Validation of instrument to assess software of patients’ risk classification

Validação de instrumento para avaliar software de classificação de risco de pacientes

Validación de instrumento para evaluar el programa (software) de clasificación de riesgo de pacientes

Leila Garcia de Oliveira Pegoraro¹, Raquel Gvozd¹, Maria do Carmo Fernandez Lourenço Haddad¹, Marli Terezinha Oliveira Vannuchi¹, Larissa Gutierrez de Carvalho Silva¹, Mariana Angela Rossaneis¹

¹ Universidade Estadual de Londrina, Health Sciences Center. Londrina, Paraná, Brazil.

How to cite this article:

Submission: 03-28-2017 Approval: 05-21-2017

ABSTRACT

Objective: Validate the content of an instrument that assesses the quality of a software applied to the risk classification of patients. Method: Methodological study, conducted in three stages: adaptation of the instrument, content validation through Delphi technique and pre-test. The results were analyzed through Content Validity Index, Overall Content Validity Index and Inter-rater Reliability. Results: The final version of the instrument comprises 8 characteristics and 28 sub-characteristics, being 37 general questions to computer experts and nurse and 7 specific questions to computer experts, including 1 question and excluding 3 questions of the original instrument. We obtained Overall Content Validity of 92% and Inter-rater Reliability of 100% in the second Delphi round. Final considerations: The instrument has content validity, allowing to assess the technical quality and functional performance of the software applied to the risk classification of patients.

Descriptors: Classification; Emergency Medical Services; Validation Studies; Validation of Computer Programs; Healthcare Quality.

RESUMO

Objetivo: Validar o conteúdo de instrumento que avalia a qualidade de um software aplicado à classificação de risco de pacientes. Método: Estudo metodológico, realizado em três fases: adaptação do instrumento, validação de conteúdo por meio da técnica Delphi e pré-teste. Os resultados foram analisados por meio do Índice de Validez de Conteúdo, Índice de Validez de Conteúdo global e Índice de Concordância Interevaluadores. Resultados: A versão final do instrumento contempla oito características, 28 subcaracterísticas, sendo 37 questões gerais aos especialistas em informática e enfermeiro e sete questões específicas aos especialistas em informática, com inclusão de uma questão e exclusão de três questões do instrumento original. Obteve-se percentual de Validez de Conteúdo Global de 92% e Índice de Concordância Interevaluadores de 100% na segunda rodada Delphi. Considerações finais: O instrumento possui validade de conteúdo permitindo avaliar a qualidade técnica e desempenho funcional de software aplicado à classificação de risco de pacientes.

Descritores: Classificação; Serviços Médicos de Emergências; Estudos de Validação; Validação de Programas de Computador; Qualidade da Assistência à Saúde.

RESUMEN

Objetivo: Validar el contenido de instrumento que evalúa la cualidad de un programa (software) aplicado a la clasificación de riesgo de pacientes. Método: Estudio metodológico, realizado en tres etapas: la adaptación del instrumento, la validación de contenido por medio de la técnica Delphi y la prueba previa. Los resultados fueron analizados por medio del Índice de Validez de Contenido, Índice de Validez de Contenido Global e Índice de Concordancia Interevaluadores. Resultados: La versión final del instrumento contempla ocho características, 28 subcaracterísticas, siendo 37 cuestiones generales a los expertos en informática y enfermero y siete cuestiones específicas a los expertos en informática, con inclusión de una cuestión y exclusión de tres cuestiones del instrumento original. Se obtuvo porcentual de Validez de Contenido Global del 92% e Índice de Concordancia Interevaluadores del 100% en la segunda ronda.
INTRODUCTION

The quality of health information is a global concern. Countries such as the United States of America, Canada and England promote initiatives for the safety in the conception, acquisition and deployment of health information technology. Advances in access to information, increased accuracy of documentation, deployment of evidence-based practice, cost reduction with return on investment, as well as improvements in quality of care and satisfaction of employees, are evidence of international experiences in the use of computerization in health.

In Brazil, the use of computer systems to management has increased exponentially in different areas. In the work process of nursing, computer science improves more and more through the development and evaluation of tools, processes and structures that assist nurses in care management. In emergency hospital services, computerization has been used for the Assessment with Risk Classification of patients. International literature brings two Assessment and Risk Classification protocols in electronic format: the Manchester Triage System and the Andorá Triage Model. In Brazil, Manchester is the most known protocol, which has inspired the computerization of the National Protocol of Risk Assessment and Classification, implemented by the Ministry of Health through the ordinance 2048/2002. However, no scientific evidences were found on the evaluation of technical quality and functional performance of software used for this computerization, both of national and international protocol.

The assessment of a software is critical to identify the weaknesses and limitations of the product, analyze its performance and diagnose the need for adjustments. This assessment is consolidated by specific regulations, such as the International Organization for Standardization (ISO), International Electrotechnical Commission (IEC) and Brazilian Association of Technical Standards (ABNT), which proposed two Brazilian standards (NBR), ISO/IEC 14598 and ISO/IEC 9126, regarding the quality of software products. In 2011, updates of these regulations gave rise to the standards ISO/IEC 25010 – System and Software engineering – System and software quality models; and ISO/IEC 25040 – System and Software engineering – Evaluation process. Such standards have been used to create software and its assessment in different contexts of Brazilian nursing.

In this sense, we identified on our work practice the need to assess the technical quality and functional performance of a software used to perform the Risk Assessment and Classification of patients in a Hospital Emergency Service (HES). To do so, our study was based on a recent survey that applied such methodology in an Electronic Document System of Nursing Processes (PROCenf-USP) through an instrument that contemplated the last resolution specific to this type of assessment.

OBJECTIVE

To validate the content of an instrument that assesses the quality of a software applied to the risk classification of patients.

METHOD

Ethical aspects

This study was approved in April 27, 2016 by the Research Ethics Committee of the Londrina State University (UEL).

Study design, location, and period

This is a methodological and applied study, with a quantitative approach, carried out from April to October 2016, consolidated in three stages, namely: instrument adaptation, validation of content and pre-test. The study was held in a medium-sized public secondary hospital, located in the State of Paraná, Brazil.

Study protocol

For validation of this instrument we opted for the Delphi technique, a systematized method of judgment of information intended to reach the consensus of opinions about a particular subject, regarding the knowledge of a committee of experts, through validations articulated in rounds, maintaining anonymity. Content validation is held by questioning various experts who analyze the representativeness of the items that integrate the object of analysis.

The instrument chosen in its original version aims to assess the software quality of the Electronic Document System of Nursing Processes PROCenf-USP, which includes the latest update from the International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC), specific to the assessment of the software.

The original instrument contains 8 characteristics and 31 sub-characteristics directed to computer experts and nurses, they are: functional adequacy (functional integrity, functional correction and functional fitness); performance efficiency (time, resources, and capacity); compatibility (coexistence and interoperability); usability (recognition of suitability, learnability, error protection, operability, user interface aesthetics and accessibility); reliability (maturity, fault tolerance, recoverability and availability); security (confidentiality, integrity, non-repudiation, accountability and authentication); maintainability: (analyzability, changeability, modularity, reusability, testability); portability (adaptability, installability and replaceability). These two last characteristics were assessed only by computer experts.
The author of this instrument granted permission for its adaptation and validation to assess the software of Risk Classification of patients.

The first stage of the study was the adaptation of the original instrument to the context of Assessment with Risk Classification of patients. In the second stage, we proceed to content validation of the instrument adapted through Delphi technique. Two Delphi rounds were needed. In the first, the committee was composed of 6 experts; of these, 3 computer experts with experience in creating software and computerization in health area and 3 nurses, 2 doctors and 1 master, both with experience in validation and adaptation of instruments. The same experts were invited for the second round, but only 5 responded to the request.

After prior contact with the experts by e-mail, the following documents were sent: an Informed Consent Form in electronic format, a summarized protocol for the presentation of the research project, the instrument adapted and instructions for completion and assessment of the questions to be reviewed by them. It was stipulated a deadline of 15 days for the return of the analyzed material.

Experts assessed each item of the instrument in +1 (suitable), -1 (suitable with amendments) or 0 (unsuitable). At the end of the assessment, it was possible to calculate a Content Validity Index (CVI) and the Overall Content Validity Index, which measure the proportion or percentage of evaluators who are in agreement on certain aspects of the instrument and its items, being considered valid the ones that obtained minimum percentage of 80% of agreement between the experts.[17-18]

The Reliability Index or Inter-rater Reliability, which is intended to assess the extent to which the evaluators are reliable when assessing the items outside the context studied, was calculated for each dimension of the instrument, by dividing the number of items that scored above 80% of agreement between evaluators by the total number of items of each dimension[17-18].

After content validation, the third stage of the study was held by pre-testing the adapted instrument. The selection of individuals for this stage occurred through purposive sampling at the institution studied, being composed by 1 computer expert, 3 nurses with experience in computerized Assessment of Classification Risk of patients and 3 nurses without experience in this activity. In this stage, participants were asked about problems regarding the understanding of each item and terms difficult to be understood; they assessed the questions through questionnaire of Practicality of Instrument[19].

**Analysis of results and statistics**

Data were tabulated in the software Microsoft Excel 2010 and analyzed from reflective reading and descriptive statistics.

**RESULTS**

In the process of adaptation of the instrument, researchers chose to combine some questions that complemented each other to better understand the content. Sub-characteristics “Adaptability” and “Installability” were grouped into a single question. “User Interface Aesthetics,” “Time,” and “Integrity” had questions grouped each one within its own sub-characteristic.

The content validation procedure consolidated the final version of the instrument with 8 characteristics, 28 sub-characteristics and 44 questions. The characteristics Functional Adequacy, Reliability, Usability, Performance Efficiency, Compatibility, and Safety appear in 37 questions and are aimed at computer experts and nurses. The characteristics Maintainability and Portability are included in 7 specific questions to computer experts. One question was included and 3 were excluded of the original instrument. From the first Delphi round, we identified the need to reformulate some questions of the instrument, justifying the second Delphi round.

Table 1 presents the Content Validity Index (CVI) of each characteristic and sub-characteristic of the instrument, the Inter-rater Reliability (IRR) and the Overall Content Validity Index in the first and second Delphi rounds, Londrina, Paraná, Brazil, 2016.

<table>
<thead>
<tr>
<th>Characteristics and sub-characteristics</th>
<th>CVI (%)**</th>
<th>IRR (%)***</th>
<th>Overall CVI (%)***</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1st Delphi round</td>
<td>2st Delphi round</td>
<td>1st Delphi round</td>
</tr>
<tr>
<td>Functional suitability</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Functional completeness</td>
<td>74</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>Functional correctness</td>
<td>50</td>
<td>100</td>
<td>67</td>
</tr>
<tr>
<td>Functional appropriateness</td>
<td>89</td>
<td>87</td>
<td></td>
</tr>
<tr>
<td>Reliability</td>
<td>83</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Maturity</td>
<td>62</td>
<td>90</td>
<td>70</td>
</tr>
<tr>
<td>Fault tolerance</td>
<td>83</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Recoverability</td>
<td>50</td>
<td>80</td>
<td>50</td>
</tr>
<tr>
<td>Availability</td>
<td>83</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Availability</td>
<td>33</td>
<td>80</td>
<td></td>
</tr>
</tbody>
</table>

To be continued
According to the assessment of experts regarding the agreement and representativeness of the items in the instrument, the overall CVI in first Delphi round was 70%, reaching 93% in the second round.

The characteristics Usability and Portability obtained the best percentages, with CVI above 80% in the first Delphi round. Safety, Reliability, Compatibility and Maintainability presented the lowest percentages of CVI (53%, 62%, 63% and 63%, respectively).

Regarding reliability, just the Portability characteristic reached satisfactory index in the first Delphi round. All the others reached IRR of 100% in the second round.

Of the 30 sub-characteristics containing the instrument in the first round, 15 (50%) reached CVI above 80%, being 5 with 100% of compliance. Availability, Accountability, and Authentication obtained the lowest percentages of CVI (0%, 0% and 33%, respectively). In the second round, the Accountability and Authentication sub-characteristics were excluded from the instrument on the recommendation of the experts.

In Chart 1 are presented the suggestions of the committee of experts on the items considered “suitable with amendment” and “unsuitable,” obtained in the first and second Delphi round.

After content validation, the instrument was submitted to a pre-test conducted with 7 participants: 1 computer expert, 3 nurses with experience in computerized Assessment of Classification Risk of patients and 3 nurses without experience in this activity. The results of this stage showed ease of understanding of the items of the instrument and pertinence of the proposal of software assessment; there are no suggestions for adjustments in the items regarding the instrument. Two participants suggested replacing the words Disagree and Agree by Yes and No expressions, but the researchers considered this change inappropriate. The time demanded by participants for completing the instrument was approximately 18 minutes to each.
DISCUSSION

The option of adapting and validating an instrument to assess the computerization of the Assessment with Risk Classification in Hospital Emergency Services (HES), using a tool that indicates the level of technical quality and functional performance, was based on the fact that, so far, no instrument was able to quantify the opinions and impressions of health professionals regarding technical quality and functional performance of this computerized instrument.

The process of adaptation of the instrument was consolidated through scientific basis on the assessment of 

### Chart 1 – Presentation of the suggestions of the committee of experts on the items considered “suitable with amendment” and “unsuitable,” Londrina, Paraná, Brazil, 2016

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Sub-characteristic</th>
<th>Analysis and suggestion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Functional suitability</td>
<td>Functional completeness</td>
<td>Rephrase the writing, standardize non-inducing verbs and replace the term “system” by “software.”</td>
</tr>
<tr>
<td>Functional appropriateness</td>
<td></td>
<td>Considered unsuitable by an expert.</td>
</tr>
<tr>
<td>Reliability</td>
<td>Fault tolerance</td>
<td>Question 8: replace the term “system” by “the same.”</td>
</tr>
<tr>
<td></td>
<td>Availability</td>
<td>Question 10: complete the question with the term “when necessary.”</td>
</tr>
<tr>
<td></td>
<td>Maturity</td>
<td>Considered unsuitable by an expert.</td>
</tr>
<tr>
<td></td>
<td>Fault tolerance</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Recoverability</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Availability</td>
<td></td>
</tr>
<tr>
<td>Usability</td>
<td>Recognition of suitability</td>
<td>Show the acronym AACR*, described at the beginning of each characteristic, to remind the reader of its meaning.</td>
</tr>
<tr>
<td></td>
<td>Learnability</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Operability</td>
<td></td>
</tr>
<tr>
<td></td>
<td>User Interface Aesthetics</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Functional completeness</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Functional correctness</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Functional appropriateness</td>
<td></td>
</tr>
<tr>
<td>Performance efficiency</td>
<td>Resource</td>
<td>Add what types of external resources are required to use the software.</td>
</tr>
<tr>
<td></td>
<td>Capacity</td>
<td>Considered unsuitable by an expert.</td>
</tr>
<tr>
<td>Compatibility</td>
<td>Interoperability</td>
<td>Question 30: rephrase the writing by standardizing the verbs. Question 31: inform with which databases it is possible to exchange information.</td>
</tr>
<tr>
<td></td>
<td>Coexistence</td>
<td>Question 32: replace the term “shared environment” by “when other screens are being simultaneously processed.” Considered unsuitable by an expert.</td>
</tr>
<tr>
<td>Safety</td>
<td>Reliability</td>
<td>Add the term “individualized and reliable user and password authentication.”</td>
</tr>
<tr>
<td></td>
<td>Accountability</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Authenticity</td>
<td>Items considered inappropriate by six experts, resulting in their removal.</td>
</tr>
<tr>
<td>Maintainability</td>
<td>Analyzability</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Changeability</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Testability</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Modularity</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reusability</td>
<td>Rephrase the writing using the term “software,” because its absence causes problems regarding clarity and transmit incomplete ideas.</td>
</tr>
<tr>
<td>Portability</td>
<td>Adaptability and installability</td>
<td>Question 43: replace the term “systems” by “operating environments.”</td>
</tr>
<tr>
<td></td>
<td>Replaceability</td>
<td>Question 44: Add in the question the possibility of updating the software.</td>
</tr>
</tbody>
</table>

Note: *AACR – Assessment with User Embracement and Risk Classification.
technical quality and functional performance of software, Assessment with Risk Classification of patients, computerization in health and methodological benchmark of validation of instruments[15-7,10,20].

In this stage, we decided to group the questions of the sub-characteristics Adaptability and Installability, presenting them in a single question. The sub-characteristics “Time” and “Integrity” had their questions grouped, each one within its own sub-characteristics. Such changes do not generate questions from the experts, nor concerns regarding clarity and representativeness, reaching a CVI above 80% since the first Delphi round.

However, the sub-characteristic “Interface Aesthetics,” which also had its questions grouped to better research, presented CVI of 67% in the first round. The experts assessed it as “suitable with amendments” and suggested replacing the term “graphic design” by “software design” and adding which characteristics must be assessed. Such suggestions were accepted, reaching a CVI of 80% in the second round.

In the adaptation of the instrument, through the analysis of the results and reflective reading of the suggestions and comments of the committee of experts, it was possible to verify problems regarding clarity and representativeness in the instrument. Considering the method used in this research, which required IRR of 100% and a minimum CVI of 80%, the results obtained in the first Delphi round showed the need for adjustments of the instrument.

In the first Delphi round, 50% of the sub-characteristics that compose the instrument did not achieve the minimum CVI recommended by the literature and adopted in this research[17-18]. Adjustments were made on the questions of the items indicated by the experts, through reformulations of writing on the questions, such as inversions of words or replacement of items by synonyms, which allowed better understanding of the items[21]. After this stage, we proceed to the second Delphi round to ensure that the adjustments made would obtain the clarity and agreement required.

In the first Delphi round, an expert suggested adding on the “Recognition of Suitability” sub-characteristic, of the Usability characteristic, a question related to the occurrence of training for the use of the software. The suggestion of the experts was accepted, since it was considered relevant within the context of the use and assessment of technical quality and functional performance; the other questions of the sub-characteristic received suggestion of reformulations and standardization of writing, which were accepted and submitted to a second round of assessment by the committee, obtaining CVI compliance percentage > 80%.

In this context, we consider that the incorporation of new technologies in the work process requires skills that must be acquired by health professionals, to have success and efficiency in the use of these tools. In this sense, training courses are important strategies, a praxis for knowledge production and success in the appropriate use of technologies, with direct impact on the quality of care provided[22-24]. The failure of the training courses or their absence may influence on the ability and desire of professionals to engage in activities of development, applicability and transformation of practices.

Still on the content analysis of the instrument, the characteristics Functional Fitness, Maintainability, Recoverability, Learnability, Capacity and Testability achieved percentage ≥ 80% in the first Delphi round, but were respectively assessed as “unsuitable” by an expert and “suitable” by the others; although there is no suggestion of reformulation and reaching favorable percentage of CVI, they were forwarded to the second round and reached percentage of 100%.

However, the “Fault tolerance,” “Availability,” and “Coexistence” sub-characteristics did not reach percentage of 80% or more in the first Delphi round, and were respectively assessed as “unsuitable” by an expert and “suitable” by the others. Changes were made on the items indicated, considering their relevance. They were forwarded to the second Delphi round, reaching favorable percentage of CVI, mostly 100% of agreement; and considered as valid, trustworthy and representative content, as advocated by the reference adopted in this study[15-19].

On the other hand, the “Accountability” and “Authentica- tion” sub-characteristics were removed from the instrument in the first Delphi round, since both were assessed as “unsuitable” by six experts, being observed problems of redundancy that can influence the results, according to the committee, justified as contemplated in others sub-characteristics of the instrument.

In the end, the indexes from the content validation process indicated high reliability and trustworthiness of the instrument and made it possible to assess the technical quality and functional performance of software applied to computerized risk classification of patients in EHS, a tool that helps professionals in the health care and that can be applied to several processes of computerization in health.

Study limitations
We considered as limitations of this study the number of experts in computer science that would contemplate the inclusion criteria established, i.e., with experience in computerization in health, given the shortage of these professionals on the staff of most health services. Health Informatics in Brazil is an area in exponential growth that has attracted professionals from diverse academic segments in search of specializations that confer competence to act. However, we found a restriction of specialization courses in this area and, among these, there are differences in their course syllabus, corroborating the limitations of our study[25].

Contributions to the Nursing field
The validated instrument brings contributes to the health sector in the acquisition of a tool that aims to assess the efficiency of software, qualify, promote improvements in its development and construction through health institutions, considering that this instrument can be continuously optimized and adapted to other computerized work processes. We emphasize the development of technical and scientific skills to healthcare professionals, specifically in the nursing field, with incentives to an entrepreneurial attitude, which is still found
empirically in many hospitals. These professionals are assisted
by the exponential growth of technologies whose purpose is
to meet their daily needs in different contexts of their practice,
facilitating the work time and improving the care provided.

**FINAL CONSIDERATIONS**

The content validation procedure consolidated the final ver-
sion of the instrument with 8 characteristics, 28 sub-character-
istics and 44 questions. The Functional Adequacy, Reliability,
Usability, Performance Efficiency, Compatibility, and Safety char-
acteristics appear in 37 questions and are aimed at computer
experts and nurses. The Maintainability and Portability charac-
teristics are included in 7 specific questions to computer experts.

The overall CVI in the first Delphi round was 70%, reaching
93% in the second round. Of the 30 sub-characteristics
containing the instrument in the first round, 15 (50%) reached
CVI > 80%, being 5 with 100% of compliance. Availability,
Accountability and Authenticity obtained the lowest percentages of CVI (0%, 0% and 33%, respectively). In the second round, the “Accountability” and “Authorization” sub-char-
acteristics were excluded from the instrument on the rec-
ommendation of the experts.

**REFERENCES**

1. Kushniruk AW, Bates DW, Bainbridge M, Househ MS, Borycki EM. National efforts to improve health information system safety in


3. Oliveira NB, Peres HHC. Evaluation of the functional performance and technical quality of an Electronic Documentation System of the
pdf/rlae/v23n2/0104-1169-rlae-2302-2548.pdf

4. Gonçalves MFS, David G. Planejamento e realização de estudo de (re)utilização da informação clínica em contexto hospitalar com
prisramatico/article/view/3102/pdf_40

5. Jiménez GJ. Clasificación de pacientes en los servicios de urgencias y emergencias: hacia un modelo de triaje estructurado de urgen-

6. Conselho Regional de Enfermagem de Santa Catarina. PARECER Nº 009/CT/2015/PT Autarquia Federal criada pela Lei Nº 5.905/73.
uploads/2015/07/Parecer-009-2015-Acolhimento-com-Classificacao%C3%A7ao-de-Risco-CT-Altac-3A%C3%A9M%20Alta-Complexidade.pdf

7. Sperandio DJ. A tecnologia computacional móvel na sistematização da assistência de enfermagem: avaliação de um software-protótipo
[Tese] [Internet]. Ribeirão Preto: Escola de Enfermagem de Ribeirão Preto da Universidade de São Paulo; 2008. [cited 2015 Sep 13].


10. ISO/IEC 25010 – System and Software engineering - System and software Quality Requirements and Evaluation (SQuaRE) - System and

11. ISO/IEC 25040 - System and Software engineering - System and software Quality Requirements and Evaluation (SQuaRE) – Evaluation
Validation of instrument to assess software of patients’ risk classification

Pegoraro LGO, Gvozd R, Haddad MCF, Vannuchi MTO, Silva LGC, Rossaneis MA.


