Validation of an emergency triage scale for obstetrics and gynaecology: a prospective study

N Veit-Rubin,[a, b, c] P Brossard,[c, d] A Gayet-Ageron,[e] C-Y Montandon,[c, d] J Simon,[f] O Irion,[c, d] OT Rutschmann,[g] B Martinez de Tejada[c, d]

[a] Department of Gynaecology and Obstetrics, Lausanne University Hospital and Faculty of Medicine, Lausanne, Switzerland [b] Faculty of Medicine, University of Geneva, [c] Department of Obstetrics and Gynaecology, Medical University Vienna, Vienna, Austria [d] Department of Gynaecology and Obstetrics, Geneva University Hospitals, Geneva, Switzerland [e] CRC & Division of Clinical Epidemiology, Department of Health and Community Medicine, Geneva University Hospitals, Geneva, Switzerland [f] Nursing Department, Geneva University Hospitals, Geneva, Switzerland [g] Department of Primary Care, Community and Emergency Medicine, Division of Emergency Medicine, Geneva University Hospitals, Geneva, Switzerland
Correspondence: N Veit-Rubin, Avenue Pierre Decker 2, 1011 Lausanne, Switzerland. Email nikolaus.veitrubin@gmail.com

Objective To evaluate the reliability of a four-level triage scale for obstetrics and gynaecology emergencies and to explore the factors associated with an optimal triage.

Design Thirty clinical vignettes presenting the most frequent indications for obstetrics and gynaecology emergency consultations were evaluated twice using a computerised simulator.

Setting The study was performed at the emergency unit of obstetrics and gynaecology at the Geneva University Hospitals.

Sample The vignettes were submitted to nurses and midwives.

Methods We assessed inter- and intra-rater reliability and agreement using a two-way mixed-effects intra-class correlation (ICC). We also performed a generalised linear mixed model to evaluate factors associated triage correctness.

Main outcome measures Triage acuity.

Results We obtained a total of 1191 evaluations. Inter-rater reliability was good (ICC 0.748; 95% CI 0.633–0.858) and intra-rater reliability was almost perfect (ICC 0.812; 95% CI 0.726–0.889). We observed a wide variability: the mean number of questions varied from 6.9 to 18.9 across individuals and from 8.4 to 16.9 across vignettes. Triage acuity was underestimated in 12.4% of cases and overestimated in 9.3%. Undertriage occurred less frequently for gynaecology compared with obstetric vignettes [odds ratio (OR) 0.45; 95% CI 0.23–0.91; $P = 0.035$] and decreased with the number of questions asked (OR 0.94; 95% CI 0.88–0.99; $P = 0.047$). Certification in obstetrics and gynaecology emergencies was an independent factor for the avoidance of undertriage (OR 0.35; 95% CI 0.17–0.70; $P = 0.003$).

Conclusion The four-level triage scale is a valid and reliable tool for the integrated emergency management of obstetrics and gynaecology patients.

Keywords Reliability, triage, validity.

Tweetable abstract The Swiss Emergency Triage Scale is a valid and reliable tool for obstetrics and gynaecology emergency triage.

Introduction

Triage is the preliminary clinical assessment process that sorts patients before full diagnosis and treatment and has become crucial in times of overcrowded emergency units and resource constraints.[1] To deliver timely, efficient and safe high-quality care, it is important to select patients according to the severity of their condition.[2] Numerous triage instruments have been developed in several countries, such as in Australia [Australasian Triage Scale (ATS)], the United Kingdom (Manchester Triage System [MTS]), Canada [Canadian Triage and Acuity Scale (CTAS)] and the USA [Emergency Severity Index (ESI)].[3–6] Obstetrics and gynaecology patients represent specific risk groups, particularly for pregnancy-associated conditions. Physiological changes related to pregnancy, such as hypotension, tachypnoea or tachycardia, make most existing triage scales unsuitable for this population. In addition, some of the most dangerous complications during pregnancy, such as pre-eclampsia, are not well known by emergency teams. This has led to the development of specific tools, such as the Maternal Fetal Triage Index (MFTI) and

the Maternal Early Warning System (MEWS) in the USA and the Obstetric Triage Acuity Scale (OTAS) in Canada.[7–9] However, these tools either lack reliability, do not include gynaecological conditions, or fail to link to existing general triage instruments. The Swiss Emergency Triage Scale (SETS) is a reliable symptom-based four-level scale used in Europe for adult emergencies.[10]

In this study, we aimed to assess the use of the same SETS in the field of obstetrics and gynaecology emergency by means of an interactive triage simulator presenting a range of 30 real-life clinical vignettes. We first intended to evaluate the inter- and intra-rater reliability of the rating obtained in a group of obstetrics and gynaecology professionals then to assess the factors associated with optimal triage, over- and undertriage. Because, each vignette was completed following a step-by-step procedure, we secondly aimed to assess the factors associated with the number of steps used for answering each clinical vignette.

Methods

We conducted a prospective study at the maternity unit of the Geneva University Hospitals in Geneva, Switzerland, a tertiary referral centre for gynaecology and obstetrics within a 1200-bed urban teaching hospital. The unit is the largest in Switzerland with over 4000 deliveries per year, as well as 1300 emergency visits per month, 700 for gynaecology and 600 for obstetrics. All patients presenting to the obstetrics and gynaecology emergency ward are first evaluated by emergency nurses or midwives. The study consisted of the assessment of 30 clinical scenarios using a computerised triage simulator (see Supporting information, Appendix S1). All triage nurses and midwives in the emergency ward during the study period were invited to participate. The study was designed as a two-phase process, comprising one test and one retest phase. All participants were trained to use the SETS; some participants had also obtained certification in obstetrics and gynaecology emergencies. The local ethics committee of the Geneva University Hospitals approved the study protocol.

Study background
The SETS is available in French, German and Italian and currently used in adult emergency departments in Belgium, France and Switzerland.[10] To support the triage process, the measurement of vital signs and interpretation of the results are fully standardised. Similar to the MTS[11] and the CTAS,[5] the SETS incorporates timeline objectives as follows.

1 Level 1 (immediately life-threatening situation): assessment and treatment must be immediate with the patient installed in a resuscitation room, labour ward, or in an operating room for caesarean section.

2 Level 2 (potentially life-threatening situation): assessment and treatment must start within 20 minutes. The patient's condition may progress to a life-threatening situation; assessment must be rapid with the patient installed in an adequate intervention room (e.g. labour ward).

3 Level 3 (stable situation): assessment and treatment within 120 minutes; time is not considered as a critical factor. The nurse or midwife will regularly re-evaluate the patient's clinical condition in the waiting room if the patient cannot be installed immediately.

4 Level 4 (non-urgent situation): the patient is usually oriented towards outpatient clinics.

In 2008, a multidisciplinary team composed of gynaecologists, obstetricians, emergency physicians, nurses and midwives expanded the SETS to include the specificities of obstetrics and gynaecology, enabling its use as a triage tool in the emergency units of both subspecialties. These specificities included a grid defining emergency levels during and outside pregnancy to take into account physiological changes occurring during pregnancy (see Supporting information, Figure S1). In 2011, a unified emergency unit for obstetrics and gynaecology was created and the SETS was implemented as a permanent operational element. A multidisciplinary frame network was put in place to provide continuous training and supervision for its users and we started an internal evaluation process for its validation in this context. At the same time, we developed an internal 40-hour training programme that led to certification in obstetrics and gynaecology emergencies for nurses and midwives.

Triage simulator
The study used interactive software that simulated the triage process to be as close as possible to real-life conditions. In brief, a triage simulator integrated 30 standardised clinical vignettes representative of 15 common gynaecology and 15 common obstetric emergency conditions. Scenarios were developed by the multidisciplinary team based on real-life events that had previously occurred in the obstetrics and gynaecology emergency unit (see Supporting information, Appendix S1). They included situations across all triage levels. For each vignette, the participant had to choose the level of triage according to the SETS, based on the initial information and additional data provided. Participants were requested to evaluate the 30 vignettes twice within 6 months. Each vignette had an expert-based attributed level.

Outcomes and variables
Primary outcomes were reliability of the triage assessed by an intra-class correlation coefficient (ICC) and the level of triage correctness for each vignette. Correctness was defined

for each vignette if participants had attributed the same level of triage as the expert. ICC was interpreted as weighted κ coefficients following the Landis and Koch scale.[12] Among the 30 vignettes, five were level 1, 11 were level 2, 13 were level 3, and one was level 4, which is representative for local epidemiology. For each vignette, the level chosen by the participant was compared with the pre-established level attributed by the multidisciplinary expert panel. Over- and undertriage were defined as all overestimation and underestimation of the emergency level by the participant compared with the expert-attributed level. The vignette was defined as correctly classified when rated at the same level as the expert panel. Secondary outcomes were the number of questions needed per vignette to select a level of triage and the time spent to complete the triage process. These data were obtained from the logs of the computer simulator.

For each participant, we obtained information on age, gender, professional category (midwife versus nurse), completion or not of the certificate in obstetrics and gynaecology emergencies, time from nurse or midwife graduation in years, experience in obstetrics and gynaecology in months, and time spent at the obstetrics and gynaecology emergency ward in months.

Statistical analysis

Because we were constrained by the number of participants and the number of clinical vignettes, we estimated that the expected variance for an expected ICC between 0.70 and 0.80 would be between 0.002 and 0.0032.[13]

Continuous variables were presented by their median (interquartile range, IQR) for the number of questions asked; categorical variables were presented by their frequency and relative proportions. We compared the number of questions asked by vignette between test and re-test phases using a Wilcoxon signed-rank test. For the overall comparison of the number of questions asked between the two phases, we used a generalised linear mixed model and two non-nested random effects (one on the vignette and the other on the rater).

To measure the agreement in the triage assessment among participants, we calculated an ICC with its 95% CI using a two-way mixed-effects model to take into account that the 30 vignettes were rated by the same set of 22 independent raters.

The second aim was to assess the factors explaining the correctness of triage. To take into account the correlation in the answers, we performed a generalised linear mixed model using a logit link function and two non-nested random effects (one on the vignette and the other on the rater) to assess the correctness of each triage. We pre-specified a list of factors that could explain the correctness of the triage: the vignette specialty (obstetrics, gynaecology or

gynaecology–obstetrics); professional category (nurse versus midwife); time spent in obstetrics and gynaecology (<24 months, 24–48 months and ≥48 months); certification in obstetrics and gynaecology emergencies; and the number of questions needed to complete the clinical vignette. For each variable, we obtained an adjusted odds ratio (OR) with its 95% CI and the random-effect estimates of the vignette and the rater as SD.

Second, we assessed the factors associated with undertriage after exclusion of all vignettes presenting a level 4 ($n = 1$ vignette and 22 observations). Similarly, we assessed the factors related to overtriage after exclusion of all vignettes presenting a level 1 ($n = 10$ vignettes and 220 observations). For the two outcomes (undertriage then overtriage), we used again a generalised linear mixed model with a logit link function and two non-nested random effects (one on the vignette and the other on the rater). We adjusted for the same variables as in the first model assessing the correctness of triage.

All analyses were performed using STATA version intercooled 14 (STATA Corp, College Station, TX, USA). Statistical significance was defined as $P < 0.05$ (two-sided).

Results

We obtained a total of 1191 ratings (99.3%), 652 in the test phase and 539 in the retest phase. Only nine vignette evaluations were missing. Twenty-two trained triage professionals (78.6% of eligible personnel; six dropped out early because of sick leave) comprising nine nurses and 13 midwives completed the first phase of the study in summer and autumn 2014. Eight nurses and ten midwives ($n = 18$) also completed the second phase in spring 2015. Seventeen participants had obtained the certificate in obstetrics and gynaecology emergencies before the study (77.3%). All participants had >5 years of experience as health professionals and a mean number of 18.8 months of experience with the triage process (Table 1). Inter-rater reliability was 0.748 (95% CI 0633–0.858) of ICC at the test phase and intra-rater reliability was 0.812 (95% CI 0.726–0.889) of ICC at the re-test phase.

Evaluation of the triage process

We observed a wide variability in the median (IQR) number of questions asked per vignette, ranging from 8.0 (7–9) (#22, a case of a Bartholin's abscess) to 15.5 (13–21) (#12, a case of hypotension at 20 weeks of gestation) in the test phase (Figure 1A), and from 7.5 (7–11) (#22) to 17.5 (11–21) (#12) in the retest phase. The median (IQR) number of questions asked varied also across individuals from 7.0 (4–9) to 19.0 (15–22) in the test phase (Figure 1B). The observed variability had no significant impact on the correctness of triage results.

Table 1. Characteristics of study participants

Health profession, n (%)	
Nurse	9 (40.9)
Midwife	13 (59.1)
Triage certificate obtained, n (%)	
No	5 (22.7)
Yes	17 (77.3)
Years of health professional experience, median (IQR)	16 (8–18)
Years of health professional experience, n (%)	
<6 years	4 (18.2)
6–12 years	4 (18.2)
≥12 years	14 (63.6)
Months of experience in gynaecology and obstetrics, median (IQR)	42 (18–120)
Months of experience in gynaecology and obstetrics, n (%)	
<24 months	7 (31.8)
24–48 months	4 (18.2)
≥48 months	11 (50.0)
Months of experience with the triage process, median (IQR)	15 (3–36)
Months of experience with the triage process, n (%)	
<1 year	10 (45.4)
>1 year	12 (54.6)

There was no significant difference between the test and retest phases in the total number of questions asked by the 18 participants who completed both phases, except for four vignettes (#10, a case of the onset of term labour; #14, a case of vaginal mycosis; #18, a case of first-trimester hyperemesis; and #29, a case of menometrorrhaghia with a fibroid uterus). Overall, the retest phase resulted in a significantly lower median number of questions asked compared with the test phase [12.0 (9–15) versus 11.0 (8–15), respectively; $P < 0.001$].
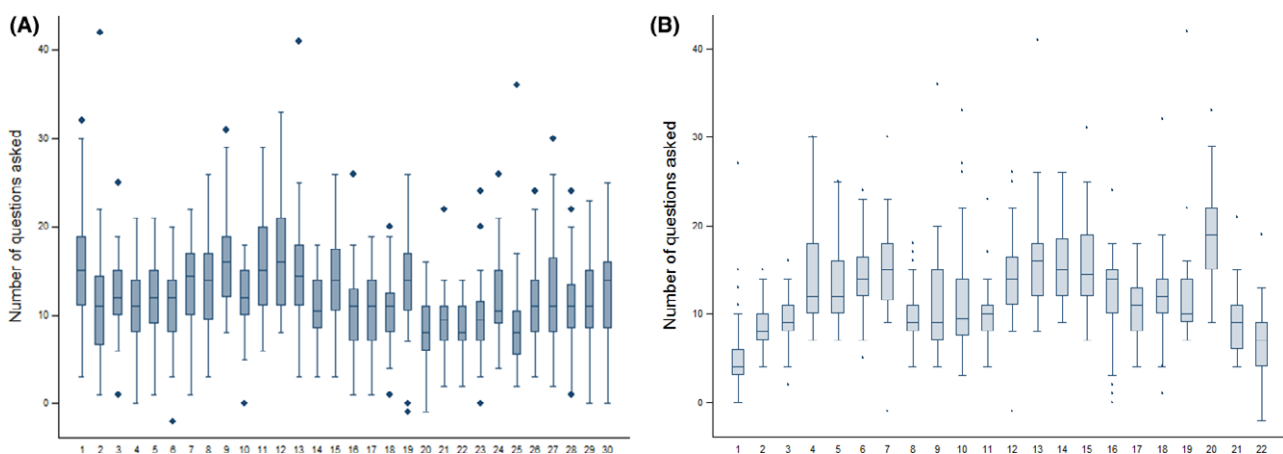
Triage performance evaluation in the test phase

Perfect agreement between the expected and the observed triage decisions was found in 78.4% of situations ($n = 511$). Overestimation of the emergency level was observed in 63 ratings (9.6%). It was especially frequent in five situations where more than 30% of participants allocated a more severe severity score. Four of the frequently overestimated cases were pregnancy-related and one dealt with sexual assault. Underestimation of the emergency level was slightly more frequent and was observed in 78 (12.0%) assessments. It often occurred in three situations where more than 30% of participants allocated a less severe severity score. Two cases were pregnancy-related and one was abdominal pain. We observed a different distribution of assessment errors depending on the vignette's emergency level ($P < 0.001$). Levels 1 and 4 were more often allocated correctly than levels 2 and 3, the latter being more prone to assessment errors (Table 2).

In univariate analysis, no factor was significantly associated with the correct allocation of the emergency level during the triage process or overtriage (Table 3). Undertriage occurred less frequently for gynaecology than obstetric vignettes (OR 0.45; 95% CI 0.23–0.91; $P = 0.035$) and underestimation of the emergency level decreased with the number of questions asked (OR 0.94; 95% CI 0.88–0.99; $P = 0.047$). The multivariate model showed that having the certificate in obstetrics and gynaecology emergencies was an independent factor for the avoidance of undertriage (OR 0.35; 95% CI 0.17–0.70; $P = 0.003$) after adjustment for the vignette specialty, profession, months of experience in gynaecology and obstetrics, and the number of questions asked.

Discussion

Main findings
We were able to confirm the reliability of the SETS in an obstetrics and gynaecology setting and to explore the

Figure 1. Variability in the number of questions asked per vignette (A), and in the number of questions asked per participant (B).

Table 2. Triage correctness, undertriage and overtriage by level of emergency presented in the vignette

Triage level	Correct triage, n (%)	Undertriage, n (%)	Overtriage, n (%)	Total
1	89 (81.6)	20 (18.3)	–	109
2	185 (77.1)	28 (11.7)	27 (11.2)	240
3	215 (76.5)	30 (10.7)	36 (12.8)	281
4	22 (100)	–	0 (0)	22
Total	511 (78.4)	78 (12.0)	63 (9.6)	652

Table 3. Factors associated with the triage correctness (univariate analyses)

	OR	95% CI
Triage level of the vignette		
1	1.00	–
2	1.54	0.55–4.31
3	1.16	0.54–2.51
4	0.19	0.03–1.19
Triage level group of the vignette		
Level 1–2	1.00	–
Level 3–4	1.51	0.41–5.56
Specialty of the vignette		
Gynaecolgy	1.00	–
Obstetrics	0.80	0.37–1.76
Gynaecology–obsterics	0.65	0.19–2.24
Health profession		
Nurse	1.00	–
Midwife	0.92	0.54–1.57
Years of health professional experience		
<6 years	1.00	–
6–12 years	1.27	0.54–2.99
≥12 years	0.75	0.38–1.46
Months of experience in gynaecology and obstetrics		
<24 months	1.00	–
24–48 months	0.66	0.31–1.41
≥48 months	0.84	0.46–1.51
Years of experience with the triage process		
<1 year	1.00	–
>1 year	1.23	0.73–2.08
Triage certificate obtained		
No	1.00	–
Yes	1.45	0.80–2.63
Number of questions asked		
<10 questions	1.00	–
10–15 questions	1.03	0.62–1.71
≥15 questions	1.18	0.66–2.13

performance of nurses and midwives in the triage process. The excellent intra-rater reliability was superior to the results obtained in the studies evaluating the ESI, CTAS and OTAS. The good inter-rater reliability was equal[9,14–16] and superior compared with the MFTI.[17]

Moreover, we observed an excellent agreement between the reference standard and evaluators in 78.4% of all assessments, which was superior to the results obtained with the ATS.[18] About 10% of assessments led to overestimation and occurred frequently in four vignettes (#4, 9, 17, 24). Underestimation occurred in 12% of ratings, which was significantly lower than in the original SETS,[10] the CTAS[19] and the ATS.[18]

Strengths and limitations

Compared with other available triage instruments in obstetrics and gynaecology, the SETS has the advantage of being an integrated scale based on an extensively tested system in a general emergency department.[10] Although using similar elements, early warning systems, such as the MEWS, constitute more of a 'track and trigger' process for inpatient and outpatient settings and use periodic observation of selected vital signs 'track' combined with predetermined criteria 'trigger' to summon experienced help.[20] Hence, they are not specifically designed for use in emergency triage. In addition, the MFTI, MEWS and OTAS only consider obstetric specificities.[7–9] Gynaecological conditions should also be adequately represented in a triage tool.

There are several limitations to our study. Our results come from a single-centre study, which might be a limitation in the study's generalisability. However, we used a standardised approach to implement the triage simulator in an obstetrics and gynaecology unit.[10] Moreover, participants represented well the characteristics of this specific population of healthcare workers and we did not suspect any selection bias in the study population, which favours the study's generalisability. The evaluation of vignettes may not represent a real-time approach where implementation is limited by overcrowded emergency units and the impossibility of simultaneous triage by a second independent evaluator. There may be a selection bias due to exclusion of the sickest, sequential evaluation and, finally, the impossibility of performing test–retest processes. Although users considered the simulator's narrative content to be close to reality, it lacked visual features, a detail previously proven to be beneficial,[18] and a revised version is currently under development. Ideally, the evaluation of a triage scale should incorporate the assessment of clinical outcome depending on triage level assignment and to address the potential adverse outcomes that would have resulted from the misassignment.

Interpretation

As three of overestimated vignettes were pregnancy-related, this could be due to an overevaluation of the danger to the unborn child. Vignette #4 described a term pregnancy driving the evaluator towards orientation to the labour suite

when easily available. Vignette #9 listed symptoms of pre-eclampsia at week 25 in a migrainous patient. Given this potentially life-threatening pathology for both the mother and child, a higher level of attention can be assumed.[21] Vignette #17 was a case of severe haemorrhage brought by ambulance at the end of the first trimester. Finally, vignette #24 illustrated a case of sexual assault, which represents a delicate situation from medical, psychological and juridical points of view.[22]

The potential impact of an underestimation would be that severe cases might be unsafely diverted to lower acuity with potentially significant negative impact on the clinical outcome. In our study, four vignettes were frequently rated with a lower level than expected (#6, 18, 21, 25). Vignette #6 was a case of haemorrhage at term with known placenta praevia, which would require level 1 care as potentially life-threatening for the newborn. We considered this underestimation as a worrying finding. Vignette #18 described a case of first-trimester hyperemesis. Associated hypotension should have prompted classifying the vignette at a higher level, although hyperemesis is usually considered a nonsevere disease. Vignette #21 was a case of a common genital infection. The high pain score should have generated the allocation to level 3 instead of 4. Finally, vignette #25 illustrated the situation of an elderly woman presenting with pelvic organ prolapse. A level 3 management for the patient's comfort would be preferable over redirection towards an outpatient setting (level 4).

The right balance between overestimation and underestimation remains a subject of debate. Many authors consider underestimation more problematic and frequently argue that patient safety is of greater value than economic efficiency.[1,2,23] Nevertheless, inefficient care due to overestimation may consume resources subsequently unavailable for concomitant severe emergencies. Hospital administrators usually favour instruments that minimise the inappropriate use of resources.[23,24] In the multivariate analysis, we confirmed that specific triage training and experience tend to reduce underestimation and overestimation, respectively.

We observed also an important variability in the number of questions asked by participants. This is possibly a reflection of the difficulties encountered by triage professionals to obtain relevant information to stratify patients and deliver the most efficient care. Another potential factor explaining this variability could be the participants' professional background (nurses and midwives) often reflecting unequal professional philosophies and values; furthermore, the raters' professional experience varied significantly. A similar variability has been described also in previous studies using a simulator or written vignettes.[14,16,25] The participants of these studies all based their assessment on the same information for each case.[14,18] However, in real life,

the triage professional must actively seek information, which is adequately simulated with our tool.[10]

Conclusion

In summary, the four-level SETS is a valid and reliable emergency triage tool suitable for the continuous training of healthcare professionals to improve women's care and cost-control in emergency units. We believe that based on our practice and the findings of our study, the SETS represents a ready-to-use tool to be adopted in a clinical setting for emergency triage in obstetrics and gynaecology.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. English translation of the pregnancy-related items of the SETS triage sheet.

Appendix S1. Thirty scenarios included in the triage simulator. ∎

References

1 Gilboy N, Travers D, Wuerz R. Re-evaluating triage in the new millennium: a comprehensive look at the need for standardization and quality. *J Emerg Nursing* 1999;25:468–73.

2 Washington DL, Stevens CD, Shekelle PG, Baker DW, Fink A, Brook RH. Safely directing patients to appropriate levels of care: guideline-driven triage in the emergency service. *Ann Emerg Med* 2000;36:15–22.

3 Health AGDo. Emergency Triage Education Kit: Australian Government Department of Health; 2013 [www.health.gov.au/internet/main/publishing.nsf/Content/casemix-ED-triage+Review+Fact+Sheet+Documents]. Accessed 26 January 2016.

4 Group MT. Manchester Triage System [www.triagenet.net/en/]. Accessed 26 January 2016.

5 (CAEP) CAOEP. Canadian Triage and Acuity Scale (CTAS) [caep.ca/resources/ctas/implementation-guidelines]. Accessed 26 January 2016.

6 Gilboy NTP, Travers D, Rosenau AM. *Emergency Severity Index (ESI): A Triage Tool for Emergency Department Care, Version 4. Implementation Handbook*. Rockville: AHRQ Publication, 2005.

7 Ruhl C, Scheich B, Onokpise B, Bingham D. Content validity testing of the maternal fetal triage index. *J Obstet Gynecol Neonatal Nurs* 2015;44:701–9.

8 Mhyre JM, D'Oria R, Hameed AB, Lappen JR, Holley SL, Hunter SK, et al. The maternal early warning criteria: a proposal from the national partnership for maternal safety. *J Obstet Gynecol Neonatal Nurs* 2014;43:771–9.

9 Smithson DS, Twohey R, Rice T, Watts N, Fernandes CM, Gratton RJ. Implementing an obstetric triage acuity scale: interrater reliability and patient flow analysis. *Am J Obstet Gynecol* 2013;209:287–93.

10 Rutschmann OT, Kossovsky M, Geissbuhler A, Perneger TV, Vermeulen B, Simon J, et al. Interactive triage simulator revealed important variability in both process and outcome of emergency triage. *J Clin Epidemiol* 2006;59:615–21.

11 Mackway-Jones K. *Manchester Triage Group: Emergency Triage*. London: BMJ Publishing Group, 1997.

12 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.

13 Shoukri M. Sample size requirements for the design of reliability study: review and new results. *Stat Methods Med Res* 2004;13:251–71.

14 Tanabe P, Gimbel R, Yarnold PR, Kyriacou DN, Adams JG. Reliability and validity of scores on The Emergency Severity Index version 3. *Acad Emerg Med* 2004;11:59–65.

15 Mirhaghi A, Heydari A, Mazlom R, Hasanzadeh F. Reliability of the Emergency Severity Index: meta-analysis. *Sultan Qaboos Univ Med J* 2015;15:e71–7.

16 Fernandes CM, McLeod S, Krause J, Shah A, Jewell J, Smith B, et al. Reliability of the Canadian Triage and Acuity Scale: interrater and intrarater agreement from a community and an academic emergency department. *CJEM* 2013;15:227–32.

17 Ruhl C, Scheich B, Onokpise B, Bingham D. Interrater reliability testing of the Maternal Fetal Triage Index. *J Obstet Gynecol Neonatal Nurs* 2015;44:710–6.

18 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.

19 Beveridge R, Ducharme J, Janes L, Beaulieu S, Walter S. Reliability of the Canadian Emergency Department Triage and Acuity Scale: interrater agreement. *Ann Emerg Med* 1999;34:155–9.

20 Swanton RD, Al-Rawi S, Wee MY. A national survey of obstetric early warning systems in the United Kingdom. *Int J Obstet Anesth* 2009;18:253–7.

21 Sperling JD, Dahlke JD, Huber WJ, Sibai BM. The role of headache in the classification and management of hypertensive disorders in pregnancy. *Obstet Gynecol* 2015;126:297–302.

22 Sande MK, Broderick KB, Moreira ME, Bender B, Hopkins E, Buchanan JA. Sexual assault training in emergency medicine residencies: a survey of program directors. *Western J Emerg Med* 2013;14:461–6.

23 Cremonesi P, di Bella E, Montefiori M, Persico L. The robustness and effectiveness of the triage system at times of overcrowding and the extra costs due to inappropriate use of emergency departments. *Appl Health Econom Health Policy* 2015;13:507–14.

24 Simonet D. Cost reduction strategies for emergency services: insurance role, practice changes and patients accountability. *Health Care Anal* 2009;17:1–19.

25 Manos D, Petrie DA, Beveridge RC, Walter S, Ducharme J. Inter-observer agreement using the Canadian Emergency Department Triage and Acuity Scale. *CJEM* 2002;4:16–22.