Original Article

Access this article online

Quick Response Code:

Website:
www.jehp.net

DOI:
10.4103/jehp.jehp_1276_24

Urgent surgical patient classification: Development and validation of a novel instrument using the Delphi approach

Priscila Buck de Oliveira Ruiz, Jaqueline Lopes Gouveia, Paula Buck de Oliveira Ruiz, João Daniel de Souza Menezes, Matheus Querino da Silva, Emerson Roberto dos Santos, Alba Regina de Abreu Lima, Vânia Maria Sabadoto Brienze, Júlio César André, Rita de Cássia Helu Mendonça Ribeiro

Abstract:

BACKGROUND: The lack of a triage system for urgent surgical patients leads to non-standardized decision-making. Developing an instrument to objectively identify the complexity of care required for each case is challenging. The aim of this report is to develop and validate an urgent surgical patient classification instrument using the Delphi technique.

MATERIALS AND METHODS: The study was conducted in several stages: (1) definition of the construct; (2) item elaboration; (3) expert analysis; (4) item selection; (5) pretest. In the first study, scale items were designed and content validity was confirmed. In the second study, the factorial structure was analyzed. In the third study, alternative measurement models were tested. In the fourth study, criterion validity was analyzed.

RESULTS: Experts validated 14 items (31.81%) from Domain 1 with 75% agreement and specific items from Domain 2 with 100% agreement. Factor analysis indicated a two-factor solution explaining 58.4% of the variance. The bifactor model presented the best fit ($\chi2/df = 1.51$; CFI = 0.95; TLI = 0.94; RMSEA = 0.051; SRMR = 0.043). Factors showed excellent internal consistency ($\alpha > 0.88$; CR > 0.90; $\omega > 0.92$).

CONCLUSION: This pioneering study developed and validated the content of the first specific instrument for classifying urgent surgical patients regarding their priority for care. The instrument was deemed valid in terms of content, based on expert consensus. Further studies are recommended to evaluate its practical application and perform additional psychometric measures. This instrument has the potential to enhance the organization of emergency services and operating theaters, promoting patient safety and efficient resource management in healthcare institutions. Its implementation should align with current health guidelines and policies to optimize the triage and prioritization process for urgent surgical patients.

Keywords:

Delphi technique, surgery, surveys and questionnaires, triage, validation studies

Faculty of Medicine De São José Do Rio Preto, Brazil

Address for correspondence:
Dr. Joao Daniel de Souza Menezes, Av. Brg. Faria Lima, 5416 - Vila Sao Pedro, São José do Rio Preto - SP, 15090-000. Brazil.
E-mail: joao.menezes@edu.famerp.br

Received: 12-07-2024
Accepted: 09-12-2024
Published: 30-04-2025

Introduction

Currently, hospitals represent one of the most intricate organizational structures in modern society. The term "hospital" derives from the Latin "hospe," meaning "one who receives," and defines an establishment suitable for the hospitalization and treatment of the sick and injured.[1] This institution's origins can be traced back to the medieval period, when individuals seeking medical attention began to turn to convents rather than individual physicians. Initially, the primary purpose of these establishments was to provide lodging or refuge for the elderly, disabled, and homeless, offering care that was predominantly religious in nature.[2,3]

How to cite this article: de Oliveira Ruiz PB, Gouveia JL, de Oliveira Ruiz PB, de Souza Menezes JD, da Silva MQ, dos Santos ER, *et al*. Urgent surgical patient classification: Development and validation of a novel instrument using the Delphi approach. J Edu Health Promot 2025;14:177.

In the contemporary context, hospitals are regarded as complex enterprises whose primary characteristic is the provision of services. As such, all fundamental principles of management—including forecasting, organization, command, coordination, and control—are applicable to these institutions.[4] Hospitals, to a greater extent than most other enterprises, encompass a multifaceted array of interconnected elements, resulting in a high degree of operational complexity. Within these environments, one can find sophisticated technological structures pertaining to both material and human resources. These elements coalesce to form an organizational system capable of delivering high-quality care, with the surgical center unit being a crucial component of this system.[4]

Figure 1 illustrates how the development of the classification instrument emerges as a response to the challenges identified in the surgical center context, aiming to achieve the expected outcomes of improved management and care for urgent surgical patients.

According to the Ministry of Health (1977), the surgical center unit is formed by the set of elements intended for surgical activities, as well as postanesthetic and immediate postoperative recovery.[5] It is composed of the surgical center (SC), the anesthesia recovery (AR), and the material and sterilization center (MSC). The operating room is an organic functional unit consisting of an integrated set of human, physical, and technical means, intended for the provision of surgical treatment or the performance of examinations that require a high level of quality.[6]

Considered one of the most complex contexts in the health area, the teams of health professionals are highly trained and qualified, exercising their activity in complementarity and interacting with advanced technology, in high-risk situations, with responsibility to respond to the needs of the surgical patient.[7] It is estimated that around 240 million surgeries are performed worldwide each year and that there will be an increase in the incidence of surgical diseases in the next decade, represented by cardiovascular diseases, trauma, and cancer, associated with the population's longer life expectancy.[8] The absence of a clinical triage system for surgical patients who require urgent interventions for their classification regarding their priority of care means that decisions are made without standards. The development of an instrument that enables urgent surgical patient care services to objectively and practically identify the complexity of care that each case requires is complex. In addition, it requires the mobilization and capacity of knowledge from various areas and consumes several resources; however, it is believed that it will allow health teams to better dimension the use of installed capacity and better conditions to judge priorities within a group of patients awaiting the surgical act.[9] Man has performed surgical practices since antiquity. The term surgery, from the Greek kheirourgia, which means "*manual work*," can be defined as the specialty that is intended for the treatment of diseases and injuries through manual and instrumental operative processes.[1] In the Middle Ages, surgeries were performed on battlefields, in surgeons' homes, or under the decks of warships. The operations were restricted to limb amputations, abscess drainage, and tumor removal and were performed on the human body only with the use of hands or the aid of instruments. Patients undergoing surgical treatments had to overcome the pain, hemorrhage, and infection generated by the procedures without anesthesia.[2,10]

Surgery had an important evolution from 1846 when anesthesia was discovered, and narcosis began to be performed by ether inhalation. The possibility of performing a painless surgical procedure became real.[11] Surgical treatment can be classified according to the operative moment, the purpose of the surgery, the cardiological risk to which the patient is subjected, the duration of the surgery, and the potential for contamination. Surgeries classified according to the operative moment are divided into elective surgeries that can be performed with prescheduled dates, emergency surgery that, due to the severity of the patient's clinical condition, requires immediate intervention without the possibility of preoperative evaluation, and urgent surgery that requires mediate intervention and can wait a few hours, during which the patient is kept under clinical and laboratory evaluation and observation.[12]

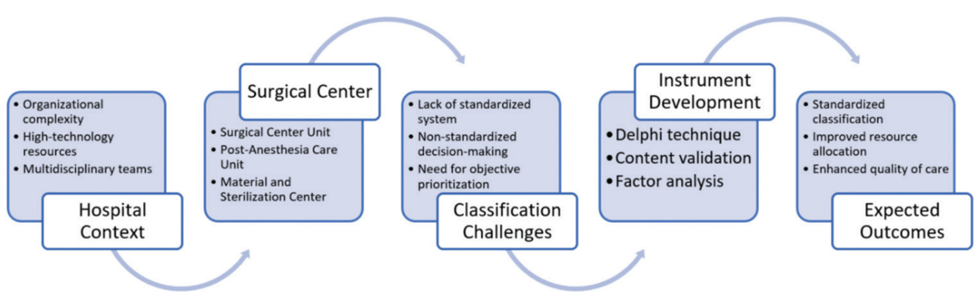

Figure 1: Theoretical framework for urgent surgical patient classification

Access to surgery is a complex issue that requires special attention within the strategic planning of health. Complex actions are understood as those that have a wide spectrum in the patient's line of care, which can be expressed from diagnostic practices to resolving treatment. Presurgical choices are permeated by discussions that involve difficult decisions, as the results can be unexpected, and there is a duty to consider ethical discussions and differences of opinion from specialist professionals in choosing the best opportunity for surgical intervention.[13]

The word "*choose*" means "*preference given to something that is found among others; predilection*."[1] We know that in the health field, the act of choosing must be supported by scientific data so that subjectivity is transformed into objectivity and assertiveness. The risk classification of a patient is a dynamic process of identifying the needs for immediate treatment according to the potential for risk, health problems, or the degree of suffering. Its importance consists of preventing complications and identifying acute conditions that imply a risk of death for individuals.[14]

Among the most used triage systems in the world, the following stand out: Emergency Severity Index (ESI), Australian Triage Scale (ATS), Canadian Triage Acuity Scale (CTAS), Manchester Triage System (MTS). These classifications are used in hospital emergency rooms for prioritization in medical care.[15] Triage for classification of care is commonly applied in catastrophes and in emergency units. However, in a surgical center scenario, with many elective and nonelective surgeries, there is no standardized nomenclature for classifying urgent surgeries. In most hospitals, obtaining an operating room for an urgent surgery depends on dialogue and negotiation; in others, urgent surgeries are performed in order of arrival.[15]

In 2013, the World Society of Emergency Surgery Study Group (WSES) recommended the use of a color system to classify urgent and emergency surgeries, aiming to reduce the loss of information and allowing the establishment of standardized language among teams. The classification of nonelective surgeries by time (Timing of Acute Care Surgery—TACS) was based on a survey conducted with a panel of experts regarding the ideal time for the most frequent emergency surgeries.[16] This classification is divided by colors, with red for immediate emergency care up to blue with surgeries for care within 48 hours, made through the analysis of the type of surgical groups, not differentiating regarding the severity of the patients.

There are instruments in which patients are assessed for the risks of complications and mortality after surgery. The following stand out: Revised Cardiac Risk Index (RCRI), Universal Surgical Risk Calculator (ACS-NSQIP), National Surgical Quality Improvement Program (NSQIP), American Society of Anesthesiologists (ASA), Acute Physiology and Chronic Health Evaluation (APACHE), 2nd Simplified Acute Physiology Score (SAPS II), Sequential Organ Failure Assessment (SOFA), Physiological and Operative Severity Score for the Enumeration of Mortality and Morbidity (POSSUM), Goldman cardiac risk, Sutton surgical risk score, Biochemistry and Hematology Outcome Models (BHOM).[17-21]

Among the various ways of developing and validating an instrument, the Delphi technique is found. In general terms, it is related to the precision of the instrument in measuring what it proposes, seeking consensus among judges who have mastery of the subject in question.[22] In general, when it comes to the validation of measurement instruments by the Delphi technique, the best-known techniques are content validity, face validity, criterion validity, and construct validity.[23]

This study proposed to develop and validate the content of an urgent surgical classification instrument using the Delphi technique.

Materials and Methods

Study design and setting

In this research program, we intend to develop an instrument to classify urgent patients who require a surgical procedure. In the first study, we designed the scale items and confirmed their content validity through expert evaluators and conducted a pilot study in a sample of the target population. In the second study, we analyzed the factorial structure of the scale to verify if the items loaded factors that describe the theoretical dimensions proposed for the instrument. In the third study, we tested several alternative measurement models to confirm the factorial structure found in Study 2. In the fourth study, we analyzed the criterion validity of the scale by testing the hypothesis that the instrument is capable of discriminating urgent surgical patients according to the severity of the clinical condition.

For the Study 1, we based ourselves on the literature mentioned in the introduction to develop the urgent surgical patient classification instrument. The elaboration of the items was done in several stages: (1) definition of the construct of urgent surgical patient classification and identification of its theoretical dimensions; (2) elaboration of items based on the definition of the theoretical dimensions of urgent surgical patient classification; (3) analysis by expert evaluators; (4) selection of the best items to describe the construct; (5) pretest with the target population to verify the clarity of the cognitive understanding of the items. We developed a comprehensive set of items for the Urgent Surgical Patient Classification Instrument, comprising two domains:

Domain 1: General approach signs and symptoms (44 items)

Domain 2: Specific signs and symptoms by specialty (six items).

These items were designed to assess how healthcare professionals would respond to and classify urgent surgical patients. The complete list of items and their predicted theoretical dimensions is presented in Table 1.

After the elaboration of the scale items, we asked the experts to assess the adequacy of each item in relation to the theoretical dimensions for which it was developed.

Study 2: Factorial validity of the Urgent Surgical Patient Classification Scale (USPCS)

The objective of this Study 2 was to investigate whether the factorial structure of the Urgent Surgical Patient Classification Scale (USPCS) empirically corresponds to the theoretical dimensions predicted for the scale: general approach signs and symptoms and specific signs and symptoms by specialty. In addition, we analyzed the quality of the items using item response theory (IRT). Next, we estimated the convergent–discriminant validity of the USPCS using measures that theoretically should be more strongly correlated with the scale (i.e., convergent validity) and measures that should be very weakly correlated with the scale (i.e., discriminant validity). We also analyzed whether beliefs about the importance of a classification instrument predict factors related to its applicability in clinical practice, which gives us a first indication of the scale's criterion validity.

In the following study, Study 3, we performed a more in-depth analysis of these parameters and tested several alternative measurement models for the factorial structure of the scale.

This study aims to expand the analysis of the factorial validity of the USPCS by testing different measurement models in its structure. First, we tested the hypothesis that the scale items reflect the multidimensional structure of urgent surgical patient classification (i.e., the two factors we identified in previous studies: general signs and symptoms, and specific signs and symptoms). We tested this factorial structure through confirmatory factor analysis, comparing it with three alternative models. The first alternative model specified a single latent factor to explore the possibility of a single general factor underlying the scale items. The second alternative model tested a hierarchical multifactorial structure in which the two first-order factors loaded a second-order general factor. Finally, alternative model 3 specified a bifactor structure. This model predicted that the scale would assess urgent surgical patient classification using two specific factors (i.e., S factors) and a general factor (i.e., G factor).

Table 1: Items of the Urgent Surgical Patient Classification Scale and their predicted theoretical dimensions (Study 1)

Dimension	Item
General signs and symptoms	1. Surgical reapproach
	2. Pyrexia (39°C to 40°C)
	3. Hyperpyrexia (above 40°C)
	4. Severe hypothermia (less than 33.9°C)
	5. Tachycardia (above 100 bpm)
	6. Oxygen saturation below 91%
	7. Hypotension
	8. Leukocytosis greater than 12,000
	9. Hemoglobin below 6 g/dl
	10. Moderate blood loss
	11. Massive blood loss
	12. Torpor
	13. Dyspnea
	14. Tachypnea
Specific signs and symptoms	General surgery
	1. Intestinal obstruction
	2. Peritonitis
	3. Pneumoperitoneum
	4. Fluid in the cavity
	5. Crepitus on palpation
	Neurosurgery
	1. Acute neurological deficit
	2. Glasgow Coma Scale 9–13 (moderate)
	3. Anisocoria
	4. ASIA C
	Orthopedics
	1. Gustilo–Anderson Grade I open fracture
	2. Gustilo–Anderson Grade II open fracture
	3. Gustilo–Anderson Grade III A open fracture
	4. Tscherne–Gotzen Grade 1 closed fracture
	5. Tscherne–Gotzen Grade 2 closed fracture
	Plastic surgery
	1. Hematoma
	Urology
	1. Renal obstruction
	2. Pyuria
	3. Hematuria
	Thoracic surgery
	1. Partial airway obstruction
	2. Hemoptysis
	3. Hemothorax

Source: Authors

Study participants and sampling

For the Study 1, the expert panel comprised 22 professionals (Mage = 42.5 years; SD = 8.2) selected based on their extensive experience with studies on urgent surgical patient classification.

A sample of the target population, totaling 30 surgeons, took part in the pretest.

Data collection tolls and techniques

In the Study 1, recruitment was conducted via email, which included a comprehensive explanation of the study objectives and a request for their participation in the item analysis process.

To assess the representativeness of the items within their respective theoretical domains, the evaluators examined 44 items from Domain 1 and six items from Domain 2. The evaluation criteria included: the item's representation of its theoretical domain, accuracy of representation, clarity of content presentation, and relevance to the proposed dimension. Assessments were made using a four-point Likert scale, ranging from 1 (unimportant) to 4 (very important) for each evaluated aspect.

Participants accessed an online questionnaire to conduct their assessments. To enhance the accuracy of item selection and determine content adequacy, we calculated the Kappa coefficient.[24,25] This coefficient estimates the appropriateness of items within each dimension of the urgent surgical patient classification measure and indicates the degree of inter-rater agreement. A higher coefficient value signifies greater agreement. For the criteria of precision, clarity, and relevance, we calculated the content validity coefficient (CVC), which estimates the consistency of each item's content among experts.[26]

For the pretest, the instrument containing the 50 items (44 from Domain 1 and six from Domain 2) was applied to a sample of 30 surgeons to assess the understanding of the items.

In the Study 2, 22 specialist physicians, aged 30 to 79 years (M = 45.5; SD = 12.3), from public and private hospitals in the cities of São José do Rio Preto-SP, São Paulo-SP, Campinas-SP, and Primavera-MT, Brazil, were observed. In a sample sensitivity analysis using the observed degrees of freedom (df = 209) and defining an acceptable hypothetical fit quality of the measurement model to the data (based on RMSEA = 0.05), our sample size (n = 22) provides a power of 0.80.

Data collection was done individually after participants signed the informed consent form (ICF). We informed them about the confidential nature of the study so that they could not identify themselves and about the possibility of interrupting the study at any time without consequences if they so desired. Completing the survey would take an average of 20 minutes. Participants responded to the instruments described below.

In the Urgent Surgical Patient Classification Scale (USPCS), developed in Study 1, participants indicated their agreement with each item on a scale ranging from 1 (Unimportant) to 4 (Very Important).

An Agreement Questionnaire containing statements for agreement analysis regarding the importance, clarity, objectivity, applicability, and performance of the USPCS items and the participants indicated their degree of agreement on a scale from 1 (Disagree) to 4 (Fully Agree).

We analyzed the data using SPSS statistical software (version 25). First, we calculated descriptive statistics to characterize the sample (frequencies, means, and standard deviations). Then, principal component factor analysis (PCA) with varimax rotation of the USPCS was performed. We conducted a regression analysis to determine whether the importance attributed to the USPCS and agreement with its attributes predicted factors related to its applicability. We estimated Cronbach's alpha and composite reliability to test internal consistency in addition to McDonald's omega. We used Mplus software to estimate the parameters of item response theory.[27,28]

For the study, a sample comprised 200 Brazilian surgeons, with ages ranging from 28 to 65 years (M = 42.5; SD = 9.8). A sensitivity analysis was conducted using Webpower to ensure adequate sample size.[29] The analysis confirmed that our sample was sufficient to detect an acceptable fit to the data, based on RMSEA < 0.08 and Power = 0.80.

We employed Mplus software version 8.3 to conduct a series of confirmatory factor analyses using the maximum likelihood estimator.[27] Model fit was assessed using multiple indices:

- $\chi 2$/df (Chi-square/degrees of freedom)
- CFI (Comparative Fit Index)
- RMSEA (Root Mean Square Error of Approximation)
- SRMR (Standardized Root Mean Square Residual)
- TLI (Tucker Lewis Index).

The following criteria were applied to evaluate model fit:[30]

- $\chi 2/df < 4.00$
- CFI > 0.95
- TLI ≥ 0.90
- RMSEA < 0.08
- SRMR < 0.10

To assess the internal consistency of the USPCS, we calculated Cronbach's alpha, composite reliability, and McDonald's omega. Additionally, we examined the bivariate correlation between the USPCS and a measure of clinical condition severity to analyze convergence.

For the bifactor model, we followed the procedures proposed by Eid *et al*. (2017)[31] for statistical identification and estimation. We specified the model with a reference indicator (SI—1), designating item 4 as the reference for the general factor.

We made all data publicly available through the Open Science Framework (DOI 10.17605/OSF.IO/EQ7HG).

Ethicals considerations

In all studies, participants signed the informed consent form (ICF), which explained the objectives and guaranteed the anonymity of their responses. All procedures used in the studies were approved by the Research Ethics Committee of the Faculty of Medicine of São José do Rio Preto (CAEE: 2652137/2018, dated 02/17/2021).

Results

In the Study 1 (Scale Development and Content Validity), the analysis yielded an intraclass correlation coefficient (ICC) of 0.85 for evaluator agreement and a Kappa of 0.78 for item adequacy. Regarding precision, clarity, and relevance, we obtained satisfactory CVC values: CVC Precision = 0.92; CVC Clarity = 0.90; CVC Relevance = 0.95. These results provide initial evidence supporting the content validity of the Urgent Surgical Patient Classification Scale items. Both the reliability of the items and the level of agreement among evaluators were deemed satisfactory, suggesting that the items possess sufficient quality for testing in the target population.[24,25,32] Additionally, the expert evaluators validated 14 items (31.81%) from Domain 1 with 75% agreement and the specific items from Domain 2 with 100% agreement for each specialty.

In the pretest, the items obtained clarity means ranging from 3.2 to 3.9 (M = 3.6; SD = 0.2) and comprehension means ranging from 3.4 to 4.0 (M = 3.7; SD = 0.2), indicating that they were well understood by the target population. There were no suggestions for modifications in the wording of the items.

In the Study 2, the exploratory factor analysis revealed a two-factor solution that accounts for 58.4% of the total variance in the Urgent Surgical Patient Classification Scale. This structure aligns with the conceptual design of the instrument:

Factor 1: Encompasses the 14 validated items from Domain 1 (general approach signs and symptoms) Factor 2: Comprises the specific validated items from Domain 2 for each specialty.

Table 2 presents a detailed visualization of the factorial structure, illustrating how each item loads onto the two identified factors.

Table 2: Factorial structure of the Urgent Surgical Patient Classification Scale (Study 2)

Item	Factor 1	Factor 2
1. Surgical reapproach	0.72	
2. Pyrexia (39°C to 40°C)	0.68	
3. Hyperpyrexia (above 40°C)	0.81	
4. Severe hypothermia (less than 33.9°C)	0.75	
5. Tachycardia (above 100 bpm)	0.66	
6. Oxygen saturation below 91%	0.79	
7. Hypotension	0.82	
8. Leukocytosis greater than 12,000	0.59	
9. Hemoglobin below 6 g/dl	0.77	
10. Moderate blood loss	0.71	
11. Massive blood loss	0.84	
12. Torpor	0.62	
13. Dyspnea	0.74	
14. Tachypnea	0.69	
15. Intestinal obstruction		0.88
16. Peritonitis		0.92
17. Pneumoperitoneum		0.86
18. Fluid in the cavity		0.79
19. Crepitus on palpation		0.83
20. Acute neurological deficit		0.91
21. Glasgow Coma Scale 9–13 (moderate)		0.87
22. Anisocoria		0.82
24. ASIA C		0.76
25. Gustilo–Anderson Grade I open fracture		0.81
26. Gustilo–Anderson Grade II open fracture		0.85
27. Gustilo–Anderson Grade III A open fracture		0.89
28. Tscherne–Gotzen Grade 1 closed fracture		0.78
29. Tscherne–Gotzen Grade 2 closed fracture		0.80
30. Hematoma		0.73
31. Absence/decrease of peripheral pulse		0.84
32. Compartment syndrome		0.90
33. Fixed cyanosis		0.86
34. Acute ischemia		0.92
35. Capillary refill >2 seconds		0.81
36. Renal obstruction		0.77
37. Pyuria		0.74
38. Hematuria		0.79
39. Partial airway obstruction		0.88
40. Hemoptysis		0.83
41. Hemothorax		0.87

Source: Authors

The analysis demonstrated that the items loaded consistently on their proposed factors, aligning with the semantic content of each dimension.

To assess the reliability of the USPCS, we employed multiple measures of internal consistency. The results were as follows:

• Factor 1: Cronbach's α = 0.89; composite reliability (CR) = 0.91 • Factor 2: Cronbach's α = 0.85; composite reliability (CR) = 0.88 • Total scale: McDonald's ω = 0.93.

Item Parameters

Table 3: Parameters of the USPCS items according to the graded response model (Study 2)

Item	A	B1	B2	B3
1	1.32	-1.85	-0.92	0.78
2	1.18	-1.62	-0.74	0.91
3	1.76	-1.39	-0.58	0.64
4	1.47	-1.51	-0.69	0.82
5	1.09	-1.93	-1.08	0.95
6	1.58	-1.44	-0.61	0.71
7	1.81	-1.28	-0.47	0.59
8	0.92	-2.12	-1.35	1.17
9	1.52	-1.48	-0.65	0.76
10	1.29	-1.88	-0.95	0.81
11	1.93	-1.19	-0.38	0.52
12	1.03	-2.05	-1.22	1.08
13	1.41	-1.56	-0.72	0.85
14	1.22	-1.79	-0.87	0.98
15	2.24	-1.02	-0.25	0.41
16	2.58	-0.87	-0.12	0.29
17	2.09	-1.11	-0.32	0.48
18	1.69	-1.35	-0.54	0.68
19	1.88	-1.22	-0.41	0.56
20	2.43	-0.93	-0.18	0.34
21	2.14	-1.08	-0.29	0.45
22	1.95	-1.17	-0.36	0.51
23	1.58	-1.44	-0.61	0.71
24	1.76	-1.39	-0.58	0.64
25	1.98	-1.15	-0.34	0.49
26	2.31	-0.98	-0,21	0.38
27	1.64	-1.39	-0.57	0.73
28	1.71	-1.33	-0.52	0.66
29	1.38	-1.59	-0.75	0.88
30	1.93	-1.19	-0.38	0.52
31	2.38	-0.95	-0.20	0.36
32	2.09	-1.11	-0.32	0.48
33	2.62	-0.85	-0.10	0.27
34	1.76	-1.39	-0.58	0.64
35	1.55	-1.46	-0.63	0.74
36	1.41	-1.56	-0.72	0.85
37	1.64	-1.39	-0.57	0.73
38	2.19	-1.05	-0.27	0.43
39	1.88	-1.22	-0.41	0.56
40	2.02	-1.13	-0.33	0.47
41	1.69	-1.35	-0.54	0.68
42	1.81	-1.28	-0.47	0.59
43	1.93	-1.19	-0.38	0.52
44	2.28	-1.00	-0.23	0.39

a=discrimination parameter; b1, b2, b3=difficulty parameters. Source: Authors

We conducted an item analysis using item response theory (IRT) to evaluate whether the exclusion of items resulting from the factor structure analysis would impact the quality of measurement of the general latent trait [Table 3].[28]

Employing Samejima's graded response model, we tested the assumption of unidimensionality of the latent trait.[33] The analysis revealed that the items demonstrated discrimination parameters ranging from moderate to high, according to Baker's (2001) criteria.[34] This finding suggests a positive impact on the formation of the latent trait, indicating that these items effectively discriminate between respondents with varying levels of the latent trait.

For subsequent analyses, we utilized these parameters as item adequacy criteria, with a threshold value > 0.65.[34] This approach aligns with the Kaiser criterion analysis, which indicated that non-validated items did not load adequately on any factor, justifying their exclusion.

Regarding item difficulty (or agreement), we observed latent variable values (theta estimates) ranging from -2.12 (item 5) to 1.89 (item 22), demonstrating a wide spectrum of item difficulty across the scale.

In the Study 3, we estimated and compared several measurement models to determine the optimal factorial structure for the Urgent Surgical Patient Classification Scale (USPCS). Table 4 presents a comprehensive comparison of fit indices for these models.

Key findings from our analysis include:
1. The unifactorial model demonstrated the poorest fit, particularly in terms of the Standardized Root Mean Square Residual (SRMR).
2. The two-factor first-order model showed a very good fit to the data, significantly outperforming the one-factor model.
3. The hierarchical multidimensional model exhibited identical fit indices to the two correlated factors model.
4. The bifactor structure model emerged as the best-fitting model, demonstrating superior fit indices, especially in terms of SRMR. This model significantly outperformed all other tested models.

Table 4 provides a detailed comparison of the fit indices across the different models:

Internal consistency

We assessed the internal consistency of the two dimensions of the USPCS, yielding the following coefficients:
- Factor 1 (General signs and symptoms): Cronbach's $\alpha = 0.91$; composite reliability (CR) =0.93; McDonald's $\omega = 0.92$
- Factor 2 (Specific signs and symptoms): Cronbach's $\alpha = 0.88$; CR = 0.90
- General factor: Cronbach's $\alpha = 0.95$; CR = 0.96.

These results indicate excellent internal consistency across all dimensions of the USPCS.

Table 4: Fit quality of the USPCS measurement models (Study 3)

Model	χ^2 (df)	χ^2/df	CFI	TLI	RMSEA (IC 90%)	SRMR
1 factor	2145.67 (902)	2.38	0.87	0.86	0.082 (0.078–0.086)	0.068
2 factors	1532.24 (901)	1.70	0.93	0.92	0.059 (0.054–0.064)	0.051
2nd order	1532.24 (901)	1.70	0.93	0.92	0.059 (0.054–0.064)	0.051
Bifactorial	1294.81 (855)	1.51	0.95	0.94	0.051 (0.046–0.056)	0.043

Source: Authors

Discussion

This study aimed to develop and validate the content of an instrument for classifying urgent surgical patients according to their priority for care. The results provide evidence supporting the content validity of the Urgent Surgical Patient Classification Scale (USPCS), demonstrating its potential as a valuable tool for healthcare professionals in emergency and surgical settings.

The Delphi technique, employed in this study, has been widely recognized as an effective method for achieving consensus among experts in various fields, including health care.[35-37] In the context of instrument development, this technique offers several advantages. Firstly, it allows for the systematic collection and aggregation of expert opinions, which is particularly valuable when empirical evidence is limited or conflicting.[38-40] Secondly, the iterative nature of the Delphi process facilitates refinement and convergence of ideas, leading to more robust and consensual outcomes.[41,42] In our study, the application of the Delphi technique enabled us to harness the collective expertise of healthcare professionals, ensuring that the USPCS items accurately reflect the critical factors in urgent surgical patient classification.[43-45] The content validation process, a crucial step in instrument development, yielded satisfactory results. The high content validity coefficient (CVC) values for precision, clarity, and relevance provide strong evidence for the quality and appropriateness of the USPCS items. These findings align with established guidelines for content validation in healthcare research, which emphasize the importance of expert evaluation in ensuring instrument validity.[46-48] Similar approaches have been successfully employed in recent studies, such as the development of a tool for assessing the quality of mobile health applications, where expert validation played a crucial role in ensuring the instrument's relevance and applicability.[37,49,50]

The factorial structure of the USPCS, as revealed by exploratory and confirmatory factor analyses, supports the theoretical underpinnings of the instrument. The emergence of two distinct factors—general signs and symptoms, and specific signs and symptoms—reflects the multifaceted nature of urgent surgical patient assessment. This structure aligns with current understanding in emergency medicine and surgical triage, which recognizes the importance of both general health indicators and specialty-specific symptoms in patient prioritization.[51-53]

The item validation process employed in this study was rigorous and multifaceted. Initially, expert evaluation through the Delphi technique provided a foundation for item selection and refinement. Subsequently, the application of item response theory (IRT) offered valuable insights into item performance. The use of Samejima's graded response model allowed for a nuanced understanding of item discrimination and difficulty, crucial aspects in assessing the quality of measurement instruments.[54-56] The observed range of discrimination parameters, from moderate to high, indicates that the USPCS items effectively differentiate between patients with varying levels of urgency. Furthermore, the wide spectrum of item difficulty ensures that the scale can accurately assess patients across different severity levels, enhancing its clinical utility.[57-59] This approach is consistent with recent advancements in healthcare instrument development, as demonstrated in a study that developed and validated a questionnaire to assess knowledge about cervical cancer and human papillomavirus, where rigorous item analysis contributed to the instrument's reliability and validity.[60]

The internal consistency measures obtained for the USPCS demonstrate its reliability as a measurement tool. The high Cronbach's alpha, composite reliability, and McDonald's omega coefficients across both factors and the general scale indicate strong inter-item correlations and overall scale coherence. These findings suggest that the USPCS provides consistent and reliable measurements, a critical attribute for any clinical assessment instrument.[61-63]

The bifactor model's superior fit, as evidenced by the confirmatory factor analysis, offers important insights into the structure of the USPCS. This model suggests the presence of a strong general factor underlying the specific dimensions, which aligns with the complex, multidimensional nature of surgical urgency assessment. The bifactor structure allows for the simultaneous consideration of both general and specific aspects of patient condition, potentially enhancing the scale's sensitivity and specificity in clinical settings.[64-66]

Limitation and recommendation

While these results are promising, it is important to acknowledge the limitations of this study. The validation process was conducted in a specific cultural and healthcare context, which may limit the generalizability of the findings. Future research should focus on cross-cultural validation and testing in diverse healthcare settings to ensure the USPCS's applicability across different populations and healthcare systems. This approach is particularly important in healthcare instrument development, as highlighted by a recent study that emphasized the need for cultural adaptation and validation of health literacy assessment tools.[67-72]

Conclusions

This pioneering study developed and validated the content of the first specific instrument for classifying urgent surgical patients regarding their priority for care. Given the paucity of instruments for this purpose, the development and validation of this tool have the potential to significantly contribute to hospital practice, facilitating a more objective and standardized patient assessment.

Content validity, established through expert consensus using the Delphi technique, constitutes a fundamental step in the construction of assessment instruments, ensuring that the tool is both safe and reliable for its intended purpose.

The results of this study indicate that the instrument is deemed valid in terms of content, based on the importance attributed by experts to the items for classifying urgent surgical patients regarding their priority for care.

Further studies are recommended to evaluate the practical application of the instrument and to perform additional psychometric measures, such as construct validity and reliability. The instrument may also be utilized in future scientific research endeavors.

The advancement represented by the development and validation of this instrument has the potential to enhance the organization of emergency services and operating theaters, promoting patient safety and efficient management of financial and temporal resources in healthcare institutions.

It is imperative to emphasize that this conclusion is predicated on health policy benchmarks, taking into account best practices and established standards in the field of hospital management and emergency care. The implementation of this instrument should be aligned with current health guidelines and policies, aiming to optimize the triage and prioritization process for urgent surgical patients.

Ethics approval

The study was approved by the Research Ethics Committee of the Faculdade de Medicina de São José do Rio Preto (FAMERP)—Brazil, under Opinion No. 4,543,158, dated 02/17/2021. All participants signed the informed consent form (ICF), after being informed about the objectives, procedures, risks, and benefits of the research, as well as the guarantee of data confidentiality.

Acknowledgements

The authors extend their sincere gratitude to all individuals and institutions that contributed to the successful completion of this study. We are particularly indebted to the panel of experts who participated in the Delphi technique, providing invaluable insights and expertise. Our heartfelt appreciation goes to the Faculdade de Medicina de São José do Rio Preto (FAMERP) for their unwavering support throughout the research process. We acknowledge the administrative staff of FAMERP for their logistical assistance, and the Hospital de Base—FUNFARME for their collaboration and support.

Financial support and sponsorship

This study was funded under registration number 88882.464241/2019-01—FACULDADE DE MEDICINA DE SÃO JOSÉ DO RIO PRETO in the social demand program, master's modality of the Coordination for the Improvement of Higher Education Personnel of the Federal Government—Brazil.

Conflicts of interest

There are no conflicts of interest.

References

1. Houaiss A. Dicionário Houaiss da Língua Portuguesa. 1st ed. Rio de Janeiro, Brasil: Objetiva: Instituto Antônio Houaiss de Lexicografia; 2009.
2. Ribeiro B, de Souza JSM. A segurança do paciente no centro cirúrgico: Papel da equipe de enfermagem. Semin Ciências Biológicas e da Saúde 2022;43:27-38.
3. Dos Reis RG, Rodrigues MCS. Post-discharge surgical site infection: Occurrence and characterization of general surgery outpatients. Cogitare Enferm 2017;22:1-10.
4. Maia E, Marquez de Paula T. Sistematização da Assistência de Enfermagem Perioperatória: Percepções e registros dos enfermeiros de um centro cirúrgico. Heal Resid J HRJ 2023;4.
5. Ministério da Saúde (MS). Conceitos e deficinições em saúde. 1977.
6. Rios A, Iniesta-Sepúlveda M. Application of systematic reviews and meta-analyses in surgical clinical practice. Cir Esp (Engl Ed). 2024;102(6):297-299.
7. Silva IMCC da. Segurança e Qualidade nos Blocos Operatórios dos Hospitais Militares. 2017.
8. Lopes TMR, Machado AVA, da Silva AS, Santos T de JX dos, Raiol IF, de Miranda SA, *et al*. Atuação do enfermeiro na segurança

do paciente em centro cirúrgico: Revisão integrativa da literatura. Rev Eletrônica Acervo Saúde 2019;(26):e769.

9. Levati S, Bonetti L, Prandi C, Bianchi M. Psychometric evaluation of the complexity of care in the home care setting instrument (COMID) in Italian Language. Int J Older People Nurs 2024;19:e12637.
10. Possari JF. Centro Cirúrgico: Planejamento, organização e gestão. Saraiva Educação; 2009.
11. Fernandes AT. O desafio das infecções hospitalares: Visão histórica, a atualidade e perspectivas. In: Infecção Hospitalar e Suas Interfaces na Área da Saúde. São Paulo: Atheneu; 2000.
12. Associação Brasileira de Enfermeiros de Centro Cirúrgico R anestésica e C de M de esterilização (SOBECC). Práticas Recomendadas: Centro de Material e Esterilização, Centro Cirúrgico, Recuperação Pós-Anestésica. São Paulo: Manole; 2017.
13. Hickland P, Clements J, Convie L, McKay D, McElvanna K. Adapting an emergency general surgery service in response to the COVID-19 pandemic. Ann R Coll Surg Engl 2021;103:487-92.
14. Alonso N, Massenburg BB, Galli R, Sobrado L, Biroloni D. Surgery in Brazilian health care: Funding and physician distribution. Rev Col Bras Cir 2017;44:202-7.
15. Sacoman TM, Beltrammi DGM, Andrezza R, Cecílio LC de O, Reis AAC dos. Implantação do Sistema de Classificação de Risco Manchester em uma rede municipal de urgência. Saúde em Debate 2019;43:354-67.
16. Achanta A, Nordestgaard A, Kongkaewpaisan N, Han K, Mendoza A, Saillant N, *et al*. Most of the variation in length of stay in emergency general surgery is not related to clinical factors of patient care. J Trauma Acute Care Surg 2019;87:408-12.
17. Raymond BL, Wanderer JP, Hawkins AT, Geiger TM, Ehrenfeld JM, Stokes JW, *et al*. Use of the American College of Surgeons national surgical quality improvement program surgical risk calculator during preoperative risk discussion: The patient perspective. Anesth Analg 2019;128:643-50.
18. Toloui A, Kiah M, Zarrin AA, Azizi Y, Yousefifard M. Prognostic accuracy of emergency surgery score: A systematic review. Eur J Trauma Emerg Surg 2024;50:723-39.
19. Crawford TC, Magruder JT, Grimm JC, Suarez-Pierre A, Sciortino CM, Mandal K, *et al*. Complications after cardiac operations: All are not created equal. Ann Thorac Surg 2017;103:32-40.
20. Ramsay G, Haynes AB, Lipsitz SR, Solsky I, Leitch J, Gawande AA, *et al*. Reducing surgical mortality in Scotland by use of the WHO Surgical Safety Checklist. Br J Surg 2019;106:1005-11.
21. Prytherch DR, Briggs JS, Weaver PC, Schmidt P, Smith GB. Measuring clinical performance using routinely collected clinical data. Med Inform Internet Med 2005;30:151-6.
22. Nascimento RAM do, Assunção MSC, Silva Junior JM, Amendola CP, Carvalho TM de, Lima EQ, *et al*. Nurses' knowledge to identify early acute kidney injury. Rev da Esc Enferm da USP 2016;50:399-404.
23. Revorêdo L da S, Maia RS, Torres G de V, Maia EMC. O uso da técnica delphi em saúde: Uma revisão integrativa de estudos brasileiros. Arq Ciênc Saúde 2015;22:16-21.
24. Martin Bland J, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;327:307-10.
25. Santos EJF, Duarte C, da Silva JAP, Ferreira RJO. The impact of fatigue in rheumatoid arthritis and the challenges of its assessment. Rheumatology 2019;58(Supplement_5):v3-9.
26. Hernández-Nieto RA. Contributions to Statistical Analysis. Mérida Univ Los Andes; 2002.
27. Muthén LK, Muthén BO. Mplus: Statistical Analysis with Latent Variables. Los Angeles: Muthén and Muthén; 2010.
28. Asparouhov T, Muthén B. IRT in Mplus. 2016.
29. Zhang Z, Yuan KH. Practical Statistical Power Analysis using WebPower and R. ISDSA Press; 2018.
30. Byrne BM. Basic concepts, applications, and programming." Structural Equation Modeling with AMOS 396. New York: Routledge; 2010.
31. Eid M, Geiser C, Koch T, Heene M. Anomalous results in G-factor models: Explanations and alternatives. Psychol Methods 2017;22:541-62.
32. Kyriazos TA, Stalikas A. Applied psychometrics: The steps of scale development and standardization process. Psychology 2018;09:2531-60.
33. Samejima F. Estimation of latent ability using a response pattern of graded scores. APA Psyc Net 1969;34:100.
34. Baker FB. The Basics of Item Response Theory. 2001;186:34-63.
35. Batistini HC. Elaboração e validação de checklist de cuidados do enfermeiro ao paciente no pós-operatório imediatode cirurgia cardíaca. 2018.
36. Wolf L, Delao A, Perhats C, Baker K, Olson CM. Development of Nurse-sensitive, emergency department–specific quality indicators using a modified Delphi technique. J Nurs Care Qual 2022;37:E59-66.
37. Moloney E, O'Donovan MR, Carpenter CR, Salvi F, Dent E, Mooijaart S, *et al*. Core requirements of frailty screening in the emergency department: An international Delphi consensus study. Age Ageing 2024;53:afae013. doi: 10.1093/ageing/afae013.
38. Porcari TA, Cavalari PCF, Roscani ANCP, Kumakura ARSO, Gasparino RC. Safe surgeries: Elaboration and validation of a checklist for outpatient surgical procedures. Rev Gaucha Enferm 2020;41:e20190321. Portuguese, English.
39. Nasa P, Jain R, Juneja D. Delphi methodology in healthcare research: How to decide its appropriateness. World J Methodol 2021;11:116-29.
40. Barrios M, Guilera G, Nuño L, Gómez-Benito J. Consensus in the delphi method: What makes a decision change? Technol Forecast Soc Change 2021;163:120484.
41. Mota A, Castilho A. Construction and psychometric validation of the questionnaire on patient safety in the operating room. Rev Enferm Ref 2019;5:67-78.
42. Brouwers MC, Spithoff K, Kerkvliet K, Alonso-Coello P, Burgers J, Cluzeau F, *et al*. Development and validation of a tool to assess the quality of clinical practice guideline recommendations. JAMA Netw Open 2020;3:e205535.
43. Niederberger M, Spranger J. Delphi technique in health sciences: A map. Front Public Health 2020;8:457.
44. Boswell CL, Minteer SA, Herasevich S, Garcia-Mendez JP, Dong Y, Gajic O, *et al*. Early prevention of critical illness in older adults: Adaptation and pilot testing of an electronic risk score and checklist. J Prim Care Community Health 2024;15:e21501319241231238.
45. Cardell CF, Peters XD, Hu QL, Robinson-Gerace A, Mistretta S, Wescott AB, *et al*. Evidence review for the American College of surgeons quality verification part III: Standardization, protocols, and achieving better outcomes for patient care. J Am Coll Surg 2024;239:494-510.
46. Rahmani A, Vaziri Nezhad R, Ahmadi Nia H, Rezaeian M. Methodological principles and applications of the Delphi Method: A narrative review. J Rafsanjan Univ Med Sci 2020;19:515-38.
47. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR, Young SL. Best practices for developing and validating scales for health, social, and behavioral research: A primer. Front Public Health 2018;6:149-161.
48. Lancaster S, Leggio WJ, Ashford S, Carhart E, McKenna KD, Crowe RP. Defining priorities for emergency medical services education research: A modified Delphi study. J Am Coll Emerg Physicians Open 2023;4:e12882. doi: 10.1002/emp2.12882.
49. Azevedo NH, Scarpa DL. Decisões envolvidas na elaboração e validação de um questionário contextualizado sobre concepções de natureza da ciência. Investig em Ensino Ciências 2017;22:57.

50. Adeniyi AO, Arowoogun JO, Chidi R, Okolo CA, Babawarun O. The impact of electronic health records on patient care and outcomes: A comprehensive review. World J Adv Res Rev 2024;21:1446-55.
51. Vieira MA, Ohara CV, Domenico EB. The construction and validation of an instrument for the assessment of graduates of undergraduate nursing courses. Rev Lat Am Enfermagem 2016;24:e2710. doi: 10.1590/1518-8345.0834.2710.
52. Takara NC, Ferreira NDC, Murakami BM, Lopes CT. Development and validation of an informative manual on venous thromboembolism for the lay population. Einstein (São Paulo) 2020;18:eAO5425. doi: 10.31744/einstein_journal/2020AO5425.
53. Pedrosa KK de A, Oliveira SA de, Machado RC. Validation of a care protocol for the septic patient in the Intensive Care Unit. Rev Bras Enferm 2018;71:1106-14.
54. Martins PR, Fonseca LF, Rossetto EG. Developing and validating the perioperative thirst discomfort scale. Rev Esc Enferm USP 2017;51:e03240. English, Portuguese.
55. Lopes JL, Baptista RCN, Domingues TAM, Ohl RIB, Barros ALBL. Development and validation of a video on bed baths. Rev Lat Am Enfermagem 2020;28:e3329.
56. Revorêdo L da S, Dantas MMC, Maia RS, Torres G de V, Maia EMC. Validação de conteúdo de um instrumento para identificação de violência contra criança. Acta Paul Enferm 2016;29:205-17.
57. Cesconetto D, Gelbcke FL, Echevarría-Guanilo ME. Validação de conteúdo do instrumento de avaliação do grau de dependência dos usuários. Enferm em Foco 2020;10:1-16.
58. Gonçalves G de MO. Fatores de risco associados às infecções em cirurgias ortopédicas: Um estudo de revisão. 2022; 1:1-67.
59. Karpenko A, Keegan J. Diagnosis of coma. Emerg Med Clin North Am 2021;39:155-72.
60. Galetsi P, Katsaliaki K, Kumar S. Exploring benefits and ethical challenges in the rise of mHealth (mobile healthcare) technology for the common good: An analysis of mobile applications for health specialists. Technovation 2023;121:102598.
61. Loh B, Lim JA, Seah M, Khan W. Perioperative management of open fractures in the lower limb. J Perioper Pract 2022;32:100-7.
62. Oliveira RV, Cruz LP, Matos MA. Comparative accuracy assessment of the Gustilo and Tscherne classification systems as predictors of infection in open fractures. Rev Bras Ortop 2018;53:314-8.
63. Motta AS da, Versiani CM, Jamil LC, Rocha PPL, Fonseca HB, Almeida AC de M. Hematoma expansivo tardio após ritidoplastia. Rev Bras Cir Plástica – Braz J Plast Surg 2019;34:73-4.
64. Chaker SC, Hung YC, Saad M, Perdikis G, Grotting JC, Higdon KK. Complications and risks associated with the different types of abdominoplasties: An analysis of 55,956 patients. Aesthetic Surg J 2024;44:965-75.
65. Guimbard-Pérez JH, Camino-Willhuber G, Romero-Muñoz LM, Peral-Alarma M, Brocca ME, Barriga-Martín A. Efficacy of early versus delayed spinal cord decompression in neurological recovery after traumatic spinal cord injury: Systematic review and meta-analysis. Rev Esp Cir Ortop Traumatol. 2024;68(5):524-36.
66. Lima de SCCE, Correia JPC, de Oliveira Silva P, de Oliveira Silva T, Petter J. Diagnóstico ecográfico de Grangrena de Fournier; 2020.
67. da Silva VCR, Guerino MR, Rodrigues FTM, Alves KFP, Trigueiro LC de L, Ferreira AP de L, *et al*. Perfil de idosos atendidos na fisioterapia em universidade federal entre 2009-2019: Estudo retrospectivo. Saúde e Pesqui 2021;14:1-12.
68. Gattinoni L, Busana M, Giosa L, Macrì M, Quintel M. Prone positioning in acute respiratory distress syndrome. Semin Respir Crit Care Med 2019;40:94-100.
69. Guérin C. Prone positioning acute respiratory distress syndrome patients. Ann Transl Med 2017;5:289.
70. Hurndall KH, Merriman H, Leatherby R, Glasgow S, Davenport R. Vascular interventions and surgery in trauma audit (VISTA). BJS Open 2023;7:ezrad012. doi: 10.1093/bjsopen/zrad012.
71. Moraes MR de S, Baptista-Silva JCC, Burihan E, Figueiredo LFP de. Trauma vascular. In: Maffei FH de A, Yoshida WB, Rollo HA, Moura R, Sobreira ML, Giannini M, *et al*., editors. Doenças Vasculares Periféricas. Rio de Janeiro, Brasil: Editora Guanabara Koogan; 2015;4:50-78.
72. Efthymiou A, Kalaitzaki A, Kondilis B, Rovithis M. Health literacy continuing education courses and tools for healthcare professionals: A scoping review. Gerontol Geriatr Educ 2024;45:212-47.