ORIGINAL ARTICLE

VALIDITY AND RELIABILITY OF THE MANCHESTER SCALE USED IN THE ORTHOPEDIC EMERGENCY DEPARTMENT

VALIDADE E CONFIABILIDADE DA ESCALA DE MANCHESTER APLICADA NO PRONTO-SOCORRO ORTOPÉDICO

Fernando Brandão Andrade-Silva[1], Renan Lyuji Takemura[1], Renato Tavares Bellato[1], Marcos de Camargo Leonhardt[1], Kodi Edson Kojima[1], Jorge dos Santos Silva[1]

1. Universidade de São Paulo, Faculdade de Medicina, Hospital das Clínicas (IOT-HCFMUSP), Instituto de Ortopedia e Traumatologia, São Paulo, SP, Brazil.

ABSTRACT

Objectives: To describe the clinical utility of the Manchester triage scale adapted for orthopedic emergency departments and to evaluate its validity in identifying patients with the need for hospital care and its reliability when reproduced by different professionals. Methods: Five triage flowcharts were developed based on the Manchester scale for the following orthopedic disorders: traumatic injuries, joint pain, vertebral pain, postoperative disorders, and musculoskeletal infections. A series of patients triaged by two orthopedists was analyzed to assess the concordance between the evaluators (reliability) and the validity of the Manchester scale as predictive of severity. Results: The reliability analysis included 231 patients, with an inter-observer agreement of 84% (Kappa = 0.77, p <0.001). The validity analysis included 138 patients. The risk category had a strong association with the need for hospital care in patients with trauma (OR = 6.57, p = 0.001) and was not significant for non-traumatic disorders (OR = 2.42; p = 0.208). The overall sensitivity and specificity were 64% and 76%, respectively. Conclusion: The evaluated system presented high reliability. Its validity was adequate, with good sensitivity for identifying patients requiring hospital care among those with traumatic lesions. However, the sensitivity was low for patients with non-traumatic lesions. **Level of Evidence III, Retrospective Study.**

Keywords: Triage. Classification. Risk. Orthopedics.

RESUMO

Objetivos: Descrever a utilização clínica da escala de triagem de Manchester adaptada ao pronto-socorro ortopédico e avaliar sua validade para detectar pacientes com necessidade de cuidados hospitalares e sua confiabilidade quando reproduzida por diferentes profissionais. Métodos: Cinco fluxogramas de triagem foram desenvolvidos com base na escala de Manchester: lesões traumáticas, dor articular, dor vertebral, transtorno pós-operatório e infecção musculoesquelética. Uma série de pacientes triados por dois ortopedistas foi analisada para avaliação da concordância entre os avaliadores (confiabilidade) e sua validade como preditivo de gravidade. Resultados: A avaliação da confiabilidade incluiu 231 pacientes, com concordância inter--observadores de 84% (Kappa = 0,77; p <0,001). A análise da validade incluiu 138 pacientes. A classificação de risco apresentou forte associação com a necessidade de cuidados hospitalares em pacientes com trauma (OR = 6,57; p = 0,001), não sendo significativa nos transtornos não-traumáticos (OR = 2,42; p = 0,208). A sensibilidade geral foi de 64% e a especificidade de 76%. Conclusão: O sistema avaliado apresentou alta confiabilidade. Sua validade foi adequada, com boa sensibilidade para detectar pacientes com necessidade de procedimentos hospitalares em lesões traumáticas, enquanto a sensibilidade foi baixa em pacientes com lesões não-traumáticas. **Nível de Evidência III, Estudo Retrospectivo.**

Descritores: Triagem. Classificação. Risco. Ortopedia.

INTRODUCTION

Triage systems aim to identify patients with severe conditions who need urgent management, prioritizing their care in emergency departments. Hospital accreditation bodies have required the implementation of clinically validated triage systems in emergency departments. Long waiting times have been associated with greater morbidity in patients who need urgent attention.[1]
Most triage systems used in practice are based on a five-level risk model whose application ideally takes place within a short interval from the arrival of the patient at the emergency department.[2] The Australasian Triage Scale, Canadian Triage and Acuity Scale, Emergency Severity Index, and Manchester Triage System (MTS) are the most commonly cited in the literature, with studies reporting adequate overall validity and reliability.[2-7] The Manchester system is one of the most widely used in practice[8] and includes 52 flow charts for various clinical conditions, including some orthopedic disorders, such as back pain, local infections, and trauma.[9] Some studies have reported specific

systems for orthopedic triage in outpatient services,[10,11] but not in emergency departments.

This study aimed to develop a triage scale based on the Manchester system adapted to orthopedic management and to evaluate the scale for inter-examiner agreement (reliability) and validity in detection of higher severity clinical conditions.

MATERIALS AND METHODS

The Manchester system divides orthopedic disorders into five groups: traumatic injuries, joint pain, vertebral pain, postoperative disorder, and musculoskeletal infection. A flow chart was developed for each group based on severity criteria, including pain level assessed with the visual analog scale (VAS), infectious signs, presence of deformity or bleeding, and time of appearance of symptoms, among others (Attachments 1 to 5). The flow charts were adapted from corresponding charts in the Manchester scale.[9] The urgency categories (assigned risk) and the maximum waiting time were standardized according to the MTS (immediate: 0 minutes; very urgent: 10 minutes; urgent: 60 minutes; routine: 120 minutes; non-urgent: 240 minutes). This study was approved by the institutional ethics committee (protocol number 1,127; registration number: 13,611).

Reliability was evaluated prospectively for this study. For seven days, all patients treated in our emergency department were triaged by two physicians: initially by an orthopedic physician of the emergency room staff (attending physician), and later by a third-year resident physician. The evaluations were carried out in separate rooms to avoid evaluation bias. The evaluation of the attending physician was used to define the patient flow after triage. Examiner agreement was calculated using the kappa coefficient. This analysis included all patients aged ≥18 years who were evaluated by both physicians, regardless of follow-up after triage.

Validity was evaluated retrospectively by analyzing the data from the medical records of patients triaged in the period described. A list of clinical outcomes considered relevant within each group was defined, and medical records were evaluated to detect patients who had one or more outcomes after triage. The association between the risk assigned to the patient and the presence of a clinical outcome was analyzed using logistic regression to evaluate the validity of the scale. Patients with incomplete medical records whose clinical course could not be identified were excluded from this analysis. Among the clinical outcomes evaluated, the need for hospitalization and the need for surgical treatment were included in all groups of injuries. The outcomes specific to each group included:

Traumatic injuries: radiographic diagnosis of fracture or dislocation, with the need for reduction or immobilization;

Joint pain: definitive diagnosis of septic or aseptic arthritis, with the need for joint puncture or intravenous analgesia;

Vertebral pain: clinical or radiological diagnosis of fracture, tumor, infection, or radiculopathy, with the need for intravenous analgesia or hospitalization;

Postoperative disorder: clinical or laboratory diagnosis of postoperative infection, implant failure, or release, requiring intravenous analgesia or hospitalization

Musculoskeletal infection: laboratory and/or radiological diagnosis of septic arthritis or osteomyelitis, with the need for intravenous antibiotic therapy.

In the statistical analysis, the triage categories were evaluated separately and grouped as follows: "high priority" (very urgent and urgent categories) and "low priority" (routine and non-urgent categories). The association between the category and the presence of a clinical outcome was calculated using univariate logistic regression analysis and the chi-square test or Fisher's test according to the number of observations in each group.

The sensitivity and specificity of the scale were calculated, in addition to the undertriage rate (number of patients with negative outcomes who were classified as "high priority" divided by the total number of patients classified as "high priority") and overtriage rate (number of patients with positive outcomes classified as "low priority" divided by the total number of patients). The sample size was determined according to the number of patients observed during the study period. Stata Software 13.0 (Stata Corp, College Station, TX, USA) was used for statistical analysis.

RESULTS

The reliability analysis included 231 patients with the following distribution: traumatic injuries: 106 patients (46%); joint pain: 58 (25%); vertebral pain: 49 (21%); postoperative disorder: 14 (6%); and musculoskeletal infection: 4 (2%). The validity analysis included 138 patients. The mean age was 42.2 years, with the same number of male and female patients.

The analysis of risk distribution showed that patients with traumatic injuries were almost equally classified as "high priority" (very urgent = 25%, urgent = 27%) and "low priority" (routine = 43%; non-urgent = 5%). In contrast, patients with joint pain and vertebral pain were mostly classified as "low priority" (93% and 84%, respectively). Patients with postoperative disorders and musculoskeletal infections were almost equally distributed, based on a limited number of cases (Table 1). The inter-examiner agreement varied among the groups: traumatic injuries: 84%; joint pain: 88%; vertebral pain: 76%; postoperative disorder: 93%; and infection: 75%. The overall reliability was 84%, with a kappa coefficient of 0.77 (p < 0.001).

The validity analysis of the overall sample showed a statistically significant association between the assigned risk and the presence of a clinical outcome, both in the individual and grouped categories (odds ratio [OR] = 5.71; confidence interval [CI] 95% = 2.60 to 12.54, p < 0.001). Similarly, the analysis of traumatic injuries showed a statistically significant association between the assigned risk and the presence of a clinical outcome, both in the individual (OR = 0.29; 95% CI = 0.15 to 0.59; p = 0.001) and grouped categories (OR = 6.57; 95% CI = 2.16 to 20.03, p = 0, 001). The analysis of non-traumatic injuries did not show a statistically significant association between the assigned risk and the presence of an outcome, both in individual and grouped categories (OR = 2.42; 95% CI = 0.61 to 9.58; p = 0.208) (Table 2).

Table 1. Urgency levels assigned by the attending and resident physicians to patients evaluated in the study according to the pathology group.

	Very Urgent	Urgent	Routine	Non-urgent
Trauma (n = 106)				
Attending	26 (25%)	29 (27%)	46 (43%)	5 (5%)
Resident	28 (26%)	29 (27%)	42 (40%)	7 (7%)
Joint pain (n = 58)				
Attending	1 (2%)	3 (5%)	24 (41%)	30 (52%)
Resident	0	5 (9%)	23 (40%)	30 (52%)
Vertebral pain (n = 49)				
Attending	0	8 (16%)	21 (43%)	20 (41%)
Resident	0	8 (8%)	29 (60%)	16 (33%)
Postoperative period (n = 14)				
Attending	2 (14%)	4 (29%)	7 (50%)	1 (7%)
Resident	2 (14%)	4 (29%)	6 (43%)	2 (14%)
Infection (n = 4)				
Attending	0	2 (50%)	1 (25%)	1 (25%)
Resident	1 (25%)	2 (50%)	1 (25%)	0
TOTAL (n = 231)				
Attending	29 (13%)	46 (20%)	99 (43%)	57 (25%)
Resident	31 (13%)	44 (19%)	101 (44%)	55 (24%)

Table 2. Patients with clinical outcomes divided according to the pathology group and the risk assigned.

	Very Urgent	Urgent	Routine	Non-urgent	P-value
Trauma (n = 29)	11 (38%)	12 (41%)	6 (21%)	0	0.002
Non-traumatic (n = 13)					
Joint pain (n = 6)	0	0	4 (67%)	2 (33%)	0.779
Vertebral pain (n = 5)	0	2 (40%)	1 (20%)	2 (40%)	0.185
Postoperative (n = 0)	0	0	0	0	-
Infection (n = 2)	0	2 (100%)	0	0	-
TOTAL (n = 42)	11 (26%)	16 (38%)	11 (26%)	4 (10%)	< 0.001

The sensitivity of the scale in the overall sample was 64% (95% CI = 48% to 78%) and specificity was 76% (95% CI = 66% to 84%). For traumatic injuries, sensitivity was 79% (95% CI = 60% to 92%) and specificity was 63% (95% CI = 46% to 78%), while in non-traumatic injuries, sensitivity was 31% (95% CI = 9% to 61%) and specificity was 85% (95% CI = 73% to 93%). The overtriage rate was 38% in patients with traumatic injuries and 69% in patients with non-traumatic injuries, while the undertriage rate was 9% in patients with traumatic injuries and 13% in patients with non-traumatic disorders.

DISCUSSION

The scale presented in this study is an alternative to orthopedic emergency services that evaluate pre-care risk. With the data from this study, we confirmed the high reliability of the adapted Manchester scale, demonstrating good reproducibility when used by different professionals. We also demonstrated adequate validity for identification of patients with severe traumatic injuries. The scale had low sensitivity for non-traumatic injuries, and its validity could not be demonstrated in this group of pathologies, which can be partially explained by the method chosen for validation. Despite this shortcoming, the scale proved to be easy to use in clinical practice, based on objective criteria for the classification of patients and guiding of flow in the emergency room.

The reliability, or reproducibility, of a classification scale and any diagnostic test represents its ability to produce consistent results when used independently.[12] A risk scale with good average validity but low reliability produces dispersed results, with little clinical use.[12] The scale described in this study showed a good agreement rate (84%; kappa = 0.77) according to previously established criteria.[13] Van der Wulp et al. reported consistent reliability (kappa = 0.62) using the MTS in a general population, but did not find an association between the experience of nurses and system reliability.[5] Van Veen et al. observed high reliability (kappa = 0.83) when using the Manchester system for the triage of children using written case scenarios.[14]

The validity of a diagnostic test should be ideally evaluated by comparing it to a gold standard. However, in the absence of this test, substitute markers are used as references.[12] In our study, we used markers indicating two types of outcomes: the definitive diagnosis of orthopedic pathology and the need for hospital care or procedures. Although they are clinically relevant outcomes, their severity was not considered, which may have influenced the validity analysis of the scale. In the groups of non-traumatic disorders, patients who developed a clinical outcome (N = 13) were mostly underdiagnosed (N = 9) and classified as "low priority," which is related to the low sensitivity of the scale in these groups of pathologies. The presence of mild clinical signs in patients requiring procedures may explain this incongruence (e.g., patients with joint effusion and low pain level who underwent joint puncture). Parenti et al. conducted a systematic review on the validity and reliability of the MTS scale, showing low safety of the method due to a high undertriage rate and low sensitivity to identify higher levels of urgency.[7] Our findings on non-traumatic injuries corroborate these data.
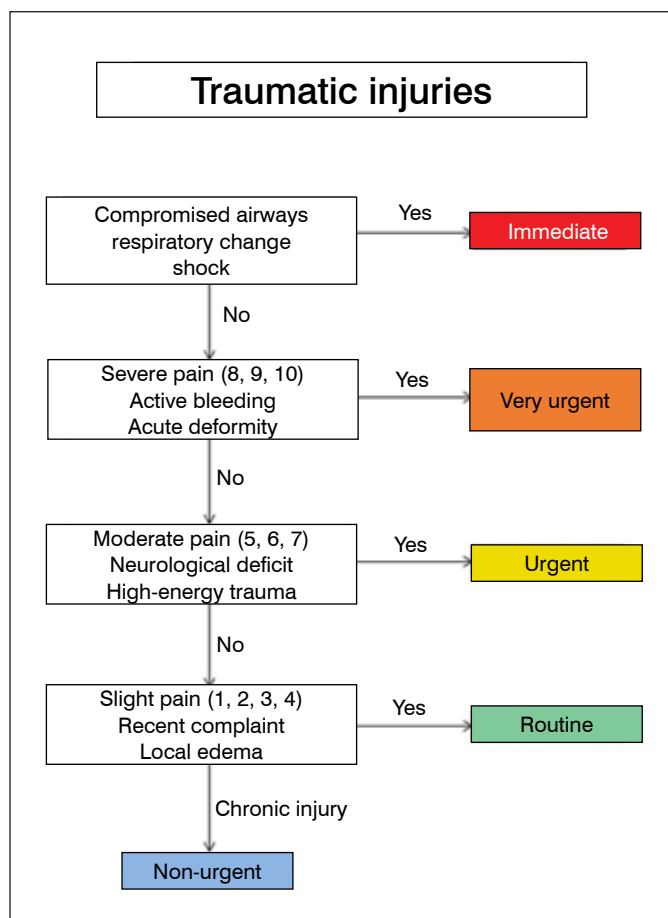
The undertriage rate of a risk scale evaluates its ability to detect patients with greater urgency correctly and is inversely proportional to their sensitivity.[15] A high undertriage rate means that patients who need urgent care are not detected, and is related to low sensitivity of the method. Ideally, undertriage rates should be less than 5%.[15] However, most studies validating risk scales show values 5%.[7] In the present study, the undertriage rate in traumatic injuries was 9% and sensitivity was 79%, which are considered adequate for clinical use. Roukema et al. reported an undertriage rate of 15% and a sensitivity of 63% when using the Manchester scale. In the assessment of emergency conditions in pediatric patients[16] with non-traumatic injuries, our data showed low sensitivity (31%) for the identification of patients with clinical outcomes and an undertriage rate of 13%. As already mentioned, this shortcoming may be related to the method used in validation, but we cannot rule out failure of the scale to identify patients with greater severity, as shown in previous studies that evaluated the Manchester system in different scenarios.[7] The overtriage rate of a diagnostic test represents the percentage of patients who do not have emergency conditions but are mistakenly classified as priority. A high overtriage rate overloads the emergency department, and may impair the care of patients with a true emergency. Ideally, overtriage rates should vary between 25 and 50%.[15] In our study, the overtriage rate for traumatic injuries was 38% and specificity was 63%; these are within the appropriate range for clinical use. In non-traumatic injuries, the overtriage rate was 69% and the specificity was 85%. Storm-Versloot et al. reported a 29% overtriage rate and great variability in the specificity of MTS, showing lower values at lower levels of urgency (level 1 = 100%; 2 = 95%; 3 = 66%; and 4 = 2%).[17] Other studies showed significant variability in sensitivity and specificity when the MTS scale was used in different scenarios according to the assigned urgency levels.[5,18-20]

This study had some limitations. Retrospective validity evaluation involved a significant loss of patient follow-up and incomplete medical records data, resulting in a limited number of cases. This may have influenced the outcome of scale validation. Specifically, the "postoperative disorder" and "musculoskeletal infection" groups had few cases, preventing adequate data analysis. As noted, comparison with a gold standard would be more appropriate for validation, and the use of clinical markers may have generated inconsistencies. In most studies, triage was performed by nurses. In our study, triage was performed by orthopedic physicians, which may limit the generalization of the data.

In conclusion, this study presented a triage system based on the Manchester system adapted to orthopedic scenarios. The system showed high reliability when used by different professionals and was efficient in detecting patients with traumatic injuries who needed hospital care. The system had low sensitivity for detecting patients with non-traumatic injuries who required hospital care. These findings do not preclude the use of the Manchester system in the orthopedic emergency department, but greater attention should be given to avoid undervaluing the clinical presentation of patients with non-traumatic injuries.
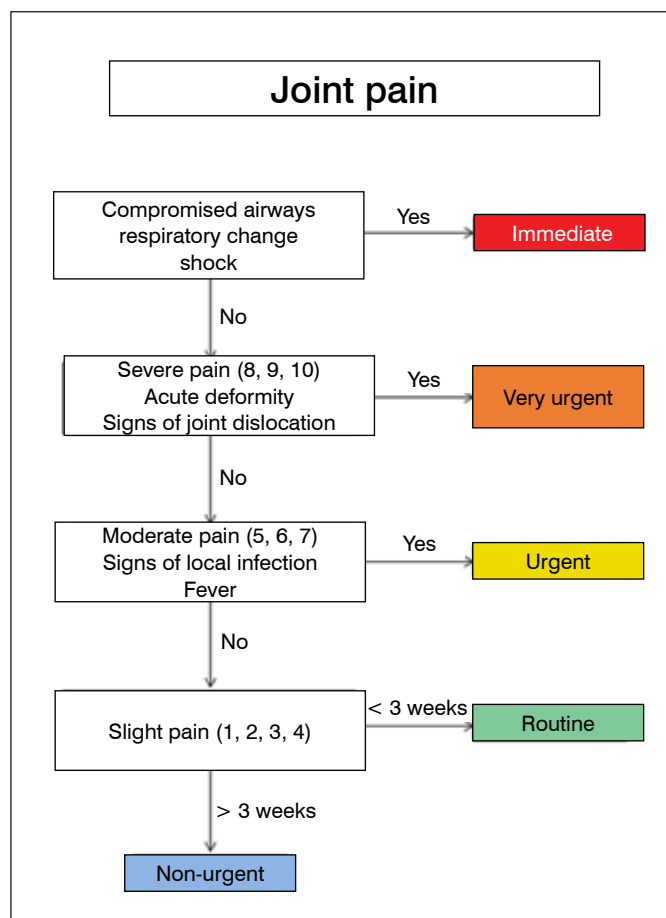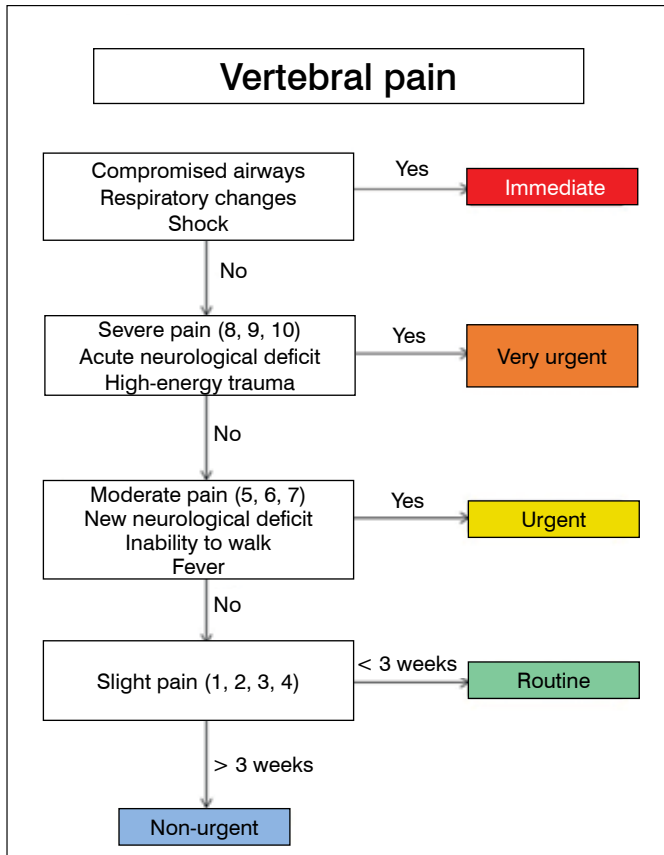
REFERENCES

1. Schuetz P, Hausfater P, Amin D, Haubitz S, Fässler L, Grolimund E, et al. Optimizing triage and hospitalization in adult general medical emergency patients: the triage project. BMC Emerg Med. 2013;13:12.
2. Christ M, Grossmann F, Winter D, Bingisser R, Platz E. Modern triage in the emergency department. Dtsch Arztebl Int. 2010;107(50):892–8.
3. 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(1):59–65.
4. Alquraini M, Awad E, Hijazi R. Reliability of Canadian Emergency Department Triage and Acuity Scale (CTAS) in Saudi Arabia. Int J Emerg Med. 2015;8(1):80.
5. van der Wulp I, van Baar ME, Schrijvers AJP. 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(7):431–4.
6. Grossmann FF, Zumbrunn T, Frauchiger A, Delport K, Bingisser R, Nickel CH. At risk of undertriage? Testing the performance and accuracy of the emergency severity index in older emergency department patients. Ann Emerg Med. 2012;60(3):317–25.
7. Parenti N, Reggiani MLB, Iannone P, Percudani D, Dowding D. A systematic review on the validity and reliability of an emergency department triage scale, the Manchester Triage System. Int J Nurs Stud. 2014;51(7):1062–9.
8. Farrohknia N, Castrén M, Ehrenberg A, Lind L, Oredsson S, Jonsson H, 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.
9. Mackway-Jones K, Marsden J, Windle J. Emergency Triage - Manchester Triage Group. Hoboken, NJ: Blackwell Publishing - BMJ Books; 2006.
10. Morris JH, James RE, Davey R, Waddington G. What is orthopaedic triage? A systematic review. J Eval Clin Pract. 2015;21(1):128–36.
11. Desmeules F, Toliopoulos P, Roy J-S, Woodhouse LJ, Lacelle M, Leroux M, et al. Validation of an advanced practice physiotherapy model of care in an orthopaedic outpatient clinic. BMC Musculoskelet Disord. 2013;14:162.
12. Pereira MG. Avaliação de testes diagnósticos. In: Andrade AL, Zicker F, editors. Métodos de Investigação Epidemiológica em Doenças Transmissíveis. Brasília: Organização Pan-Americana da Saúde; 1997. p. 22–40.
13. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159–74.
14. van Veen M, Moll HA. Reliability and validity of triage systems in paediatric emergency care. Scand J Trauma Resusc Emerg Med. 2009;17:38.
15. Rotondo MF, Cribari C, Smith RT. Resources For Optimal Care Of the Injured Patient. American College of Surgeons; 2014.
16. Roukema J, Steyerberg EW, van Meurs A, Ruige M, van der Lei J, Moll HA. Validity of the Manchester Triage System in paediatric emergency care. Emerg Med J. 2006;23(12):906–10.
17. Storm-Versloot MN, Ubbink DT, Kappelhof J, Luitse JSK. Comparison of an informally structured triage system, the emergency severity index, and the manchester triage system to distinguish patient priority in the emergency department. Acad Emerg Med. 2011;18(8):822–9.
18. van Veen M, Steyerberg EW, Ruige M, van Meurs AHJ, Roukema J, van der Lei J, et al. Manchester triage system in paediatric emergency care: prospective observational study. BMJ. 2008;337:a1501.
19. Twomey M, Wallis LA, Myers JE. Limitations in validating emergency department triage scales. Emerg Med J. 2007;24(7):477–9.
20. Cooke MW, Jinks S. Does the Manchester triage system detect the critically ill? J Accid Emerg Med. 1999;16(3):179–81.
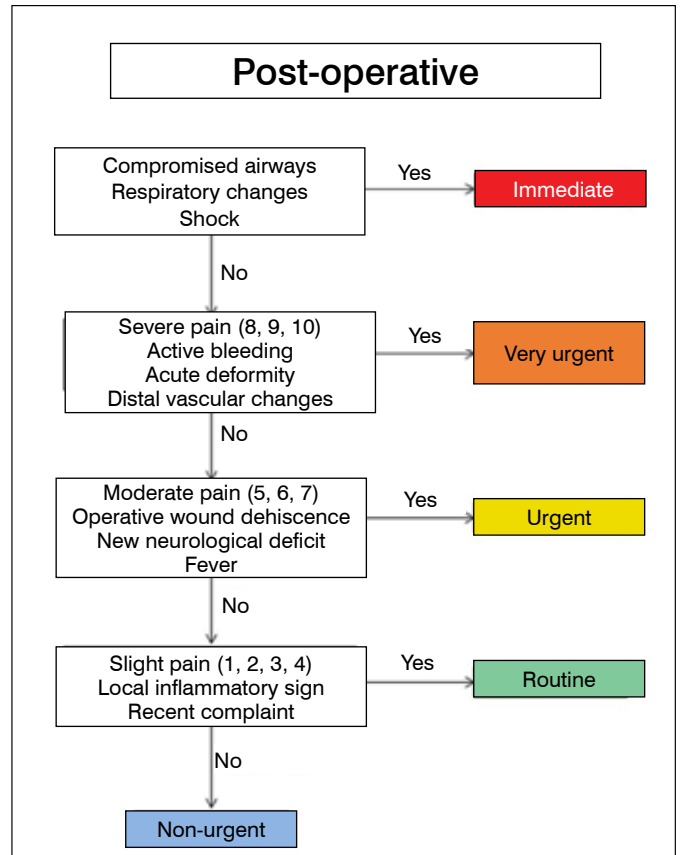
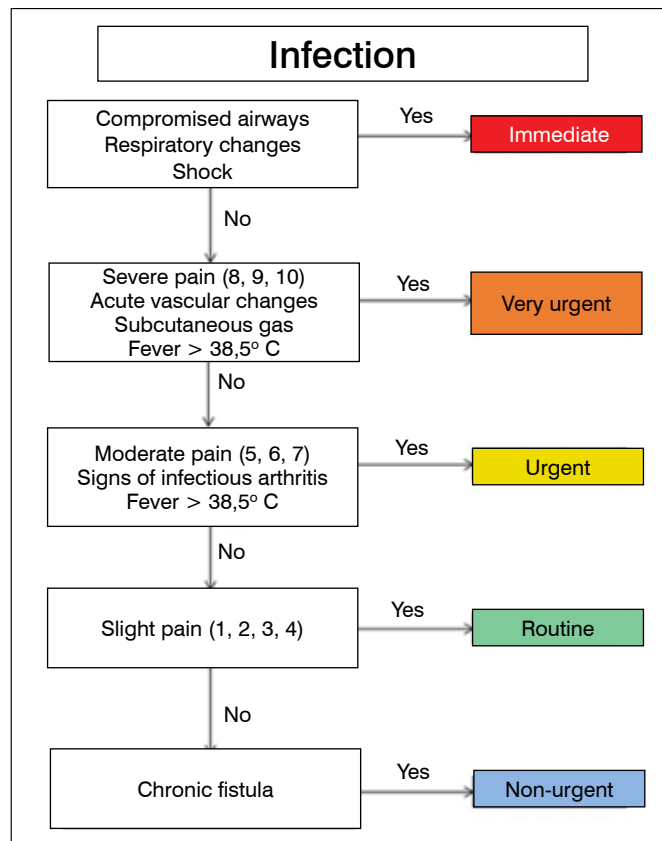**Annex 1.** Triage flow chart - Traumatic injuries.

Annex 2. Triage flow chart - Joint pain.

Vertebral pain

Compromised airways / Respiratory changes / Shock	Yes →	**Immediate**

No ↓

Severe pain (8, 9, 10) / Acute neurological deficit / High-energy trauma	Yes →	**Very urgent**

No ↓

Moderate pain (5, 6, 7) / New neurological deficit / Inability to walk / Fever	Yes →	**Urgent**

No ↓

Slight pain (1, 2, 3, 4)	< 3 weeks →	**Routine**

> 3 weeks ↓

Non-urgent

Annex 3. Triage flow chart - Vertebral pain.

Post-operative

Compromised airways / Respiratory changes / Shock	Yes →	**Immediate**

No ↓

Severe pain (8, 9, 10) / Active bleeding / Acute deformity / Distal vascular changes	Yes →	**Very urgent**

No ↓

Moderate pain (5, 6, 7) / Operative wound dehiscence / New neurological deficit / Fever	Yes →	**Urgent**

No ↓

Slight pain (1, 2, 3, 4) / Local inflammatory sign / Recent complaint	Yes →	**Routine**

No ↓

Non-urgent

Annex 4. Triage flow chart - Postoperative disorder.

Infection

Compromised airways / Respiratory changes / Shock	Yes →	**Immediate**

No ↓

Severe pain (8, 9, 10) / Acute vascular changes / Subcutaneous gas / Fever > 38,5º C	Yes →	**Very urgent**

No ↓

Moderate pain (5, 6, 7) / Signs of infectious arthritis / Fever > 38,5º C	Yes →	**Urgent**

No ↓

Slight pain (1, 2, 3, 4)	Yes →	**Routine**

No ↓

Chronic fistula	Yes →	**Non-urgent**

Annex 5. Triage flow chart - Musculoskeletal infection.