further validation before wider use.[103]

OTT is a novel scale designed and implemented across multiple hospitals in Cambodia since 2013.[89] It has the advantage of needing only heart rate and blood oxygen saturation instead of full vital signs; however, it has shown unacceptably high levels of under-triage and only fair reliability with a Fleiss kappa of 0.308. Its use more widely will need further studies to ensure patient safety.

In early 2017 the WHO convened a consensus group of experts with a view of developing a triage scale for hospital settings in low-resource settings. It was agreed to be based on the SATS 2012 version and integrate elements of ETAT for quick identification of high acuity patients. The tool will consist of two charts that cover both adults (12 years of age and older) and children (under 12 years of age). It will consist of three levels – Red, Yellow and Green (with another colour for dead on arrival).

A unique property of the tool will be the capacity to be converted into a four-level scale for settings with higher resources. To be converted into a four-level scale, certain discriminators from both the Red and Yellow levels are amalgamated to create an Orange level. Currently this scale is in its pilot phase, and WHO is aiming to launch it in 2018. Once launched, it would be important to assess the reliability and validity across multiple low-resource settings.

Chapter 9: Conclusion and Recommendations

Conclusion

Triage is an essential component of ECS in low-resource settings, aiming to allow limited emergency care resources to be deployed to the maximum benefit. While multiple triage scales exist, SATS is one of very few that have been designed specifically for low-resource settings and that can be used in both adults and children.

Triage scales are most commonly assessed through reliability and validity. Before this thesis, there was little published on the reliability and validity of SATS – or indeed most triage tools – across multiple low-resource settings.

The 2011, MSF implementation of the SATS in their emergency centres allowed for this more in-depth assessment of SATS. Through this thesis we have shown:

- how to implement the SATS in a low-resource MSF facility – broad stakeholder engagement for cultural considerations is critical;
- that a set of standardised paediatric reference vignettes can improve reliability testing, although they have to be locally contextualised to be properly meaningful;
- that SATS has moderate to substantial reliability across five MSF sites in three countries; and
- that SATS is valid across five sites in three countries, in both trauma-only and mixed medical and trauma emergency settings, but that it's validity in paediatric only settings needs further work.

In addition, we have suggested a novel way to report validity when using surrogate markers in low-resource settings.

Our findings have added substantially to the body of evidence around SATS, which is a suitable tool for MSF and, more widely, for hospitals in low-resource settings.

Recommendations

Recommendations for MSF

- While our work demonstrated good results for SATS, it used the 2008 version, the 2012 version of SATS is significantly easier to use, and preliminary studies suggest it has similar test performance characteristics.
    - We recommend the implementation and studying of the 2012 version of SATS across MSF settings.
- Continuous learning is a key part of system change. To maintain the quality of triage in the various settings it is vital to have people able to train in the local language.
    - We recommend a train the trainer certification program for nursing and EC managers.
- Regular audits and validity assessments are recommended by the SATS group. Having robust data collection allows this to easily occur.
    - We recommend the implementation of a standard data collection tool across all MSF ECs, including triage acuity and outcomes.

Recommendations for further research

- This thesis provides further evidence that SATS is a reliable and valid triage scale across various low-resource settings. However, the settings in this thesis do not cover every context.
    - We recommend continuous reliability and validity studies as SATS is rolled out further, to ensure patient safety.
- We identified challenges with both acceptable levels of reliability and context- and disease-appropriate vignettes for triage scales.
    - We recommend studies to identify acceptable reliability levels in low-resource settings.
    - We recommend the development of a set of disease specific vignettes to ensure easy reliability testing across diverse low-resource settings.

- In Chapter 7 we proposed a novel approach for validation when using surrogate markers. Having the ability to standardise the validation process would help future research to compare across settings.
    - We recommend further studies assessing the acceptable proportion of admission rates across different acuity levels in low-resource settings is needed.

Bibliography:

1 Reynolds TA, Sawe H, Andrés M. Rubiano, *et al. Disease Control Priorities, Third Edition (Volume 9): Improving Health and Reducing Poverty*. The World Bank 2017. doi:10.1596/978-1-4648-0527-1

2 World Health Organization. Emergency care systems. Dep. Manag. NCDs, Disabil. Violence Inj. Prev. World Heal. Organ. http://www.who.int/emergencycare/systems/en/ (accessed 1 Feb2018).

3 Razzak J a, Hyder A a, Akhtar T, *et al.* Assessing emergency medical care in low income countries: a pilot study from Pakistan. *BMC Emerg Med* 2008;**8**:8. doi:10.1186/1471-227X-8-8

4 Razzak J a, Kellermann AL. Emergency medical care in developing countries: is it worthwhile? *Bull World Health Organ* 2002;**80**:900–5.

5 Molyneux E, Ahmad S, Robertson A. Improved triage and emergency care for children reduces inpatient mortality in a resource-constrained setting. Bull. World Health Organ. 2006;**84**:314–9. doi:10.2471/BLT.04.019505

6 Farrohknia N, Castrén M, Ehrenberg A, *et al.* Emergency Department Triage Scales and Their Components: A Systematic Review of the Scientific Evidence. *Scand J Trauma Resusc Emerg Med* 2011;**19**:42. doi:10.1186/1757-7241-19-42

7 Cooren F, Brummans BHJM, Charrieras D. The coproduction of organizational presence: A study of Mdecins Sans Frontires in action. *Hum Relations* 2008;**61**:1339–70. doi:10.1177/0018726708095707

8 Oredsson S, Jonsson H, Rognes J, *et al.* A systematic review of triage-related interventions to improve patient flow in emergency departments. *Scand J Trauma Resusc Emerg Med* 2011;**19**:43. doi:10.1186/1757-7241-19-43

9 Cronin JG. The introduction of the Manchester triage scale to an emergency department in the Republic of Ireland. Accid. Emerg. Nurs. 2003;**11**:121–5. doi:10.1016/S0965-2302(02)00218-7

10 Beveridge R, Clarke B, Janes L, *et al.* Implementation Guidelines for The Canadian Emergency Department Triage & Acuity Scale (CTAS). 1998.

11 Gilboy N, Tanabe P, Travers D, *et al. Emergency Severity Index (ESI) A Triage Tool for Emergency Department Care*. AHRQ Publication No. 12-0014 2011.

12 Considine J, LeVasseur SA, Villanueva E. The Australasian Triage Scale: Examining emergency department nurses' performance using computer and paper scenarios. *Ann Emerg Med* 2004;**44**:516–23. doi:10.1016/j.annemergmed.2004.04.007

13 Grouse A, Bishop R, Bannon A. The Manchester Triage System provides good reliability in an Australian emergency department. *Emerg Med J* 2009;**26**:484–6. doi:10.1136/emj.2008.065508

14 van der Wulp I, Schrijvers AJP, van Stel HF. Predicting admission and mortality with the Emergency Severity Index and the Manchester Triage System: a retrospective observational study. *Emerg Med J* 2009;**26**:506–9. doi:10.1136/emj.2008.063768

15 World Health Organization. Emergency Triage Assessment and Treatment (ETAT) course. Dep. Matern. Newborn, Child Adolesc. Heal. WHO. http://www.who.int/maternal_child_adolescent/documents/9241546875/en/ (accessed 1 Feb2018).

16 Twomey M. The South African Triage Scale (SATS). Emerg. Med. Soc. South Africa. http://emssa.org.za/sats/ (accessed 1 Feb2018).

17 Wildschut A, Mqolozana T. Shortage of nurses in South Africa: Relative or Absolute? 2008.

18 Rosedale K, Smith Z. The effectiveness of the South African Triage Score in a rural emergency department. *South African Med J* 2011;**101**:537–40.

19 Twomey M, de Sá A, Wallis L, *et al.* Inter-rater reliability of the South African Triage Scale: Assessing two different cadres of health care workers in a real time environment. *African J Emerg Med* 2011;**1**:113–8. doi:10.1016/j.afjem.2011.08.003

20 Twomey M, Cheema B, Buys H, *et al.* Vital signs for children at triage: A multicentre validation of the revised South African Triage Scale (SATS) for children. *South African Med J* 2013;**103**:3–8. doi:10.7196/samj.6877

21 Twomey M, Wallis LA, Thompson M Lou, *et al.* The South African triage scale (adult version) provides valid acuity ratings when used by doctors and enrolled nursing assistants. *African J Emerg Med* 2012;**2**:3–12. doi:10.1016/j.afjem.2011.08.014

22 Twomey M, Wallis LA, Myers JE. Limitations in validating emergency department triage scales. *Emerg Med J* 2007;**24**:477–9. doi:10.1136/emj.2007.046383

23 Nakao H, Ukai I, Kotani J. A review of the history of the origin of triage from a disaster medicine perspective. *Acute Med Surg* 2017;:379–84. doi:10.1002/ams2.293

24 Jenkins JL, McCarthy ML, Sauer LM, *et al.* Mass-Casualty Triage: Time for an Evidence-Based Approach. *Prehosp Disaster Med* 2008;**23**:3–8. doi:10.1017/S1049023X00005471

25 Hodge A, Hugman A, Varndell W, *et al.* A review of the quality assurance processes for the Australasian Triage Scale (ATS) and implications for future practice. *Australas Emerg Nurs J* 2013;**16**:21–9. doi:10.1016/j.aenj.2012.12.003

26 Travers DA, Waller AE, Bowling JM, *et al.* Five-level triage system more effective than three-level in tertiary emergency department. *J Emerg Nurs* 2002;**28**:395–400. doi:10.1067/men.2002.127184

27 Parenti N, Reggiani MLB, Iannone P, *et al.* A systematic review on the validity and reliability of an emergency department triage scale, the Manchester Triage System. *Int J Nurs Stud* 2014;**51**:1062–9. doi:10.1016/j.ijnurstu.2014.01.013

28 Ng CJ, Yen ZS, Tsai JCH, *et al.* Validation of the Taiwan triage and acuity scale: A new computerised five-level triage system. *Emerg Med J* 2011;**28**:1026–31. doi:10.1136/emj.2010.094185

29 Marsden J, Windle J, Mackway-Jones K, *et al. Emergency Triage*. 3rd ed. Chichester, UK: : John Wiley & Sons Ltd 2013. doi:10.1002/9781118299029

30 Wuerz RC, Milne LW, Eitel DR, *et al.* Reliability and Validity of a New Five-level Triage Instrument. *Acad Emerg Med* 2000;**7**:236–42. doi:10.1111/j.1553-2712.2000.tb01066.x

31 de Magalhães-Barbosa MC, Robaina JR, Prata-Barbosa A, *et al.* Validity of triage systems for paediatric emergency care: a systematic review. *Emerg Med J* 2017;:711–9. doi:10.1136/emermed-2016-206058

32 Pourasghar F, Tabrizi J, Ala A, *et al.* Validity of the electronic triage system in predicting patient outcomes in Tabriz, Iran: A cross-sectional study. *Bull Emerg Trauma* 2016;**4**:211–5.

33 Ng CJ, Hsu KH, Kuan JT, *et al.* Comparison between Canadian Triage and Acuity Scale and Taiwan Triage System in emergency departments. *J Formos Med Assoc* 2010;**109**:828–37. doi:10.1016/S0929-6646(10)60128-3

34 Henning B, Lydersen S, Døllner H. A reliability study of the rapid emergency triage and treatment system for children. *Scand J Trauma Resusc Emerg Med* 2016;**24**:1–6. doi:10.1186/s13049-016-0207-6

35 Chang YC, Ng CJ, Wu CT, *et al.* Effectiveness of a five-level Paediatric Triage System: An analysis of resource utilisation in the emergency department in Taiwan. *Emerg Med J* 2013;**30**:735–9. doi:10.1136/emermed-2012-201362

36 Westergren H, Ferm M, Häggström P. First evaluation of the paediatric version of the Swedish rapid emergency triage and treatment system shows good reliability. *Acta Paediatr* 2014;**103**:305–8. doi:10.1111/apa.12491

37 Pérez N, Nissen L, Nielsen RF, *et al.* The predictive validity of RETTS-HEV as an acuity triage tool in the emergency department of a Danish Regional Hospital. *Eur J Emerg Med* 2016;**23**:33–7. doi:10.1097/MEJ.0000000000000173

38 Parenti N, Ferrara L, Bacchi Reggiani ML, *et al.* Reliability and validity of two four-level emergency triage systems. *Eur J Emerg Med* 2009;**16**:115–20. doi:10.1097/MEJ.0b013e328310b594

39 Taboulet P, Moreira V, Haas L, *et al.* Triage with the French Emergency Nurses Classification in Hospital scale: reliability and validity. *Eur J Emerg Med* 2009;**16**:61–7. doi:10.1097/MEJ.0b013e328304ae57

40 Bruijns SR, Wallis LA, Burch VC. A prospective evaluation of the Cape triage score in the emergency department of an urban public hospital in South Africa. *Emerg Med J* 2008;**25**:398–402. doi:10.1136/emj.2007.051177

41 Hansoti B, Jenson A, Keefe D, *et al.* Reliability and validity of pediatric triage tools evaluated in Low resource settings: A systematic review. *BMC Pediatr* 2017;**17**. doi:10.1186/s12887-017-0796-x

42 Opiro K, Wallis L, Ogwang M. Assessment of hospital-based adult triage at emergency receiving areas in hospitals in Northern Uganda. *Afr Health Sci* 2017;**17**:481–90. doi:10.4314/ahs.v17i2.23

43 Baker T, Lugazia E, Eriksen J, *et al.* Emergency and critical care services in Tanzania: a survey of ten hospitals. *BMC Health Serv Res* 2013;**13**:140. doi:10.1186/1472-6963-13-140

44 Dünser MW, Baelani I, Ganbold L. A review and analysis of intensive care medicine in the least developed countries*. *Crit Care Med* 2006;**34**:1234–42. doi:10.1097/01.CCM.0000208360.70835.87

45 Gove S, Tamburlini G, Molyneux E, *et al.* Development and technical basis of simplified guidelines for emergency triage assessment and treatment in developing

countries. *Arch Dis Child* 1999;**81**:473–7. doi:10.1136/adc.81.6.473

46 Hategeka C, Mwai L, Tuyisenge L. Implementing the Emergency Triage, Assessment and Treatment plus admission care (ETAT+) clinical practice guidelines to improve quality of hospital care in Rwandan district hospitals: healthcare workers' perspectives on relevance and challenges. *BMC Health Serv Res* 2017;**17**:1–12. doi:10.1186/s12913-017-2193-4

47 Hategeka C, Shoveller J, Tuyisenge L, *et al.* Pediatric emergency care capacity in a low-resource setting: An assessment of district hospitals in Rwanda. *PLoS One* 2017;**12**:1–13. doi:10.1371/journal.pone.0173233

48 MacLeod JBA, Kobusingye O, Frost C, *et al.* A Comparison of the Kampala Trauma Score (KTS) with the Revised Trauma Score (RTS), Injury Severity Score (ISS) and the TRISS Method in a Ugandan Trauma Registry: Is Equal Performance Achieved with Fewer Resources? *Eur J Trauma* 2003;**29**:392–8. doi:10.1007/s00068-003-1277-5

49 Kobusingye O, Frost C, Lett R, *et al.* Kampala Trauma Score (KTS): Is it a New Triage Tool? *East Cent African J Surg* 2007;**12**:74–82.

50 Wallis LA, Twomey M. Workload and casemix in Cape Town emergency departments. *South African Med J* 2007;**97**:1276–80.http://www.ncbi.nlm.nih.gov/pubmed/18264609

51 Gottschalk SB. The cape triage score: a new triage system South Africa. Proposal from the Cape Triage Group. *Emerg Med J* 2006;**23**:149–53. doi:10.1136/emj.2005.028332

52 Twomey M. *Performance characteristics of the South African Triage Scale (Adult version)*. 2011.http://hdl.handle.net/11427/11629

53 Harrison H-L, Raghunath N, Twomey M. Emergency triage, assessment and treatment at a district hospital in Malawi. *Emerg Med J* 2012;**29**:924–5. doi:10.1136/emermed-2011-200472

54 Rominski S, Bell SA, Oduro G, *et al.* The implementation of the South African Triage Score (SATS) in an urban teaching hospital, Ghana. *African J Emerg Med* 2014;**4**:71–5. doi:10.1016/j.afjem.2013.11.001

55 Bergeron S, Gouin S, Bailey B, *et al.* Agreement among pediatric health care professionals with the pediatric Canadian triage and acuity scale guidelines. Pediatr. Emerg. Care. 2004;**20**:514–8.http://www.ncbi.nlm.nih.gov/pubmed/15295246

56 Streiner DL, Norman GR, Cairney J. *Health Measurement Scales: A practical guide to*

their development and use. 5th ed. New York: : Oxford University Press 2015. doi:10.1093/acprof

57 Kottner J, Audigé L, Brorson S, *et al.* Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *J Clin Epidemiol* 2011;**64**:96–106. doi:10.1016/j.jclinepi.2010.03.002

58 van der Wulp I. *Reliability and validity of emergency department triage systems*. 2010.http://igitur-archive.library.uu.nl/dissertations/2010-0209-200146/UUindex.html

59 Twomey M, Wallis L, Thompson M Lou, *et al.* The South African Triage Scale (adult version) provides reliable acuity ratings. *Int Emerg Nurs* 2012;**20**:142–50. doi:10.1016/j.ienj.2011.08.002

60 Worster A, Gilboy N, Fernandes CM, *et al.* Assessment of inter-observer reliability of two five-level triage and acuity scales: a randomized controlled trial. *CJEM* 2004;**6**:240–5. doi:10.1017/S1481803500009192

61 Nelson JC, Pepe MS. Statistical description of interrater variability in ordinal ratings. *Stat Methods Med Res* 2000;**9**:475–96. doi:10.1177/096228020000900505

62 Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data Data for Categorical of Observer Agreement The Measurement. *Biometrics* 1977;**33**:159–74.

63 Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. *Psychol Bull* 1979;**86**:420–8. doi:10.1037/0033-2909.86.2.420

64 Berk RA. Generalizability of behavioral observations: a clarification of interobserver agreement and interobserver reliability. *Am J Ment Defic* 1979;**83**:460–72.

65 Worster A, Sardo A, Eva K, *et al.* Triage tool inter-rater reliability: a comparison of live versus paper case scenarios. *J Emerg Nurs* 2007;**33**:319–23. doi:10.1016/j.jen.2006.12.016

66 Eitel DR, Travers DA, Rosenau AM, *et al.* The Emergency Severity Index triage algorithm version 2 is reliable and valid. *Acad Emerg Med* 2003;**10**:1070–80. doi:10.1197/S1069-6563(03)00350-6

67 Göransson K, Ehrenberg A, Marklund B, *et al.* Accuracy and concordance of nurses in emergency department triage. *Scand J Caring Sci* 2005;**19**:432–8. doi:10.1111/j.1471-6712.2005.00372.x

68 Jafari-Rouhi AH, Sardashti S, Taghizadieh A, *et al.* The Emergency Severity Index, version 4, for pediatric triage: A reliability study in Tabriz Children's Hospital, Tabriz, Iran. *Int J Emerg Med* 2013;**6**:4–8. doi:10.1186/1865-1380-6-36

69 Twomey M, Wallis LA, Myers JE. Evaluating the construct of triage acuity against a set of reference vignettes developed via modified Delphi method. *Emerg Med J* 2014;**31**:562–6. doi:10.1136/emermed-2013-202352

70 Twomey M, Mullan PC, Torrey SB, *et al.* The Princess Marina Hospital accident and emergency triage scale provides highly reliable triage acuity ratings. *Emerg Med J* 2012;**29**:650–3. doi:10.1136/emermed-2011-200503

71 Dalwai MK, Twomey M, Maikere J, *et al.* Reliability and accuracy of the South African Triage Scale when used by nurses in the emergency department of Timergara Hospital, Pakistan. *South African Med J* 2014;**104**:372. doi:10.7196/samj.7604

72 van der Wulp I, van Baar ME, Schrijvers a JP. Reliability and validity of the Manchester Triage System in a general emergency department patient population in the Netherlands: results of a simulation study. *Emerg Med J* 2008;**25**:431–4. doi:10.1136/emj.2007.055228

73 Olofsson P, Gellerstedt M, Carlström ED. Manchester Triage in Sweden - interrater reliability and accuracy. *Int Emerg Nurs* 2009;**17**:143–8. doi:10.1016/j.ienj.2008.11.008

74 Wuerz RC, Travers D, Gilboy N, *et al.* Implementation and refinement of the emergency severity index. *Acad Emerg Med* 2001;**8**:170–6. doi:10.1111/j.1553-2712.2001.tb01283.x

75 Gravel J, Gouin S, Goldman RD. The Canadian Triage and Acuity Scale for Children : A Prospective Multicenter Evaluation. *YMEM* 2012;**60**:71–77.e3. doi:10.1016/j.annemergmed.2011.12.004

76 Rotondo M, Cribari C, Smith S. *Resources for optimal care of the injured patient*. Chicago: : American College of Surgeons, Committee on Trauma 2014.

77 Mohan D. Assessing the Feasibility of the American College of Surgeons' Benchmarks for the Triage of Trauma Patients. *Arch Surg* 2011;**146**:786. doi:10.1001/archsurg.2011.43

78 Rouhani SA, Aaronson E, Jacques A, *et al.* Evaluation of the implementation of the South African Triage System at an academic hospital in central Haiti. *Int Emerg Nurs*

2017;**33**:26–31. doi:10.1016/j.ienj.2016.12.004

79 Aladwani AM. Change management strategies for successful ERP implementation. *Bus Process Manag* 2001;**7**:266–75. doi:https://doi.org/10.1108/14637150110392764

80 Joseph R, Reigeluth C. The Systemic Change Process in Education: A Conceptual Framework. *Contemp Educ Technol* 2010;**1**:97–117.

81 Lewis LK. An Organizational Stakeholder Model of Change Implementation Communication. *Commun Theory* 2007;**17**:176–204. doi:10.1111/j.1468-2885.2007.00291.x

82 Kobusingye OC, Hyder A, Bishai D, *et al.* Emergency medical systems in low- and middle-income countries: recommendations for action. *Bull World Health Organ* 2005;**83**:626–31. doi:/S0042-96862005000800017

83 Dalwai M, Valles P, Twomey M, *et al.* Is the South African Triage Scale valid for use in Afghanistan, Haiti and Sierra Leone? *BMJ Glob Heal* 2017;**2**:e000160. doi:10.1136/bmjgh-2016-000160

84 Sunyoto T, Van den Bergh R, Valles P, *et al.* Providing emergency care and assessing a patient triage system in a referral hospital in Somaliland: a cross-sectional study. *BMC Health Serv Res* 2014;**14**:531. doi:10.1186/s12913-014-0531-3

85 Jiménez JG, Murray MJ, Beveridge R, *et al.* Implementation of the Canadian Emergency Department Triage and Acuity Scale (CTAS) in the Principality of Andorra: Can triage parameters serve as emergency department quality indicators? *CJEM* 2003;**5**:315–22.http://www.ncbi.nlm.nih.gov/pubmed/17466139

86 Maritz D, Hodkinson P, Wallis L. Identification of performance indicators for emergency centres in South Africa: results of a Delphi study. *Int J Emerg Med* 2010;**3**:341–9. doi:10.1007/s12245-010-0240-6

87 Hodkinson PW, Wallis LA. Emergency Medicine in the Developing World: A Delphi Study. *Acad Emerg Med* 2010;**17**:765–74. doi:10.1111/j.1553-2712.2010.00791.x

88 Wallis L, Carley S, Hodgetts CT. A procedure based alternative to the injury severity score for major incident triage of children: results of a Delphi consensus process. *Emerg Med J* 2006;**23**:291–5. doi:10.1136/emj.2005.025312

89 Khan A, Mahadevan S, Dreyfuss A, *et al.* One-two-triage: validation and reliability of a novel triage system for low-resource settings. *Emerg Med J* 2016;**33**:709–15.

doi:10.1136/emermed-2015-205430

90 Gyedu A, Agbedinu K, Dalwai M, *et al.* Triage capabilities of medical trainees in Ghana using the South African triage scale: An opportunity to improve emergency care. *Pan Afr Med J* 2016;**24**. doi:10.11604/pamj.2016.24.294.8728

91 Tshitenge ST, Ogunbanjo GA, Mbuka DO. The effectiveness of the South African Triage Toll use in Mahalapye District Hospital - Emergency Department, Botswana. *African J Prim Heal care Fam Med* 2016;**8**:e1–5. doi:10.4102/phcfm.v8i1.1030

92 Dalwai M, Tayler-Smith K, Twomey M, *et al.* Developing a reference standard for assessing paediatric triage scales in resource poor settings. *African J Emerg Med* 2015;**5**:181–4. doi:10.1016/j.afjem.2015.08.002

93 Dalwai MK, Tayler-Smith K, Trelles M, *et al.* Implementation of a triage score system in an emergency room in Timergara, Pakistan [Short communication]. *Public Heal Action* 2013;**3**:43–5. doi:10.5588/pha.12.0083

94 Rutschmann OT, Kossovsky M, Geissbühler A, *et al.* Interactive triage simulator revealed important variability in both process and outcome of emergency triage. *J Clin Epidemiol* 2006;**59**:615–21. doi:10.1016/j.jclinepi.2005.11.003

95 Lowsby R, Kamara C, Kamara M, *et al.* Evaluation du triage réalisé par les infirmières à l'hôpital de Connaught, en Sierra Leone, dans la période suivant directement la crise Ebola. *African J Emerg Med* 2017;**7**:51–5. doi:10.1016/j.afjem.2016.10.003

96 Kuriyama A, Urushidani S, Nakayama T. Five-level emergency triage systems: variation in assessment of validity. *Emerg Med J* 2017;**34**:703–10. doi:10.1136/emermed-2016-206295

97 Haac B, Varela C, Geyer A, *et al.* The utility of the Kampala trauma score as a triage tool in a sub-Saharan African trauma cohort. *World J Surg* 2015;**39**:356–62. doi:10.1007/s00268-014-2830-6

98 Pinto Júnior D, Salgado PDO, Chianca TCM. Predictive validity of the Manchester Triage System: evaluation of outcomes of patients admitted to an emergency department. *Rev Lat Am Enfermagem* 2012;**20**:1041–7. doi:10.1590/S0104-11692012000600005

99 Tamburlini G, Di Mario S, Maggi RS, *et al.* Evaluation of guidelines for emergency triage assessment and treatment in developing countries. *Arch Dis Child* 1999;**81**:478–82.

100 Gardner A, Forson PK, Oduro G, *et al.* Diagnostic accuracy of the Kampala Trauma Score using estimated Abbreviated Injury Scale scores and physician opinion. *Injury* 2017;**48**:177–83. doi:10.1016/j.injury.2016.11.022

101 Naicker S, Plange-Rhule J, Tutt RC, *et al.* Shortage of healthcare workers in developing countries--Africa. *Ethn Dis* 2009;**19**:S1-60–4.

102 Mirhaghi A, Kooshiar H, Esmaeili H, *et al.* Outcomes for emergency severity index triage implementation in the emergency department. *J Clin Diagnostic Res* 2015;**9**:OC04-OC07. doi:10.7860/JCDR/2015/11791.5737

103 Hansoti B. *Prioritization of critically unwell children in Low resource Primary Healthcare Centres in Cape Town, South Africa*. 2017.http://hdl.handle.net/11427/25008

104 Hansoti B, Dalwai M, Katz J, *et al.* Prioritising the care of critically ill children: a pilot study using SCREEN reduces clinic waiting times. *BMJ Glob Heal* 2016;**1**:e000036. doi:10.1136/bmjgh-2016-000036

Appendices

Appendix 1: MSF, Ethics Review Board approval

Ethics Review Board
Instituted by Médecins Sans Frontières

Dr Rony Zachariah
Medical Department
Médecins Sans Frontières
Rue Dupré, 94/ Dupréstraat, 94
1090 Brussels, Belgium

25 November 2012

Re: Approval of the study protocols *Reliability of the South African Triage Score system in an Emergency Room setting in Pakistan* and *Validity of the South African Triage Score system in an Emergency Room setting in Pakistan*, no date (submitted 13 October 2012)

Dear Dr Zachariah,

Many thanks for submitting the above-mentioned proposals to our review. Ethical issues are minimal and well addressed. We thus approve both protocols for a period of 12 months from initiation of the study. The study must be initiated within the next 12 months. If this is not the case the approval of the protocols is no longer valid.

Any subsequent changes you might wish to make to the project must be notified to the Ethics Review Board for further consideration and approval. We would appreciate receiving the final research report.

We would like to draw your attention to the fact that the ERB will routinely check the reported and published outcome measure(s) against the outcome measure(s) initially approved in the protocol. If the outcome measures published differ from the proposal, the ERB should be consulted beforehand. There may be good reasons for the change, but as any other alteration in the approved protocol it should be assessed on ethical grounds.

We wish you much success with the research.

Yours sincerely,

Doris Schopper
Chairperson, Ethics Review Board

Appendix 2: Sierra Leone, Ethics approval

GOVERNMENT OF SIERRA LEONE
Office of the Sierra Leone Ethics and Scientific Review Committee
Directorate of Training, Non-Communicable Diseases and Research
Connaught Hospital
Ministry of Health and Sanitation

5th November, 2012

Dr. Yvonne Nzomukunda
Medical Coordinator
MSF-OCB
9 Sall Drive, Cockle Bay
Aberdeen Road, Freetown

Dear Dr. Nzomukunda,

Implementation of the South African Triage Scale Use among Children and Infants in Bo, Sierra Leone

This letter refers to the above proposed study submitted for review.

The Committee hereby grants Ethical and Scientific clearance for this study to be conducted in Sierra Leone

The Committee stipulates as follows: that,

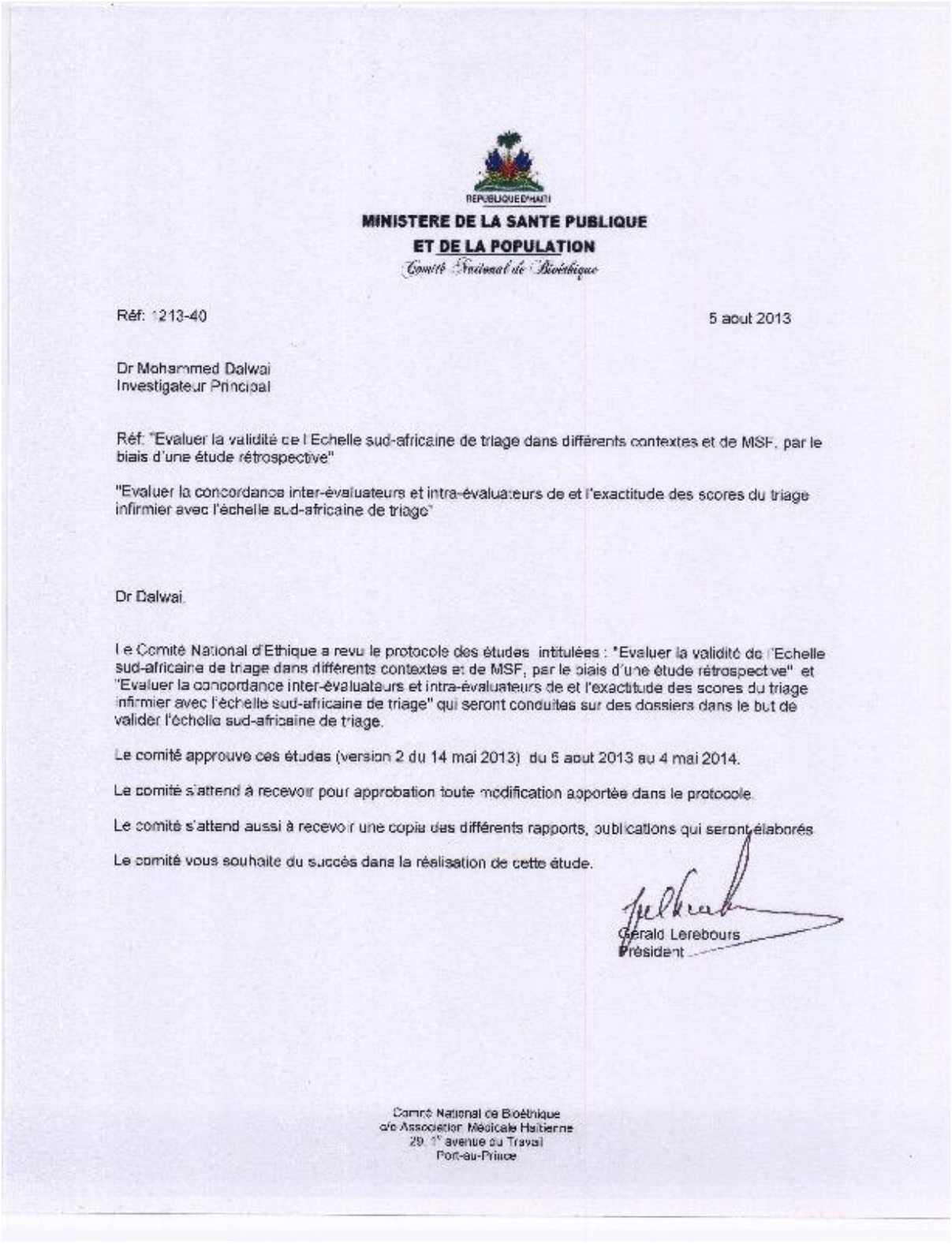
1. It must be notified in advance, if you decide to amend the research design and/or methodology at any time during the conduct of the study.
2. It must be informed if for any reason, the study is terminated prematurely.
3. On the conclusion of the study, you submit a report or any publication based on the study.

Yours sincerely,

Professor Hector G. Morgan
Chairman, SLESRC

Email: hgmorg2007@yahoo.com / williettav@yahoo.com

Appendix 3: Haiti, Ethics approval



REPUBLIQUE D'HAITI

MINISTERE DE LA SANTE PUBLIQUE
ET DE LA POPULATION
Comité National de Bioéthique

Réf: 1213-40

5 aout 2013

Dr Mohammed Dalwai
Investigateur Principal

Réf: "Evaluer la validité de l Echelle sud-africaine de triage dans différents contextes et de MSF, par le biais d'une étude rétrospective"

"Evaluer la concordance inter-évaluateurs et intra-évaluateurs de et l'exactitude des scores du triage infirmier avec l'échelle sud-africaine de triage"

Dr Dalwai.

Le Comité National d'Ethique a revu le protocole des études intitulées : "Evaluer la validité de l'Echelle sud-africaine de triage dans différents contextes et de MSF, par le biais d'une étude rétrospective" et "Evaluer la concordance inter-évaluateurs et intra-évaluateurs de et l'exactitude des scores du triage infirmier avec l'échelle sud-africaine de triage" qui seront conduites sur des dossiers dans le but de valider l'échelle sud-africaine de triage.

Le comité approuve ces études (version 2 du 14 mai 2013) du 5 aout 2013 au 4 mai 2014.

Le comité s'attend à recevoir pour approbation toute modification apportée dans le protocole.

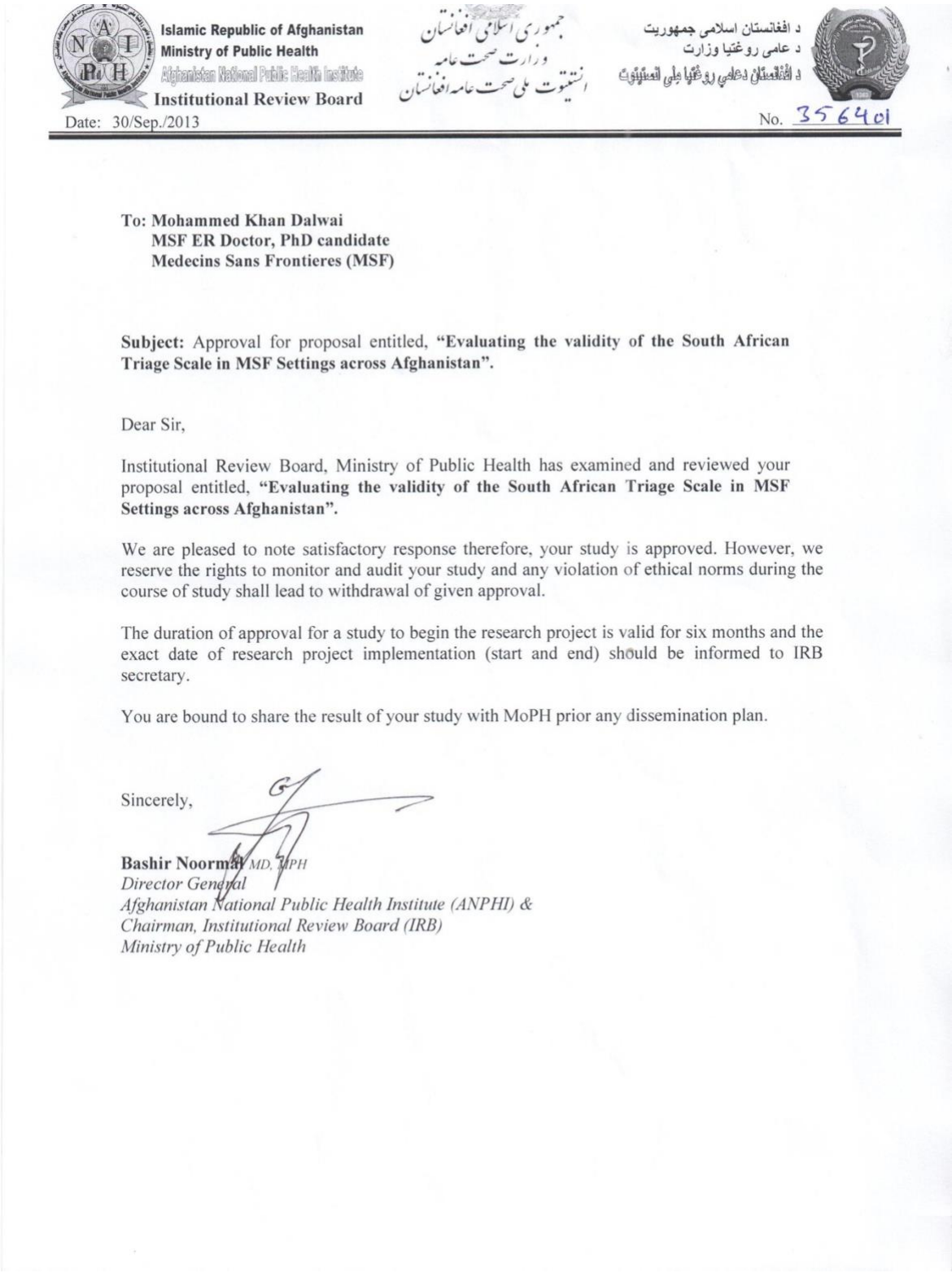
Le comité s'attend aussi à recevoir une copie des différents rapports, publications qui seront élaborés.

Le comité vous souhaite du succès dans la réalisation de cette étude.

Gérald Lerebours
Président

Comité National de Bioéthique
c/o Association Médicale Haitienne
29, 1ère avenue du Travail
Port-au-Prince

Appendix 4: Afghanistan, Ethics approval



Islamic Republic of Afghanistan
Ministry of Public Health
Afghanistan National Public Health Institute
Institutional Review Board

جمهوری اسلامی افغانستان
وزارت صحت عامه
انستیتوت ملی صحت عامه افغانستان

د افغانستان اسلامی جمهوریت
د عامی روغتیا وزارت
د افغانستان د عامي روغتيا ملي انستيتوت

Date: 30/Sep./2013

No. 356401

To: Mohammed Khan Dalwai
MSF ER Doctor, PhD candidate
Medecins Sans Frontieres (MSF)

Subject: Approval for proposal entitled, **"Evaluating the validity of the South African Triage Scale in MSF Settings across Afghanistan".**

Dear Sir,

Institutional Review Board, Ministry of Public Health has examined and reviewed your proposal entitled, **"Evaluating the validity of the South African Triage Scale in MSF Settings across Afghanistan".**

We are pleased to note satisfactory response therefore, your study is approved. However, we reserve the rights to monitor and audit your study and any violation of ethical norms during the course of study shall lead to withdrawal of given approval.

The duration of approval for a study to begin the research project is valid for six months and the exact date of research project implementation (start and end) should be informed to IRB secretary.

You are bound to share the result of your study with MoPH prior any dissemination plan.

Sincerely,

Bashir Noormal *MD, MPH*
Director General
Afghanistan National Public Health Institute (ANPHI) &
Chairman, Institutional Review Board (IRB)
Ministry of Public Health

To simplify the process we submitted both the reliability and validity under one protocol named ''Evaluating the validity of the South African Triage Scale in MSF settings across Afghanistan.

Appendix 5: UCT, Human Research Ethics Committee approval

UNIVERSITY OF CAPE TOWN

Faculty of Health Sciences
Human Research Ethics Committee
Room E52-24 Groote Schuur Hospital Old Main Building
Observatory 7925
Telephone [021] 406 6338 • **Facsimile** [021] 406 6411
e-mail: linsey.samuels@uct.ac.za
Website: www.health.uct.ac.za/research/humanethics/forms

18 September 2013

HREC REF: 570/2013

Dr M Dalwai
c/o Dr M Twomey
Emergency Medicine
Gatesville

Dear Dr Dalwai

PROJECT TITLE: ASSESSING THE INTER, INTRA-RATER RELIABILITY OF NURSE TRIAGE RATINGS WHEN USING THE SOUTH AFRICAN TRIAGE SCALE IN DIFFERENT MSF EMERGENCY CENTRES

Thank you for submitting your study to the Faculty of Health Sciences Human Research Ethics Committee for review.

It is a pleasure to inform you that the HREC has **formally approved** the above-mentioned study.

Approval is granted for one year till the 30th September 2014

Please submit a progress form, using the standardised Annual Report Form if the study continues beyond the approval period. Please submit a Standard Closure form if the study is completed within the approval period.
(Forms can be found on our website: www.health.uct.ac.za/research/humanethics/forms)

Please note that the ongoing ethical conduct of the study remains the responsibility of the principal investigator.

Please quote the HREC. REF in all your correspondence.

Yours sincerely

PROFESSOR M BLOCKMAN
CHAIRPERSON, FHS HUMAN ETHICS
Federal Wide Assurance Number: FWA00001637.
Institutional Review Board (IRB) number: IRB00001938
This serves to confirm that the University of Cape Town Human Research Ethics Committee complies to the Ethics Standards for Clinical Research with a new drug in patients, based on the Medical Research Council (MRC-SA), Food and Drug Administration (FDA-USA), International Convention on Harmonisation Good Clinical Practice (ICH GCP) and Declaration of Helsinki guidelines.
The Human Research Ethics Committee granting this approval is in compliance with the ICH Harmonised Tripartite Guidelines E6: Note for Guidance on Good Clinical Practice (CPMP/ICH/135/95) and FDA Code Federal Regulation Part 50, 56 and 312.

L Samuels

UNIVERSITY OF CAPE TOWN

Faculty of Health Sciences
Human Research Ethics Committee
Room E52-24 Groote Schuur Hospital Old Main Building
Observatory 7925
Telephone [021] 406 6338 • **Facsimile** [021] 406 6411
e-mail: linsey.samuels@uct.ac.za
Website: www.health.uct.ac.za/research/humanethics/forms

18 September 2013

HREC REF: 571/2013

Dr M Dalwai
c/o Dr M Twomey
Emergency Medicine
Gatesville

Dear Dr Dalwai

PROJECT TITLE: ASSESSING THE VALIDITY OF SOUTH AFRICA TRIAGE SCALE IN DIFFERENT MSF EMERGENCY CENTRES, BY RETROSPECTIVE AUDIT REVIEW

Thank you for submitting your study to the Faculty of Health Sciences Human Research Ethics Committee for review.

It is a pleasure to inform you that the HREC has **formally approved** the above-mentioned study.

Approval is granted for one year till the 30th September 2014

Please submit a progress form, using the standardised Annual Report Form if the study continues beyond the approval period. Please submit a Standard Closure form if the study is completed within the approval period.
(Forms can be found on our website: www.health.uct.ac.za/research/humanethics/forms)

Please note that the ongoing ethical conduct of the study remains the responsibility of the principal investigator.

Please quote the HREC. REF in all your correspondence.

Yours sincerely

PROFESSOR M BLOCKMAN
CHAIRPERSON, FHS HUMAN ETHICS
Federal Wide Assurance Number: FWA00001637.
Institutional Review Board (IRB) number: IRB00001938
This serves to confirm that the University of Cape Town Human Research Ethics Committee complies to the Ethics Standards for Clinical Research with a new drug in patients, based on the Medical Research Council (MRC-SA), Food and Drug Administration (FDA-USA), International Convention on Harmonisation Good Clinical Practice (ICH GCP) and Declaration of Helsinki guidelines.
The Human Research Ethics Committee granting this approval is in compliance with the ICH Harmonised Tripartite Guidelines E6: Note for Guidance on Good Clinical Practice (CPMP/ICH/135/95) and FDA Code Federal Regulation Part 50, 56 and 312.

L Samuels

UNIVERSITY OF CAPE TOWN

Faculty of Health Sciences
Human Research Ethics Committee
Room E52-24 Groote Schuur Hospital Old Main Building
Observatory 7925
Telephone [021] 406 6626 • **Facsimile** [021] 406 6411
e-mail: shuretta.thomas@uct.ac.za
Website: www.health.uct.ac.za/research/humanethics/forms

1 March 2013

HREC REF: 130/2013

Dr M Dalwai
c/o Prof L Wallis
Emergency Medicine
PO Box 8
Gatesville
7764

Dear Dr Dalwai

PROJECT TITLE: FORMULATION OF PAEDIATRIC REFERENCE VIGNETTES USING THE MODIFIED DELPHI PROCESS

Thank you for submitting your study to the Faculty of Health Sciences Human Research Ethics Committee for review.

Before the study can be approved the PI need to address the following issues raised by the HREC:

1. Please explain what procedures will be put in place to ensure that the panel of experts are blinded to each other.
2. The participant information sheet has been submitted without a consent document (it seems to be more a recruitment email). Please supply both the participant information sheet and consent documents.
3. Please explain how confidentiality and anonymity will be maintained in the participant information sheet.
4. Please describe any risks and benefits associated with taking part in the study.
5. Please include contact details for the researcher and supervisors in the information sheet.
6. Please include contact details of the Human Research Ethics Committee in the information sheet, in case participants have any questions or concerns regarding their rights as research participants.

Please quote the HREC REF in all your correspondence.

Please note that the ongoing ethical conduct of the study remains the responsibility of the principal investigator.

Yours sincerely

pp TBurgess

PROFESSOR M BLOCKMAN
CHAIRPERSON, FHS HUMAN ETHICS

S Thomas

UNIVERSITY OF CAPE TOWN

Faculty of Health Sciences
Human Research Ethics Committee
Room E52-24 Groote Schuur Hospital Old Main Building
Observatory 7925
Telephone [021] 406 6338 • **Facsimile** [021] 406 6411
e-mail: shuretta.thomas@uct.ac.za
Website: www.health.uct.ac.za/research/humanethics/forms

16 May 2013

HREC REF: 281/2013

Dr M Dalwai
c/o Prof L Wallis & Dr M Twomey
Emergency Medicine
Surgery, J-Floor
OMB

Dear Dr Dalwai

PROJECT TITLE: RELIABILITY AND VALIDITY OF THE SOUTH AFRICAN TRIAGE SCALE IN AN EMERGENCY ROOM SETTING IN PAKISTAN

Thank you for responding to the issues raised by the Faculty of Health Sciences Human Research Ethics Committee in your letter dated 15th May 2013.

It is a pleasure to inform you that the HREC has **formally approved** the above-mentioned study.

Approval is granted for one year till the 28th May 2014

Please submit a progress form, using the standardised Annual Report Form if the study continues beyond the approval period. Please submit a Standard Closure form if the study is completed within the approval period.
(Forms can be found on our website: www.health.uct.ac.za/research/humanethics/forms)

Please note that the ongoing ethical conduct of the study remains the responsibility of the principal investigator.

Please quote the HREC. REF in all your correspondence.

Yours sincerely

pp TuBuy BS
PROFESSOR M BLOCKMAN
CHAIRPERSON, FHS HUMAN ETHICS
Federal Wide Assurance Number: FWA00001637.
Institutional Review Board (IRB) number: IRB00001938
This serves to confirm that the University of Cape Town Human Research Ethics Committee complies to the Ethics Standards for Clinical Research with a new drug in patients, based on the Medical Research Council (MRC-SA), Food and Drug Administration (FDA-USA), International Convention on Harmonisation Good Clinical Practice (ICH GCP) and Declaration of Helsinki guidelines.
The Human Research Ethics Committee granting this approval is in compliance with the ICH Harmonised Tripartite Guidelines E6: Note for Guidance on Good Clinical Practice (CPMP/ICH/135/95) and FDA Code Federal Regulation Part 50, 56 and 312.

s.thomas

Appendix 6: Paediatric Reference Vignettes

Paediatric Vignettes that achieved >80% consensus

Emergency = 4	Needs to be seen immediately
1. A panicked mother runs into the triage area with a floppy child in her arms. She is screaming that her child cannot breath. The child is 2 years old and has abnormal breathing sounds. It sounds like he is grunting. You look at his tongue and it looks bluish. He has chest in drawing and has nasal flaring. Vital Signs: RR 41, HR 111 and a temp of 37.	
2. A 6-month-old boy presents with difficulty in breathing. His mother says he has been sick since last night. At triage you find he has rapid shallow breathing with in drawing of chest. He is grunting and looks unwell. He is still alert when you examine him. Vital Signs: RR, 56 HR, 200 and a temp of 40.	
3. A mother complains that her 4-year-old son has a rash. He developed it over the last 2 days. At triage you find a rash that is non blanching and non painful. It looks like bruises under the skin. The child is drowsy and lying still in his mothers arms. Vital Signs: RR 19, HR 95, Temp 36.9	
4. A mother says her 4 year old is having a very high fever for the last 3 days. No coughing or vomiting just a high fever. At triage you find the patient is weak and disorientated. He doesn't respond to you when you call his name and had to be brought into the ER on a stretcher. He is breathing quietly and slowly. Vital Signs: RR 20, HR 137, Temp 38.9	

Very Urgent = 8	Needs to be seen within 10mins
1. A first time mother is worried and brings in her 4 day old newborn because he is yellow. He has been yellow since birth. On examination you find a yellow child, he is still drinking, but seems sleepy. No other problems found. Vitals Signs: RR 52, HR163, Temp 36.6	
2. A 3 years old is brought in by her aunt. The child was at school but had to leave early because she was having very bad stomach cramps. The aunt also noticed that her niece was very hot and is feeling nauseous. The child looks to be in severe pain and not looking very happy (7/10). She is not having any diarrhoea but says her whole stomach is paining. Vital Signs: RR46, HR 157, Temp 38.9	
3. A grandmother brings her 6-month-old granddaughter into the hospital. The grandmother complains that the child is vomiting and having a lot of loose stools for 3 days now. She says that she is not drinking anything anymore. The child looks weak but reacts when you call her name. She has sunken eyes and her mouth looks very dry. Her hands are warm but her pulse is weak and thready. Vital Signs: RR 36, HR 146, Temp 36,5	
4. A mother brings her 10-month-old child to the triage area. She gives the history that he has been coughing, vomiting and had a fever since yesterday. You observe that the child is active, with no signs of dehydration. He is breathing fast and has mild intercostal recession. RR, 52 HR 195, Temp 38.8	
5. A 3-year-old presents with coughing for the last 3 days. His mother says that you must please give him something for his cold so she can go back to work. At triage you find a child with a nose full of snot, he is breathing fast and has both intercostal and subcostal recession. Vital Signs: RR 59, HR 152, and Temp 38.	
6. A 6-week-old girl known with TB, currently on TB meds is brought in by her grandmother. She has been yellow for the last	

(0/10) or obstructing her sight and she can still open and close her eye. Vital Signs: RR 48, HR 145, temp 36.2
6. An 8-year-old girl walks into the Emergency department with her mother, complaining of constipation for the last 4 days. She is not vomiting and doesn't have any pain (0/10). At triage you don't find anything significant. Vital Signs: RR26, HR 76, temp 36.9
7. A 3-month-old boy presents with a lump on his forehead, his aunt says he fell on his head yesterday. According to the aunt his older brother pushed him off a chair. At triage you find an active child. You see a small lump on his forehead but you are not sure what it is. The child has minimal pain (2/10). Vital Signs: RR 42, HR 166, Temp 37.1
8. A 2 yrs. old girl, brought in for a follow up visit. She presented last month and was treated for Severe acute malnutrition. Her mother says she is doing much better. She says she is eating much better now and has more energy than before. On examination she looks better than when you last saw her, but is still very small. She has no oedema present. Vital Signs: RR 18, HR 127, temp 37
9. A granny brings in her grandchild complaining of constipation. She informs you that he hasn't passed stool for 2 days. At triage you find a 1-year-old child. He wakes up when you start taking his vital signs. He appears well and you assess that he has no pain currently (0/10) Vital Signs: RR 29, HR 149, Temp 37,8
10. An active child is brought in by his mother; she complains that he has a rash. It has been there since birth and is itchy. He is now 14months old. There are no other problems. The skin is not broken with any sores present. It is dry and looks shiny. Vital Signs: RR, 32 HR, 122 and Temp 36.6
11. A mother complains that her 3 year old has been coughing and having a fever since yesterday. He is eating well and has no other problems. At triage you find a chubby 3 years old, playing on his mothers lap. No signs of respiratory distress seen at triage. Vital Signs: RR 20, HR 123, Temp 35,2
12. A 5 yrs. old patient known with diabetes presents to the EC, he father says he is having flu. The child's nose is running and he has had a cough for the last day. At Triage you find a thin 5 yrs. old with a runny nose. No respiratory distress seen. His glucose is within normal range. Vital Signs: RR 20 HR 80 Temp 36
13. A 18-month-old boy is brought in by his mother, she complains that he has been vomiting since yesterday and not eating as he normally does. There has been no blood in the vomit and no diarrhoea. He is still drinking fluids. At triage you find an active and playful child, there are no signs of dehydration that you can see. Vital Signs RR, 30 HR, 129 and Temp 36.
14. A 9-month-old girl is brought in by her mother, she reports that the girl cries whenever she passes stool, there is no reported blood or diarrhoea present. At triage you find a playful child who looks well. She has no sunken eyes and her skin pinch is normal. Vital Signs: RR36, HR, 131 and Temp 36,2
15. A 9 years old comes in complaining of itchy eyes for the last week. She says it started after she went to a flower show. At triage you find a healthy looking girl, with slightly red eyes. Vital Signs: RR 20, HR 100, Temp 36.5
16. A mother brings in her 1-year-old child who has diarrhoea and vomiting for the last day. He is happily drinking a cool drink when you see him in the triage area. At triage you find no signs of dehydration, Vital Signs: RR 36, HR 133, Temp 36,7

when you play with him. He starts to cry when you come close but has no pain now. Vital Signs: RR 36, HR 132, Temp 38,9
9. A mother runs into the emergency department with her 3-year-old child and says that he was fitting at home. His whole body was shaking and she was worried that he was going to die. This was about 2 hours ago, transport was a problem and they only arrived now. At triage you find a reactive child who is playing with the BP cuff. He is now alert and has no other signs that you can see. Vital Signs: RR 36 HR 130, Temp 37,7
10. An 11-year-old girl comes in and complains that she stood on a piece of glass. She cannot walk properly because of the pain. You assess it to be moderate (5/10) She has a bandage around her foot and its full of blood. It is slowly bleeding through the bandage. She keeps telling you how painful it is and how she can't dance now that she hurt her foot. Vital Signs: RR 16, HR 54, Temp 36.9
11. A tired frustrated mother comes to the Emergency department and complains that her baby won't stop crying. She has tried everything but nothing helps. This is her first child and is 2months old. On examination the baby is crying and even after soothing, continues to cry. It's difficult for you to assess anything with all the crying. Her vital signs are RR 46, HR 162 and temp 36,7
12. A mother brings her 2 year old in with a fever and cough for the last week. The mother tells you how the child got wet in the rain the week before and that's why she is sick now. At triage you find a generally healthy looking child who is playing with her mothers bag. You don't see any signs of respiratory distress. Vital Signs: RR 32, HR 132, Temp 39.2

Routine = 16	Needs to be seen within 240mins (4hours)
1. A mother brings a well looking 9-month-old child to the triage area; she says the babies' eyes are very red. She reports that the child hasn't been eating well since yesterday. The mother tried to wash the eyes with water but it just made it worse. The child currently has no pain (0/10). At Triage you find a child with red eyes, no other signs seen. Vital Signs: RR 32, HR 160, temp 37.1	
2. A 9-year-old girl vomits on your shoe as you walk into the waiting room. Her mother says she has been vomiting since last night. The patient walks into your triage area and looks tired. She is irritable and moody. You find no sunken eyes or dry mucous membranes and a normal skin pinch. She does not have pain currently. Vital Signs: RR 26, HR 126, Temp 37,3	
3. A distressed mother brings in a 3-year-old child and says she is very worried, he has just swallowed 50 ml of calamine/(anti itch) solution. He swallowed this about 4 hours ago but they couldn't get here sooner because of transport issues. At triage you find a normal child with no signs of distress. The child has no pain now (0/10). Vital Signs: RR 26, HR 124, Temp 37.3	
4. A mother brings in her 9-month-old baby and complains that she has been having diarrhoea for the last 2 days. She reports no vomiting and no blood seen in the diarrhoea. On examination you find a healthy looking child with slightly dry mucus membranes, but no sunken eyes and a normal skin pinch. Vital Signs: RR 36, HR 166, temp 37,3	
5. A mother brings in her 6-month-old child. The mother is very worried about a lump on her left eyelid. Otherwise she is drinking well and has just started eating solids. At triage you find she has a small, soft lump on her left eyelid. It is not painful	

few days. At triage you find she is alert but looks small for her age. She is breathing fast with mild intercostal recession, she has no airway noises. She is pink and her hands are warm. Vital Signs: RR 52, HR 153 and Temp 36.3
7. A 3-month-old presents with diarrhoea and vomiting for the last 2 days. She is not tolerating any feeds according to her mother. At Triage you see a weak and tired looking child, she has dry mucous membranes but a normal skin pinch. She has cool hands and her breathing is fast with intercostal recession. Vital Signs: RR42, HR 205, Temp 37,6
8. An 8-year-old boy limps towards you in the triage area while holding the right side of his abdomen. He says he has severe pain for the last 2 days. He vomited once this morning but is having no diarrhoea. At triage you find he is in a lot of pain, but alert and talking to you. You touch the right side of his stomach and he shouts in pain. He looks sweaty and tired. Vital Signs: RR 37, HR 149, Temp 40.7

Urgent = 12	Needs to be seen within 60mins
1. A 3-year-old presents with abdominal pain and vomiting. Her mother says she hasn't been eating or drinking anything since yesterday. At triage you find she has a dry mouth and is crying when you examine her but with no tears. She has no other significant signs. Vital Signs: RR 20 HR 112, Temp 36.8	
2. A 6-year-old boy is brought in by his mother. He complains of a headache for the last two days. He says he cannot concentrate in school because of this. He hasn't vomited and his vision is fine. At triage you find a well looking boy who is holding his head in his hands. He has moderate pain (6/10). Vital Signs: RR 16, HR 98, Temp 36	
3. A 2-month-old child known with TB presents with vomiting. His mother says, he is still drinking but has been vomiting for one day. He has not been able to take any medication. At triage you find a slow skin pinch >2 s but not other signs of dehydration, child is reactive but looks wasted. Vital signs: RR28, HR 162, Temp 37.	
4. A 3-year-old girl comes in with a 'tight chest' and no history of previous episodes. She is coughing and her mother says that the child is short of breath for the last 2 days. They are very poor and at night it is very cold where they stay. The child looks uncomfortable but alert, she has no chest in drawing but you can hear her wheezing. Vital Signs: RR 28, HR 135 and a temp 36.1	
5. A 3-year-old presents to your triage area. The history is that she was playing on the monkey bars/ jungle gym and fell off. The mother is worried about a broken bone. At triage you find she is keeping her hand very still and close to her body. She cried when you try to touch it. The arm looks slightly abnormal but no breaks in the skin are seen. She is in moderate pain when she keeps her arm still (5/10). Vital Signs: RR 20, HR 94, temp 35.4	
6. A mother complains that her child has been yellow for the last 3 weeks now. She says the previous doctor told her to keep him in the sun but it is not working. The child is 1 month old. At triage you find a yellow looking child, he is active and awake but cries when you start taking his vital signs. Vital Signs: RR 44, HR 133, Temp 36,7	
7. A mother complains that her 3-year-old son has been having a tight chest and difficulty breathing since yesterday. According to the mother he is a known 'asthmatic' On examination he is sitting up and tries to grab your cell phone. You lift up his shirt and see mild intercostal recessions with no other signs are seen. Vital signs: RR 49, HR 155, Temp 37.6	
8. A mother says her 3-year-old child has been complaining of abdominal cramps for the last 2 days. He has also been coughing for the last 2 weeks and is sweating at night. At triage you find a small for age child. He looks wasted; he is alert but very shy	

Reliability and validity of the South African Triage Scale in low resource settings

ORIGINALITY REPORT

17%

SIMILARITY INDEX

PRIMARY SOURCES

1	emssa.org.za Internet	602 words — 3%
2	fieldresearch.msf.org Internet	557 words — 3%
3	www.samj.org.za Internet	306 words — 1%
4	www.traumasa.co.za Internet	266 words — 1%
5	www.triagesa.co.za Internet	198 words — 1%
6	doaj.org Internet	164 words — 1%
7	Mohammed Dalwai, Pola Valles, Michele Twomey, Yvonne Nzomukunda et al. "Is the South African Triage Scale valid for use in Afghanistan, Haiti and Sierra Leone?", BMJ Global Health, 2017 Crossref	125 words — 1%
8	Twomey, MichÃ¨le, Lee A. Wallis, Mary Lou Thompson, and Jonathan E. Myers. "The South African Triage Scale (adult version) provides reliable acuity ratings", International Emergency Nursing, 2012. Crossref	97 words — < 1%
9	eprints.gla.ac.uk Internet	