

Reliability of Dutch Obstetric Telephone Triage

Bernice Engeltjes^{1,2}Ageeth Rosman²Loes CM Bertens³Eveline Wouters⁴Doug Cronie²Fedde Scheele^{1,5}

¹Athena Institute for Transdisciplinary Research, Faculty of Science, VU University, Amsterdam, the Netherlands;

²Department of Healthcare Studies, Rotterdam University of Applied Sciences, Rotterdam, the Netherlands;

³Department of Obstetrics and Gynecology, Erasmus MC, University Medical Centre Rotterdam, Rotterdam, the Netherlands; ⁴Department of Tranzo, Tilburg School of Social and Behavioral Sciences, Tilburg University, Tilburg, the Netherlands; ⁵Department of Healthcare Education, OLVG Teaching Hospital, Amsterdam, the Netherlands

Background: Safety and efficiency of emergency care can be optimized with a triage system which uses urgency to prioritize care. The Dutch Obstetric Telephone Triage System (DOTTS) was developed to provide a basis for assessing urgency of unplanned obstetric care requests by telephone. Reliability and validity are important components in evaluating such (obstetric) triage systems.

Objective: To determine the reliability of Dutch Obstetric Telephone Triage, by calculating the inter-rater and intra-rater reliability.

Methods: To evaluate the urgency levels of DOTTS by testing inter-rater and intra-rater reliability, 90 vignettes of possible requests were developed. Nineteen participants, from hospitals where DOTTS had been implemented, rated in two rounds a set of ten vignettes. The five urgency levels and five presenting symptoms had an equal spread and had to be entered in accordance with DOTTS per vignette. Urgency levels were dichotomized into high urgency and intermediate urgency. Inter-rater reliability was rated as degree of agreement between two different participants with the same vignette. Intra-rater reliability was rated as agreement by the same participants at different moments in time. The degree of inter-rater and intra-rater reliability was tested using weighted Cohen's Kappa and ICC.

Results: The agreement of urgency level between participants in accordance with predefined urgency level per vignette was 90.5% (95% CI 87.5–93.6) [335 of 370]. Agreement of urgency level between participants was 88.5% (95% CI 84.9–93.0) [177 of 200] and 84.9% (95% CI 78.3–91.4) after re-rating [101 of 119]. Inter-rater reliability of DOTTS expressed as Cohen's Kappa was 0.77 and as ICC 0.87; intra-rater reliability of DOTTS expressed as Cohen's Kappa was 0.70 and as ICC 0.82.

Conclusion: Inter-rater and intra-rater reliability of DOTTS showed substantial correlation, and is comparable to other studies. Therefore, DOTTS is considered reliable.

Keywords: obstetrics, triage system, inter-observer agreement, intra-observer agreement, undertriage and overtriage

Introduction

A triage system that prioritizes care according to urgency is known to have a favorable effect on safety and efficiency of emergency care.^{1–4} Triage systems contain background information about presenting symptoms and urgency levels, which aim to indicate the maximum acceptable medical waiting time. Triage is applied during a telephone and/or physical contact when registering for an emergency department. Triage systems such as the Manchester Triage System (MTS), the Emergency Severity Index (ESI) and the Canadian Triage and Acuity Scale (CTAS) are commonly used for triage in emergency departments worldwide.^{5–8}

However, general triage systems are not sufficiently specific for use in obstetrics. Therefore, in recent years physical (face-to-face) triage systems have been

Correspondence: Bernice Engeltjes
Department of Healthcare Studies,
Rotterdam University of Applied Sciences,
Rotterdam, the Netherlands
Tel +31 6 41804881
Email engeltjesbernice@gmail.com

developed specifically for obstetrics.^{6,9–15} The Obstetric Triage Acuity Scale (OTAS) from Canada,^{6,10} Swiss Emergency Triage Scale (SETS),¹¹ Birmingham Symptom specific Obstetric Triage System (BSOTS) from United Kingdom¹² and Maternal Fetal Triage Index (MFTI) from the United States of America^{13,14} are well-established obstetric physical triage systems. More recently, the Iranian Obstetric Triage Index (IOTI) was developed and published (2020).¹⁵ The inter-rater reliability of the existing physical obstetric triage systems is moderate to good (ranging between Kappa 0.69–0.86 and intraclass correlation (ICC) 0.75–0.96). Intra-rater reliability showed an ICC of 0.81 for SETS¹¹ and a Kappa of 0.65 for OTAS (2016).⁶ Intra-rater correlations are unknown for BSOTS, MFTI and IOTI.^{9,12,13,15} Due to the heterogeneity of methods, results and quality of the studies, it is difficult to compare these studies.⁹

All of the obstetric systems discussed have been developed for physical (face-to-face) triage. In practice, in western society, it is usual for most women to first make a telephone call asking whether it is necessary to have a consultation at the emergency department.^{16,17} Therefore, in most instances, the first triage is performed by telephone and occurs before the pregnant woman is clinically rated. In order to apply the correct level of priority, accurate rating of the urgency is crucial. The Dutch Obstetric Telephone Triage System (DOTTS) aims to provide a uniform and practical triage system, and was developed through a multi-phase multi-center study in consultation with all stakeholders.¹⁸ DOTTS is an evidence-based triage system, which uses presenting symptoms to classify the level of urgency. Recently published research into validity of DOTTS showed an acceptable diagnostic validity with room for improvement. The overall sensitivity was 76%, and DOTTS compared to a reference standard had an agreement of 53%, and overtriage in 30% and undertriage in 16% of the cases.¹⁹ DOTTS was introduced in 2015 and is currently used in 26% of all Dutch hospitals (n=20/78).^{18,20} The purpose of this study is to determine the reliability of DOTTS.

Materials and Methods

This study aims to evaluate the reliability of DOTTS by testing inter-rater reliability (IRR) and intra-rater reliability (ITR) using vignettes.

DOTTS is comparable to other triage systems. It consists of five urgency levels: 1) resuscitation and life threatening, 2) emergency, 3) urgent, 4) non-urgent and 5) self-care. It uses

five presenting symptoms: 1) fluid loss, 2) vaginal bleeding, 3) abdominal pain, 4) non-somatic symptoms and 5) other physical symptoms. In this study, we focused on the reliability of assigning the correct urgency levels.

Participants and Development Vignettes

From hospitals where DOTTS was implemented, triage staff (obstetrical nurses or doctor's assistants) were asked to participate. Each participant had completed practical training in the use of DOTTS at the time of implementation in their hospital and had a minimum work experience of 3 months with DOTTS.

In order to further guarantee a basic knowledge level of DOTTS, completion of an interactive e-learning developed for this study was mandatory. In the e-learning information was given about DOTTS, after which this knowledge was quizzed. In case of incorrect answers, new questions were asked, until the participant demonstrated sufficient knowledge of DOTTS. A certificate was given after completion of the e-learning. Only certified participants received vignettes.

Ninety vignettes were developed using real-life clinical situations. The vignettes described cases with one of the five urgency levels and the five presenting symptoms as used by DOTTS. The urgency levels and presenting symptoms were equally distributed ([Supplementary Material, Table A](#)). An expert panel, comprising seven midwives with expertise in DOTTS and obstetric emergency skills training, reviewed all vignettes for accuracy, credibility, and completeness. The vignettes were modelled to standardize the order of the information and incorporated into an online questionnaire (Qualtrics®).

These 90 vignettes were divided into nine sets. Each participant received a set of ten vignettes per round. In each round, each vignette was judged by a minimum of two participants. Each participant was blinded by the ratings of others. The minimum number of participants was set at 18 participants. This number was determined based on feasibility for participants. The expected time needed to complete both rounds was two hours.

Urgency levels and presenting symptoms had to be entered in accordance with DOTTS. To avoid recall bias, the contents of the sets in the second round differed from the first round, with three vignettes changed, and an adjusted order of the other seven vignettes.

For reliability, a distinction is made between inter-rater reliability (IRR) and intra-rater reliability (ITR). IRR of a triage system is the degree of agreement between different professionals, whereas ITR is agreement of the same

professionals between different moments in time.⁹ To determine IRR, the first round was sent between June and August 2020. After at least two months (September–October 2020) the vignettes were present for the second round to determine ITR.

Data Collection and Statistical Analysis

Collected participant characteristics were as follows: age, professional category (nurse or doctor's assistant), hospital, obstetric experience (years) and number of hours and patients per week in the triage ward. Analyses of participants' characteristics were presented as numbers (N) with percentages (%) or median with interquartile ranges (IQR) and ranges. All analyses were performed using SPSS, version 25.

Based on the information presented in the vignettes, participants were asked to assign an urgency level based on presenting symptoms. Agreement with DOTTS was analyzed by comparison of the urgency level. Agreed triage was defined as triage by the participant in accordance with the predefined level of urgency in DOTTS. Disagreement in triage was considered undertriage when the participant indicated a lower level of urgency and overtriage when a participant assigned a higher urgency level.

For statistical analyses, urgency levels were dichotomized into high urgency (U1, U2) and intermediate urgency (U3, U4 and U5). This resulted in 40 vignettes in the high urgency category and 50 vignettes in the intermediate urgency category ([Supplementary Material, Table A](#)).

Inter-rater reliability (IRR) and intra-rater reliability (ITR) were rated by using a weighted Cohen's Kappa to account for agreement in classifications based on chance alone, for multiple raters and multiple categories. Also, two-way-mixed intraclass correlation coefficient (ICC) was calculated, to enable comparison of the reliability of DOTTS with other published triage systems. Interpretation of Cohen's Kappa was done according to the arbitrary scaling of Landis and Koch, with a kappa between 0.61 and 0.80 indicating substantial correlation, and the values 0.81–1.0 indicating near perfect correlation.^{6,9–13,21} Interpretation of ICC values was based on the scaling of Koo and Li, meaning good reliability (0.75–0.9) and moderate reliability (0.5–0.75).²²

Ethical Approval

The study was approved by the daily Boards of the Medical Research Ethics Committees United (MEC-U) and the Medical Ethics Committee of Leiden University

Medical Center (LUMC) Act (W.16.053 and P17.075/Pg/pg).

All participants provided digital informed consent to use the data for analyses. All data was anonymously processed. Participants were able to withdraw consent at any time, without any statement of reasons.

Results

Overall, 19 participants took part, 15 (79%) nurses and 4 (21%) doctor's assistants. One professional did not participate in round two. To enable inclusion of all vignettes in calculation of IRR, the set of vignettes of the dropped out professional was rated in round two by another professional. This made a total of 370 ratings of vignettes, for the IRR 200 ratings were available and for ITR 119 ([Figure 1](#)).

The participants had a median age of 53 years [IQR 44–55], and a median work experience in obstetrics of 20 years [IQR 8–33]. An overview of basic characteristics of participants, such as participation per hospital, working hours and experience with triage, is given in [Table 1](#). Distribution of urgency levels and presenting symptoms were approximately equally divided ([Table 1](#)).

In total, 370 ratings were made. The overall agreement of urgency category was 90.5% (n=335). Undertriage was present in 4.3% of cases (n=16), overtriage was 5.2% (n=19) ([Figure 1](#)).

In total, 200 ratings were available to calculate IRR. In total 88 high urgency vignettes and 112 intermediate urgency vignettes were rated ([Figure 1](#)). Overall, in 88.5% (n=177 of 200) the urgency categories were the same between two participants: IRR Kappa 0.77 95% CI 0.68–0.86 and ICC 0.87 (95% CI 0.83–0.90), respectively. The level of agreement between participants in high urgency and intermediate urgency category was similar: 90.8% (n=79 of 87) and 86.7% (n=98 of 113), respectively ([Table 2](#)).

One hundred and nineteen vignettes were rated twice by the same participants. Of these vignettes, 51 had a high urgency level and 68 an intermediate urgency level. The ITR was calculated on these 119 paired ratings ([Figure 1](#)). Overall, in 84.9% (n=101 of 119) of the urgency categories were rated the same in the first and second round: ITR Kappa 0.70 (95% CI 0.57–0.83) and ICC 0.82 (95% CI 0.74–0.88). In both rounds, the participants scored 90.1% (n=46 of 51) the same in the high urgency category. In the intermediate urgency category, this was 80.9% (n=55 of 68) ([Table 2](#)).

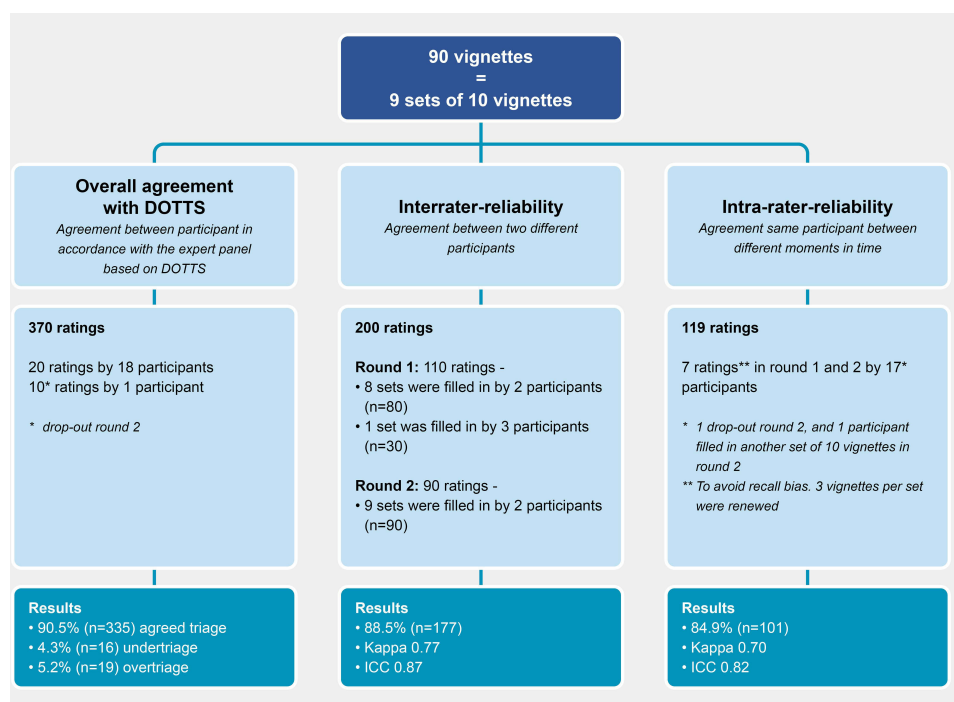


Figure 1 Schematic overview of participants, vignettes and results.

Discussion

Overall agreement of urgency category was 90.5% (n=335). Agreement between the different participants (IRR) in using DOTS was 88.5%, with weighted Kappa 0.77 and ICC 0.87. Agreement of the same participants between different moments in time (ITR), was 84.9%, with weighted Kappa 0.70 and ICC 0.82. Therefore, according to Landis and Koch's scale,²¹ our results demonstrate a substantial correlation and a good level of reliability according to Koo and Li.²² A triage system is only beneficial if the reliability has been demonstrated by research.⁹ These results confirm the internal consistency of DOTS, the use of both measurements indicates the systematic reliability.

The reliability achieved for the DOTS telephone triage system is comparable to that of physical (face-to-face) obstetric triage systems. In two studies in which reliability was reported, IRR of OTAS-2013 expressed as Kappa was 0.71 and that of SETS was expressed as ICC 0.75,^{10,11} this corresponds with the results of the reliability of DOTS. Research of the ITR of OTAS-2016 showed a weighted Kappa of 0.65, and of SETS an ICC of 0.81.^{6,11} In their recent review, Moudi et al⁹ showed that for obstetric triage systems, the quality of evidence is moderate to low, with only two systems (OTAS-2013 and SETS)

presenting psychometric properties. Compared to these two triage systems, DOTS shows similar results ([Supplementary Material - Table B](#)).

The increased volume of obstetric emergency care and the pursuit of high-quality interpretation and documentation of unplanned obstetric care consultations require improvement of current care processes.^{9,14,23} Nowadays, obstetric triage systems are being used more often in clinical practice.^{6,9-11,13,15,18} A telephone triage system adds to this development. In addition, the use of a valid and reliable telephone triage system contributes to the correct distribution of patients and resources. This is increasingly necessary due to the growing concentration of acute care in obstetrics in general and is particularly relevant during the current COVID-19 pandemic.²⁴

Currently, DOTS already has a digital application that supports clinical decision-making with algorithms suitable for use in every electronic patients' dossier. This is comparable to other triage systems that incorporate clinical decision support systems, to aid in the evaluation of patients' health conditions.² In future, DOTS algorithms may benefit from more supporting technologies such as automatically calling of an ambulance and adding home-measurements of vital parameters such as saturation, blood pressure and fetal assessment by cardiotocography (CTG).²⁵ Also, video observation and communication by

Table 1 Characteristics of Participants and Vignettes

Participants, n (%)	19 (100)
Age, years median [IQR] (Range)	53.0 [44–55] (31)
Work experience in obstetrics, years median [IQR] (Range)	20.0 [8–33] (37)
Professional category	
Obstetrical nurse, n (%)	15 (78.9)
Doctor's assistants, n (%)	4 (21.1)
Hospital	
Academic hospital, n (%)	4 (21.1)
Teaching hospital, n (%)	9 (47.4)
Non-teaching hospital, n (%)	6 (31.6)
Exposure (average) to triage per week	
≥ 16 hours, n (%)	7 (36.8)
9–15 hours, n (%)	6 (31.6)
≤ 8 hours, n (%)	6 (31.6)
Exposure (average) to patients per week	
20–49 consults, n (%)	6 (31.6)
10–19 consults, n (%)	9 (47.4)
0–9 consults, n (%)	4 (21.1)
Vignettes - Urgency levels, n (%)	90 (100)
High urgency, n (%)	40 (44.4)
Intermediate urgency, n (%)	50 (55.5)
Vignettes - Presenting symptoms, n (%)	90 (100)
Abdominal pain, n (%)	20 (22.2)
Anxious pregnant woman/non-somatic symptoms, n (%)	16 (17.8)
Other physical symptoms, n (%)	17 (18.9)
Vaginal bleeding, n (%)	20 (22.2)
Vaginal fluid loss, n (%)	17 (18.9)

healthcare professionals provide additional information such as assessment of the clinical status of the patient and/or the observation of vital signs such as the amount

of blood loss. Currently, this is not yet available in the telephone triage system, which means that the professionals need to make assumptions exclusively based on the patient's self-report.^{24–27} In future, such developments are likely to further improve the telephone triage systems and further increase reliability.

Strengths and Limitations

A strength of this study is that it mirrors the clinical situation as closely as possible. The vignettes were based on real clinical situations and were collected from hospitals where DOTS was used. In addition, they were assessed for accuracy by experts. Another strength of the study is the use of an e-learning prior to the start of the study. Participants' competency level was therefore ensured. In addition, the design of the questionnaire required the participants to complete answers to all questions, thus ensuring that completeness. Also, our results were generated from participants from a wide range of hospitals who actually use the system, which enhances generalizability.^{2–4,28–32}

A potential limitation of the study is that it was undertaken with written vignettes, as opposed to a spontaneous conversation between patient and triage staff member. Participants could not continue to ask questions if anything was unclear. Also, the study environment differed from the reality of the (often overcrowded) triage ward. Severity of complaints, patient characteristics and follow-up are various factors which influence the situation in real-life situations. In addition, due to the small sample size, no statement can be made about the outcomes per sort of hospital or work experience in obstetrics of the triagist. In this study, the triagists were found to have a wide range

Table 2 Inter-Rater and Intra-Rater Reliability Measures of DOTS

	Inter-Rater Reliability Different Participants with the Same Vignette	Intra-Rater Reliability Same Participant at Different Moment in Time
Agreed triage, Total % (95% CI) [n]	88.5 (95% CI 84.9–93.0) [177/200]	84.9 (95% CI 78.3–91.4) [101/119]
High urgency category	90.8 (95% CI 84.6–97.0) [79/87]	90.1 (95% CI 81.8–98.5) [46/51]
Intermediate urgency category	87.5 (95% CI 80.3–93.1) [98/113]	80.9 (95% CI 71.3–90.4) [55/68]
Weighted Kappa**	0.77 (95% CI 0.68–0.86)	0.70 (95% CI 0.57–0.83)
Intraclass correlation coefficient+	0.87 (95% CI 0.83–0.90)	0.82 (95% CI 0.74–0.88)

Notes: **Scale references by Landis and Koch:²¹ 0.61–0.80 = substantial correlation, 0.81–1.0 = near perfect correlation. +Scale reference by Koo and Li:²² 0.5–0.75 = moderate reliability and 0.75–0.9 good reliability.

of experience. Further research would be needed to establish any potential effect of this experience on reliability. Lastly, reading skills as opposed to listening skills of the participants may have influenced the results of this study.^{11,13,15,18,28,29,32}

Recommendations for Further Research

Triage is intended to indicate a correct level of urgency and to prioritize patients with high urgency. In this study, undertriage and overtriage were minimal, 4.3% and 5.2%, respectively. An obstetric triage system should help to reduce undertriage, because the potential consequence of undertriage could be irreversible health damage. Overtriage should also be avoided, as this can lead to work overload and inefficient use of resources. Moving forward, it is important to pay attention to all aspects of safety of triage in all hospital settings as well as to the patient experiences of such.

Conclusion

Inter-rater and intra-rater reliability of DOTTS showed substantial correlation, and is comparable to other studies. Therefore, DOTTS can be considered a reliable obstetric telephone triage system. This telephone triage tool gives priority to care based on urgency before physical examination, further increasing the quality and efficiency of obstetric care.

Abbreviations

BSOTS, Birmingham Symptom specific Obstetric Triage System; CTAS, Canadian Triage and Acuity Scale; DOTTS, Dutch Obstetric Telephone Triage System; ESI, Emergency Severity Index; IOTI, Iranian Obstetric Triage Index; ICC, Intraclass correlation coefficient; IRR, Inter-rater reliability; ITR, Intra-rater reliability; LUMC, Leiden University Medical Center; M, Mean; MEC-U, Medical Research Ethics Committees United; MFTI, Maternal Fetal Triage Index; MTS, Manchester Triage System; N, Numbers; OTAS, Obstetric Triage Acuity Scale; SD, standard deviation; SETS, Swiss Emergency Triage Scale; U, Urgency.

Ethical Approval

The study was approved by the Medical Research Ethics Committees United (MEC-U) and the Medical Ethics Committee of Leiden University Medical Center (LUMC, 2016) Act (W.16.053 and P17.075/PG/pg). The study was

submitted to and approved in 2016 by the daily Boards MEC-U and the Medical Ethics Committee of Leiden University Medical Center (LUMC, 2016). As a result of these reviews, the boards declared that the rules laid down in the Medical Research Involving Human Subjects Act (also known by its Dutch abbreviation WMO) do not apply to the study. No additional approval was requested for the current study since it is not based upon a clinical study or patient data. All participants gave digital informed consent to use the data for analyses. All data was anonymously processed. Participants were able to resign at any time, without any statement of reasons.

Acknowledgments

The authors wish to thank all participating obstetric professionals of the following hospitals: OLVG Amsterdam, Leiden University Medical Center Leiden, Elisabeth Tweesteden hospital Tilburg, Jeroen Bosch hospital 's Hertogenbosch, Tjongerschans hospital Heerenveen, Ikazia hospital Rotterdam, University Medical Center Maastricht, Antonius hospital Utrecht and IJsselland hospital Capelle aan de IJssel. The authors also wish to thank Dr Rudy Rijke, Drs Corlijn van Dijk for feedback and editing assistance.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation. And took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agreed to be accountable for all aspects of the work.

Disclosure

Drs Bernice Engeltjes reported grants from the Netherlands Organization for Scientific Research (NWO) (Award number 023.009.030) during the conduct of the study. The authors declare that they have no other potential conflicts of interest in this work.

References

1. van Ierland Y, van Veen M, Huibers L, Giesen P, Moll HA. Validity of telephone and physical triage in emergency care: the Netherlands Triage System. *Fam Pract*. 2011;28(3):334–341. doi:10.1093/fampra/cmz097
2. Lake R, Georgiou A, Li J, et al. The quality, safety and governance of telephone triage and advice services - an overview of evidence from systematic reviews. *BMC Health Serv Res*. 2017;17(1):614. doi:10.1186/s12913-017-2564-x

3. Huibers L, Keizer E, Giesen P, Grol R, Wensing M. Nurse telephone triage: good quality associated with appropriate decisions. *Fam Pract.* **2012**;29(5):547–552. doi:10.1093/fampra/cms005
4. Graversen DS, Pedersen AF, Carlsen AH, Bro F, Huibers L, Christensen MB. Quality of out-of-hours telephone triage by general practitioners and nurses: development and testing of the AQTT - an assessment tool measuring communication, patient safety and efficiency. *Scand J Prim Health Care.* **2019**;37(1):18–29. doi:10.1080/02813432.2019.1568712
5. Huibers L, Giesen P, Wensing M, Grol R. Out-of-hours care in western countries: assessment of different organizational models. *BMC Health Serv Res.* **2009**;9:105. doi:10.1186/1472-6963-9-105
6. Gratton RJ, Bazaracai N, Cameron I, et al. Acuity assessment in obstetrical triage. *J Obstet Gynaecol Can.* **2016**;38(2):125–133. doi:10.1016/j.jogc.2015.12.010
7. Zachariasse JM, Seiger N, Rood PP, et al. Validity of the Manchester Triage System in emergency care: a prospective observational study. *PLoS One.* **2017**;12(2):e0170811. doi:10.1371/journal.pone.0170811
8. Zachariasse JM, van der Hagen V, Seiger N, Mackway-Jones K, van Veen M, Moll HA. Performance of triage systems in emergency care: a systematic review and meta-analysis. *BMJ Open.* **2019**;9(5):e026471. doi:10.1136/bmjopen-2018-026471
9. Moudi A, Iravani M, Najafian M, Zareian A, Forouzan A, Mirghafourvand M. Obstetric triage systems: a systematic review of measurement properties (Clinimetric). *BMC Pregnancy Childbirth.* **2020**;20(1):275. doi:10.1186/s12884-020-02974-0
10. 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(4):287–293. doi:10.1016/j.jogc.2013.03.031
11. Veit-Rubin N, Brossard P, Gayet-Ageron A, et al. Validation of an emergency triage scale for obstetrics and gynaecology: a prospective study. *BJOG.* **2017**;124(12):1867–1873. doi:10.1111/1471-0528.14535
12. Kenyon S, Hewison A, Dann SA, et al. The design and implementation of an obstetric triage system for unscheduled pregnancy related attendances: a mixed methods evaluation. *BMC Pregnancy Childbirth.* **2017**;17(1):309. doi:10.1186/s12884-017-1503-5
13. Ruhl C, Scheich B, Onokpise B, Bingham D. Interrater reliability testing of the maternal fetal triage index. *J Obstet Gynecol Neonatal Nurs.* **2015**;44(6):710–716. doi:10.1111/1552-6909.12762
14. Ruhl C, Scheich B, Onokpise B, Bingham D. Content validity testing of the maternal fetal triage index. *J Obstet Gynecol Neonatal Nurs.* **2015**;44(6):701–709. doi:10.1111/1552-6909.12763
15. Moudi A, Iravani M, Najafian M, Zareian A, Forouzan A, Mirghafourvand M. The development and validation of an obstetric triage acuity index: a mixed-method study. *J Matern Fetal Neonatal Med.* **2020**;1–11. doi:10.1080/14767058.2020.1768239
16. Manning NA, Magann EF, Rhoads SJ, Ivey TL, Williams DJ. Role of telephone triage in obstetrics. *Obstet Gynecol Surv.* **2012**;67(12):810–816. doi:10.1097/OGX.0b013e318277dd9c
17. Bailey CM, Newton JM, Hall HG. Telephone triage in midwifery practice: a cross-sectional survey. *Int J Nurs Stud.* **2019**;91:110–118. doi:10.1016/j.ijnurstu.2018.11.009
18. Engeltjes B, Wouters E, Rijke R, Scheele F. Obstetric telephone triage. *Risk Manag Healthc Policy.* **2020**;13:2497–2506. doi:10.2147/RMHP.S277464
19. Engeltjes B, Van Dijk C, Rosman A, Rijke R, Scheele F, Wouters E. Validation of Dutch Obstetric Telephone Triage System: a prospective validation study. *Risk Manag Healthc Policy.* **2021**;14:1–9. doi:10.2147/RMHP.S306390
20. Engeltjes B, Walinga RM, Van Dijk CFA, et al. Successful implementation of a telephone obstetric triage system: the Dutch Triage Guide to Midwifery. *Dutch J Obstetr Gynecol.* **2019**;32:192–195.
21. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* **1977**;33(1):159–174. doi:10.2307/2529310
22. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med.* **2016**;15(2):155–163. doi:10.1016/j.jcm.2016.02.012
23. Poeran J, Borsboom GJ, de Graaf JP, et al. Does centralisation of acute obstetric care reduce intrapartum and first-week mortality? An empirical study of over 1 million births in the Netherlands. *Health Policy (New York).* **2014**;117(1):28–38. doi:10.1016/j.healthpol.2014.03.009
24. Madden N, Emeruwa UN, Friedman AM, et al. Telehealth uptake into prenatal care and provider attitudes during the COVID-19 pandemic in New York City: a quantitative and qualitative analysis. *Am J Perinatol.* **2020**;37(10):1005–1014. doi:10.1055/s-0040-1712939
25. van den Heuvel JFM, Ayubi S, Franx A, Bekker MN. Home-based monitoring and telemonitoring of complicated pregnancies: nationwide cross-sectional survey of current practice in the Netherlands. *JMIR Mhealth Uhealth.* **2020**;8(10):e18966. doi:10.2196/18966
26. DeNicola N, Grossman D, Marko K, et al. Telehealth interventions to improve obstetric and gynecologic health outcomes: a systematic review. *Obstet Gynecol.* **2020**;135(2):371–382. doi:10.1097/AOG.0000000000003646
27. Tate K, Antheunis ML, Kanters S, Nieboer TE, Gerritse MB. The effect of screen-to-screen versus face-to-face consultation on doctor-patient communication: an experimental study with simulated patients. *J Med Internet Res.* **2017**;19(12):e421. doi:10.2196/jmir.8033
28. Smits M, Hanssen S, Huibers L, Giesen P. Telephone triage in general practices: a written case scenario study in the Netherlands. *Scand J Prim Health Care.* **2016**;34(1):28–36. doi:10.3109/02813432.2016.1144431
29. Huibers L, Smits M, Renaud V, Giesen P, Wensing M. Safety of telephone triage in out-of-hours care: a systematic review. *Scand J Prim Health Care.* **2011**;29(4):198–209. doi:10.3109/02813432.2011.629150
30. Bunn F, Byrne G, Kendall S. Telephone consultation and triage: effects on health care use and patient satisfaction. *Cochrane Database Syst Rev.* **2004**;4:CD004180. doi:10.1002/14651858.CD004180.pub2
31. Siddiqui N, Greenfield D, Lawler A. Calling for confirmation, reassurance, and direction: investigating patient compliance after accessing a telephone triage advice service. *Int J Health Plann Manage.* **2020**;35(3):735–745. doi:10.1002/hpm.2934
32. Tam HL, Chung SF, Lou CK. A review of triage accuracy and future direction. *BMC Emerg Med.* **2018**;18(1):58. doi:10.1186/s12873-018-0215-0

Risk Management and Healthcare Policy

Dovepress

Publish your work in this journal

Risk Management and Healthcare Policy is an international, peer-reviewed, open access journal focusing on all aspects of public health, policy, and preventative measures to promote good health and improve morbidity and mortality in the population. The journal welcomes submitted papers covering original research, basic science, clinical & epidemiological studies, reviews and evaluations,

guidelines, expert opinion and commentary, case reports and extended reports. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/risk-management-and-healthcare-policy-journal>