

RESEARCH PAPER

Validation of the International Resident Assessment Instrument - Acute Care (InterRAI-AC) for the Portuguese population

Validação do *International Resident Assessment Instrument – Acute Care* (InterRAI-AC) para a população portuguesa
Validación del *International Resident Assessment Instrument – Acute Care* (InterRAI-AC) para la población portuguesa

António Fernando Salgueiro Amaral*; Pedro Lopes Ferreira**; Leonard C. Gray***

Abstract

Context: In an international context of cost reduction, it is necessary to measure nursing outcomes with reliable and valid instruments to produce indicators of effectiveness, improve the quality of care and ensure accountability. **Aim:** To validate the International Resident Assessment Instrument - Acute Care (InterRAI-AC) for the Portuguese population. **Methodology:** Longitudinal, descriptive and correlational study. Data were collected between March and July 2012 by nurses of 26 medical and surgical services of four hospitals in the Central Region of Portugal. The sample was composed of 1764 patients hospitalized for a period equal to or greater than three days. All ethical procedures were observed. **Results:** The instrument has proved to be valid and reliable, and can be used to obtain data for different purposes: quality indicators; care planning; evidence base for decision-making, among others. **Conclusion:** The main barrier identified was the time required for its completion, but it proved to be a key tool in the process of collecting information on patients.

Keywords: validation studies; nursing care; acute disease.

Resumo

Num contexto internacional de redução de custos, torna-se necessário medir os resultados de enfermagem com instrumentos fiáveis e válidos, para produzir indicadores de efetividade, melhorar a qualidade dos cuidados e garantir a prestação de contas.

Estudo longitudinal, descritivo e correlacional, cujo objetivo foi validar para a população portuguesa o *International Resident Assessment Instrument – Acute Care* (InterRAI-AC). Os dados foram recolhidos entre março e julho de 2012 pelos enfermeiros de 26 serviços médicos e cirúrgicos de quatro hospitais da região centro de Portugal, tendo-se constituído uma amostra de 1764 doentes internados por um período igual ou superior a 3 dias. Foram garantidos todos os procedimentos éticos.

O instrumento revelou-se válido e fiável e pode ser utilizado para a obtenção de dados com diferentes fins: indicadores de qualidade; planeamento dos cuidados; base de evidência para a tomada de decisão, entre outros. O principal obstáculo identificado foi o tempo necessário para o seu preenchimento, mas revelou-se uma ferramenta fundamental no processo de recolha de informação sobre os doentes.

Palavras-chave: estudos de validação; cuidados de enfermagem; doença aguda.

Resumen

En un contexto internacional de reducción de costos es necesario medir los resultados de enfermería con instrumentos fiables y válidos para producir indicadores de eficacia, mejorar la calidad de la atención y garantizar la rendición de cuentas.

Estudio longitudinal, descriptivo y correlacional cuyo objetivo fue validar para la población portuguesa el *International Resident Assessment Instrument – Acute Care* (InterRAI-AC). La recolección de datos se llevó a cabo de marzo a julio de 2012 por los enfermeros de 26 servicios médicos y quirúrgicos de cuatro hospitales del centro de Portugal. La muestra estuvo formada por 1.764 pacientes hospitalizados durante un período igual o superior a 3 días. Se garantizaron todos los procedimientos éticos.

Se demostró que el instrumento es válido y fiable, y se puede utilizar para obtener datos para diferentes propósitos: indicadores de calidad, planificación de la atención; base de evidencia para tomar decisiones, entre otros. El principal obstáculo identificado fue el tiempo necesario para realizarlo, sin embargo resultó ser una herramienta clave en el proceso de recolección de información sobre los pacientes.

Palabras clave: estudios de validación; atención de enfermería; enfermedad aguda.

* Coordinating Professor of the Nursing School of Coimbra
Master in Management and Economy in Health. Address: Escola Superior de Enfermagem de Coimbra Av. Bissaya Barreto 3001 - 091 Coimbra, Portugal [amaral@esenfc.pt].

** Associate Professor with Aggregation at the Faculty of Economics of the University of Coimbra. Ph.D. in Industrial Engineering – Theory of Decision/ Health Systems, University of Coimbra Centre for Health Studies and Research, Faculty of Economics of the University of Coimbra, 165 3004-512 Coimbra, Portugal [pedrolf@fe.uc.pt].

*** MD, PhD, Professor in Geriatric Medicine, The University of Queensland, Level 3, Foundation Building Royal Children's Hospital, Herston, 4029. Australia [len.gray@uq.edu.au].

Received for publication: 20.11.13

Accepted for publication: 10.02.14

Introduction

The measurement of health outcomes as a way to produce indicators of effectiveness, to improve the quality of care and to guarantee the accountability of healthcare providers and units has become an essential practice in the different settings where healthcare are provided, and is closely related to the funding processes of organizations and to an international trend of cost reduction. This is evident in nursing care since services provided by nurses represent a significant fraction of healthcare and managers require the justification of the associated costs and, above all, their connection to patient outcomes (Donabedian, 1992).

Hence, the measurement of these outcomes must be a common goal which unites the interests and the activities of all stakeholders, for it has professional, financial and political effects. It is, therefore, necessary to design reliable and valid measuring instruments (Porter, 2010).

In order to analyse the effect of nursing care in obtaining the best patient outcomes, an assessment of nursing-sensitive outcomes is necessary. These are considered by Doran, Sidani, Watt-Watson, Laschinger, and Hall (2001) as a status, a behaviour or a perception of a patient or family, measurable in a continuum, which occurs as a response to a nursing intervention.

Health organizations and systems have been facing major challenges due to population ageing and associated increase in chronic-degenerative diseases, with ever more limited resources. According to Statistics Portugal (2013), the Ageing Index rose from 103 to 128 between 2001 and 2011. This represents a challenge because elderly individuals, as well as persons with disability, are more vulnerable to health problems which require multidisciplinary approaches to care (Hirdes et al., 2008). The frailty of the person and the complexity of the health status increase the risk of hospitalization-related adverse outcomes, such as falls and pressure ulcers. A range of socioeconomic costs include increased length of hospital stay and expenditure, higher mortality rate, decline after discharge, high rates of hospital readmission, the loss of ability to perform activities of daily living, among others (Brand et al., 2011).

The complexity and multidimensionality of the situations require the use of comprehensive

instruments which measure more than one dimension of the person and, especially, reflect a holistic vision of care (Hirdes et al., 2008). However, its use poses multiples challenges at the moment of validation, since the usual tests included in instruments which only measure one dimension of the person cannot be applied in the same way (Wellens, Milisen, Flamaing, & Moons, 2010).

Validation and reliability

Instrument validation processes have been developed in line with the evolution of the concept of validity, which went from a static property of a measure, in which a test was valid or not according to the existing correlation between the measure and another external criterion measure, to another perspective in which validity was categorized into specific types: content validity, construct validity and criterion-related validity (which included concurrent and predictive validities). Later on, a multi-method approach, introducing convergent and discriminant types of validity as part of construct validity, was presented. In the 1980s, a new meaning of validity, led by Cronbach (1980) and Messick (1980), emphasized the inferences and decisions made from test scores. This process led to the standardization of psychological tests, which, in turn, led to a definition of validity as the appropriateness, meaningfulness, and usefulness of specific inferences made from test scores, and this accumulation of evidence provides a sound scientific basis for score interpretations (Goodwin & Leech, 2003). Hence, the most recent standards for instrument validation indicate five types of validity evidence: evidence based on text content; evidence based on response processes; evidence based on internal structure; evidence based on relations to other variables; and evidence based on the consequences of testing (Goodwin & Leech, 2003).

Regarding reliability, which refers to the capacity of a given instrument to consistently and reliably measure any population attribute, regardless of the time of measurement and the person who measures it, comprehensive instruments also pose several challenges (Wellens, Milisen, Flamaing, & Moons, 2011; Streiner & Norman, 2008). Thus, the authors advocate the need to measure reliability evidence based on the equivalence of the measures undertaken by two different observers, evidence based on stability

and evidence based on internal consistency. Evidence based on reliability is very important to check whether a measurement instrument focuses on a single idea and whether the variables which compose the scale are internally consistent (Streiner & Norman, 2008).

InterRAI AC

InterRAI is an international research group which develops comprehensive assessment instruments oriented towards persons who need care, especially elderly people (InterRAI, 2012). Resident Assessment Instruments (RAI) are used internationally in various settings and with different purposes, such as care planning and measurement of quality outcomes and indicators (Glenny & Stolee, 2009).

There are currently 18 InterRAI assessment instruments for different care provision settings (palliative care, home care, nursing home care, post-acute care, community mental health care, among others). The InterRAI AC, introduced in 2006, was developed to be used in acute care hospitals, where people can benefit from a comprehensive assessment carried out by various specialists (Gray et al. 2008).

The InterRAI AC instrument was developed to be applied to persons with functional limitations in their personal activities, such as Activities of Daily Living (ADL) and Instrumental Activities of Daily Living (IADL); persons with cognitive impairment; persons with diseases together with loss of functional capacity in the performance of personal activities; or persons aged over 75 years as they are more at risk of developing these problems. It also provides holistic information on the complexity and frailty of inpatients (Gray et al., 2006, 2008).

InterRAI instruments are composed by a set of items which are common to all instruments (about 70; e.g. ADL and Pain); about 100 items are present in almost all instruments and some are specific for special settings and vary according to the instruments (Wellens et al., 2012).

InterRAI AC consists of several dimensions which assess different clinical domains, at four moments: Preadmission, i.e. within a three-day period prior to the onset of the situation which precipitated admission (in this case the informants may be relatives if patients are not able to do it); Admission, i.e. within the 24 hours following admission; on the 14th day of hospitalization; and at discharge, in which the assessment relates to the 24 hours prior to discharge. Hence, the specificity of this instrument in what

concerns its comprehensiveness and structure is the possibility to assess the hospitalization as a whole. The assessment of the situation prior to the disease allows not only the establishment of a connection with the current status, but it may also be used as a reference for rehabilitation and treatment. Furthermore, the measurement of outcomes requires a reference point, and, according to Sidanni, Doran and Mitchell (2004), the status of the patient prior to the situation which led him/her to the hospital influences the outcomes of care. Since this measure is essentially directed to supporting clinical practice, it is not possible to reach a final score after total completion of the instrument. The scores obtained in each of the scales which reflect disease severity and its evolution in selected domains can, nevertheless, be calculated (Wellens et al., 2010; Hirdes et al., 2008).

The interRAI Acute Care - Portuguese version (InterRAI AC-PT) consists of 16 sections: Identification; Intake and Initial History; Assessment Dates; Cognition; Communication, Vision and Hearing; Mood and Behaviour; Functional Status; Continence; Diagnoses; Health Conditions; Nutritional and Oral Status; Skin Condition; Treatments and Procedures; Discharge Potential; Discharge; Assessment Information.

From the sections of the instrument, and according to the User Manual, it is possible to produce a set of scales through item-mix summations and algorithms which measure: Cognitive Performance; Communication; Pain; IADL; Delirium; ADL Hierarchy scale; short ADL scale; Depression Rating Scale (short version); Body Mass Index.

The Cognitive Performance Scale (CPS) is based on an algorithm which combines the variables “cognitive skills for daily decision making, “memory/recall – good short term memory”, “memory/recall – good procedural memory”, “making self understood” and “ADL self-performance - eating”. This algorithm generates a set of 6 performance levels, from level 0 (intact) to level 6 (very severe impairment).

The Communication Scale (COMUNIC) assesses the persons’ ability to communicate and is based on the variables “making self understood” and “ability to understand others”. The summation of these two variables generates a set of 8 scores for the ability to communicate, from 0 (independent) to 8 (very severe inability).

The Pain Scale (DOR) assesses the pain mentioned by the person and is based on an algorithm which

combines codes of the variables “pain symptoms – how often the person complains or shows evidence of pain” and “pain symptoms –intensity of highest level of pain present”. This algorithm originates a range of scores from 0 (no pain) to 4 (excruciating daily pain), with intermediate scores of 1 (less than daily pain), 2 (daily pain but not severe) and 3 (daily severe pain). The Instrumental Activities of Daily Living (IADL) scale assesses the persons’ level of dependence in the performance of activities such as preparing meals, doing house chores or managing personal finances, and is based on the recoding and summation of the variables Self-Performance in IADLs and capacity for “meal preparation”, “ordinary housework”, “managing finances”, “managing medications”, “phone use”, “stairs”, “shopping” and “transportation”. This scale produces a total score ranging from 0 to 48, where higher scores represent greater dependence and lower scores independence or low dependence in the performance of IADLs.

The Delirium Scale (DELIR) assesses the presence of delirium symptoms and is based on the recoding and summation of the variables “periodic disordered thinking or awareness – easily distracted”, “periodic disordered thinking or awareness – episodes of disorganized speech”, “periodic disordered thinking or awareness – mental function varies over the course of the day” and “acute mental status change”. The recoding and summation of these variables generates a scale from 0 to 4, where higher scores indicate of a higher probability of delirium.

The Activities of Daily Living Hierarchy Scale (ADL_H) assesses the person’s level of dependence in the performance of daily activities, as using the toilet, walking or eating, and is based on an algorithm which combines the variables “self-performance in personal hygiene”, “self-performance in locomotion”, “self-performance in toilet use”, “self-performance in eating”. This scale uses the concepts of “early-loss”, “middle-loss” and “late-loss” in the performance of ADLs. In the recoding of the variables, the recoded self-performance in personal hygiene corresponds to the variable “early-loss”, the recoded performance in locomotion corresponds to the variable “middle-loss”, the recoded performance in eating corresponds to the variable “late-loss”, and the recoded performance in toilet use corresponds to the variable “Toilet use”. This scale is divided into seven different performance levels: “Independent”, “Supervision required”,

“Limited impairment”, “Extensive assistance required – 1”, “Extensive assistance required – 2”, “Dependent”, “Total dependence”.

The Short ADL Scale (AVD – BR) uses the variables: “self-performance in personal hygiene”, “self-performance in mobility”, “self-performance in toilet use” and “self-performance in eating”, which are recoded and summed to range from 0 to 16. Higher scores mean greater dependence and lower scores lower dependence in the performance of ADL. This scale is more useful to assess the function as a whole, while the hierarchy scale is more useful when the goal is to find the person’s performance pattern.

The Depression Rating Scale – short version (DRS) assesses the presence of depression symptoms and is based on the recoding and summation of the variables “Self-reported mood –little interest or pleasure in doing things s/he normally enjoys”, “Self-reported mood – anxious, restless or uneasy” and “Self-reported mood – sad, depressed or hopeless”. The summation of these recoded variables creates a range from 0 (no symptoms of depression) to 3 (all symptoms present on the last 3 days/24 hours), with intermediate values 1 (one symptom sometimes present) and 2 (at least two symptoms present on the last 3 days/24 hours).

The Body Mass Index (BMI) is a nutritional measure which relates the person’s weight and height, originating a score which represents obesity when high and weakness and potential malnutrition when low. The formula to calculate BMI is $(\text{Weight}/\text{Height}^2)$. Taking into account the methodological requirements and the complexity of this measuring instrument, the goals of this study are as follows: to build a Portuguese version of the InterRAI AC, which we called InterRAI-AC-PT, translate it and ensure its validation to the Portuguese population; and to validate the InterRAI AC-PT.

Methodology

Permission to use the instrument InterRAI-AC within a research project aiming to measure the effectiveness of nursing care was required from and granted by the InterRAI organization. As this instrument was designed to be mostly applied to older people and as most hospitalised patients are elderly, its application was considered both pertinent and useful for the analysis of collected data has the potential to improve

quality as it allows the assessment of outcomes and case, benchmarking and planning analysis, (Gray et al., 2008). Both the instrument and its User Manual were translated.

Cultural and linguistic validation

Linguistic validation was achieved through a translation/back-translation process with expert analysis. Content validity in terms of relevance, clarity and clinical significance in responding to the goal of the study, as well as the translation of its User Manual and the coding system, were carried out with the collaboration of a group of experts composed of nurses and nursing teachers, as is proposed by Goodwin and Leech (2003).

In order to assess item clarity and cultural equivalence, an analysis of non-responses by variable or of invalid responses due to wrong coding was performed. According to Wellens et al. (2010), the existence of missing data may derive from items not clear enough or culturally inadequate, difficulty to collect data, or an irrelevant construct. On the other hand, if responses have many invalid codes, that is because its elaboration needs improvement or the coding system is not clear.

Internal structure validity

Taking into account the nature of this instrument, it is important to verify to what extent there is consistency between the responses to questions belonging to different variables which are interrelated (twin scores). It is then necessary to analyse interrelations between different variables which assess a specific part of the same construct. The coding of some variables must therefore be consistent with the

coding of other variables (Goodwin & Leech, 2003; Wellens et al., 2010).

Therefore, response consistency was assessed by analysing the coding of the interrelated variables. To this end, twenty-five hypotheses were formulated with the purpose of reflecting those interrelations (Table 1).

Reliability

The internal reliability of InterRAI-AC was assessed through the analysis of two key-aspects: interrater reliability and internal consistency (Wellens et al., 2011).

Evidence of reliability based on equivalence was tested through interrater reliability. Thus, double assessments of a group of 31 patients were carried out by nurses of the services and by the research team. These assessments were randomly performed within a period of less than 48 hours and the second rater was not aware of the previous coding.

Cohen's Kappa Coefficient was used to analyse the level of agreement in data classification. According to Landis and Koch (1977), k values lower than 0 indicate no agreement; between 0 and 0.20 a slight agreement; between 0.21 and 0.40 a fair agreement; between 0.41 and 0.60 a moderate agreement; between 0.61 and 0.80 a substantial agreement; above 0.81 an excellent or almost perfect agreement, i.e. both raters codified the same questions similarly or almost perfectly. To assess the values of kappa, the levels of agreement proposed by Landis and Koch (1977) were used. However, as there were few values corresponding to the first (< 0), second ($0 - 0.20$) and third ($0.21 - 0.40$) groups, these three groups were gathered into one.

TABLE 1 – Evidence based on internal structure – Analysis of inter-variable correlations.

No.	If	So
R1	Intake = Long-term care facility	Living arrangements = With non-relative(s)
R2	Confined to bed for medical reasons	Primary mode of locomotion = Bedbound
R3	Primary mode of locomotion = Bedbound	Refused to do the timed 4 meter walk test or not tested
R4	Bladder continence controlled with any catheter or ostomy	Urinary collection device
R5	Did not walk	Refused to do the timed 4 meter walk test or not tested
R6	Due to diminished energy, unable to start some normal day-to-day activities	Maximal assistance, total dependence or meal preparation did not occur
R7	Due to diminished energy, unable to start some normal day-to-day activities	Maximal assistance, total dependence or ordinary housework did not occur
R8	Due to diminished energy, unable to start some normal day-to-day activities	Maximal assistance, total dependence or managing personal finances did not occur
R9	Due to diminished energy, unable to start some normal day-to-day activities	Maximal assistance, total dependence or managing medications did not occur
R10	Due to diminished energy, unable to start some normal day-to-day activities	Maximal assistance, total dependence or phone use did not occur
R11	Due to diminished energy, unable to start some normal day-to-day activities	Maximal assistance, total dependence or use of stairs did not occur
R12	Due to diminished energy, unable to start some normal day-to-day activities	Maximal assistance, total dependence or doing the shopping did not occur
R13	Due to diminished energy, unable to start some normal day-to-day activities	Maximal assistance, total dependence or transportation use did not occur
R14	Due to diminished energy, unable to start some normal day-to-day activities	Maximal assistance, total dependence or bathing did not occur
R15	Due to diminished energy, unable to start some normal day-to-day activities	Maximal assistance, total dependence or personal hygiene did not occur
R16	Due to diminished energy, unable to start some normal day-to-day activities	Maximal assistance, total dependence or walking did not occur
R17	Due to diminished energy, unable to start some normal day-to-day activities	Maximal assistance, total dependence or transfer to toilet did not occur
R18	Due to diminished energy, unable to start some normal day-to-day activities	Maximal assistance, total dependence or no activity regarding bathroom use
R19	Due to diminished energy, unable to start some normal day-to-day activities	Maximal assistance, total dependence or bed mobility did not occur
R20	Due to diminished energy, unable to start some normal day-to-day activities	Maximal assistance, total dependence or eating did not occur
R21	No signs of pain	Pain Intensity = No pain
R22	Pain Intensity = No pain	Pain Consistency = No pain
R23	Benefits from community assistance to perform housework	Minimum ordinary housework and limited assistance
R24	Benefits from community assistance with meal services	Minimum meal preparation and limited assistance
R25	Eating activity did not occur	Nutritional intake did not occur

Cronbach's α can be used to assess internal consistency; however, this method does not apply to the InterRAI AC as a whole due to its multidimensional nature. Its variables should not be summed to achieve a total score, and Cronbach's α may be significantly affected by a great number of variables, as in this case. Nevertheless, this technique can be used to test the internal consistency of the various scales which compose the instrument, but only with those which result from the sum of variables, since the ones composed of algorithms do not ensure an equal contribution of each variable to the total score

(Wellens et al., 2011).

The internal consistency using Cronbach's α of the Communication Scale, the IADL Scale, the Delirium Scale, the Depression Rating Scale – short version, and the Short ADL Scale was measured at the three moments (Preadmission, Admission and Discharge), taking into account how each scale was elaborated (Wellens et al., 2011). Another way to test the scales' internal consistency is to verify the existing correlations between the variables and the total score. Pearson's correlation coefficient (r), Spearman's

Rho and Cramer's V were used for this purpose. The variables are expected to have a significant and strong correlation with the scale to which they belong (Goodwin & Leech, 2003; Streiner & Norman, 2008).

Study design and sample

This is a descriptive correlational longitudinal study. Data collection took place between March and July 2012 in general surgery and medicine units of four acute care hospitals in the centre region of Portugal. The nurses who were part of the teams of these services were asked to collaborate in data collection. The head nurse of each service appointed two nurses to coordinate data collection among the patients and families and to collaborate with the research team.

Data could be recorded on paper or electronically. A software application was designed and used in the services which requested it.

All nurses went through several training sessions on the InterRAI AC and its administration process.

A member of the research team was assigned to each unit to clarify any aspects and collect patients' data in order to assess the consistency of the records.

The sample consisted of all patients hospitalised in the units for three or more days during the data collection period.

Ethical procedures

Permission to carry out the study was required from the Board of Directors of the hospitals. This permission was granted after assent from the respective Ethics Committees.

Permission to participate in the study was asked from all patients or relatives through a declaration of consent. This permission followed a clarification session on the study and its goals. Data confidentiality was also assured, as well as their participation involving no costs or harm.

Statistical analysis

Data were analysed using the Statistical Package for the Social Sciences (SPSS) software, version 19.0.

Frequency and descriptive measures were used, such as the mean and standard deviation. According to the degree of measurability of the variables, the chi-square test to assess independence between qualitative variables, the one-way ANOVA to compare means, and the Pearson's *r* and the Spearman's Rho correlation coefficients to analyse the relationship between

numeric and categorical variables, respectively, and the Cramer's V were used.

Results

Cultural and linguistic equivalence

This process was carried out by a group of PhD nursing teachers and a group of nurses with a professional experience of at least 5 years, who could all speak English. The instrument was designed with slight changes from the original instrument, namely the elimination of the 14th day of hospitalisation as an assessment moment, considering that the average hospitalization period in Portuguese hospitals with the same type of sample is 7.64 days, as according to the Final Report of the Technical Group on Hospital Reform created by Order No. 10.601/2011 of the Ministry of Health (Direcção Geral da Saúde, 2011).

The component regarding medication was also removed for that analysis would not be useful for this study. In what concerns skin condition, a question regarding the risk of acquiring a pressure ulcer was included. This variable may be used as evidence of the effectiveness of nursing care since it may establish an association between risk diagnosis and the presence or not of a pressure ulcer.

The following questions were excluded from the identification section since they were not considered relevant: ethnicity/race, name of the treating doctor, national numeric identifier, facility/agency provider number and current payment sources for in-person stay. In the section on intake and initial history, the code for Mental Health Residence was also eliminated as it does not exist in the Portuguese Health System.

Characteristics of the sample

A total of 1823 questionnaires were collected. After reviewing the questionnaires to eliminate those which were very incomplete or did not meet inclusion criteria, a total of 1764 questionnaires were gathered. When comparing the samples from the 4 hospitals by using the chi-square test, no significant differences in the percentage of men and women were found ($\chi^2=6.626$; $p=0.085$). The mean age was 70.78 years, and ANOVA showed that there were no significant differences in the mean age among the 4 hospitals ($F=0.604$; $p=0.612$). Table 2 shows that most patients come from medicine units (64%), and that 12.3% of the patients live alone.

TABLE 2 – Sample Distribution.

Variable	Values	No	%
Gender	Male	849	50.0%
	Female	850	50.0%
Marital Status	Married	944	55.0%
	Widowed	495	28.8%
	Single	194	11.2%
Age	Mean	70.8	
	Standard deviation	17.0	
Mother tongue	Portuguese	1731	98%
Intake	Private or rented home	1430	81.9%
	Nursing home	194	11.1%
Living arrangements	With spouse /partner only	530	30%
	With spouse/partner and others	377	21.4%
	With child	268	15.4%
	With non-relative(s)	233	13.4%
Ward	Alone	214	12.3%
	Medicine	1129	64%
	Surgery	635	36.0%

Interrater reliability

A total of 151 items of the InterRAI-CA-PT, distributed across 13 sections, collected at the three moments – Preadmission, Admission and Discharge – were assessed. The proportion of agreement was almost perfect in 44% (66) of the items, substantial in 41% (62) of the items, moderate in 14% (22) of the items and poor in 1% (1) of the items (Table 3).

The most interesting items are those in which the rater interprets data collected from the patient, such as those on cognitive performance, mood, ADLs,

IADLs, and pain (Gray et al, 2008). In the section on “Mood and Behaviour”, which includes mood-related items, the level of agreement in 80% of the items was higher than substantial, i.e. the values of kappa were higher than 0.61. In the section on “Functional Status”, which includes items related to ADLs and IADLs, the level of agreement in 94% of the items was higher than substantial. In the section on “Health Conditions”, which includes pain-related items, the level of agreement in 67% of the items was higher than substantial.

TABLE 3 – Interrater agreement.

Section	Total of variables by level of agreement (K)				Total of variables tested by section
	Poor <0.40	Moderate 0.41–0.60	Substantial 0.61–0.80	Excellent >0.80	
Identification	0	0	0	4	4
Intake and initial history	0	0	1	5	6
Cognition	0	0	5	14	19
Communication, vision and hearing	0	2	5	3	10
Mood and behaviour	0	1	6	3	10
Functional status	0	3	23	21	47
Continence	0	3	2	7	12
Health Conditions	0	6	11	1	18
Oral and nutrition status	0	1	1	1	3
Skin condition	0	2	3	0	5
Treatments and procedures	0	0	3	1	4
Discharge potential	1	4	2	5	12
Discharge	0	0	0	1	1
Total of variables by level of agreement	1	22	62	66	151

Internal consistency

The Cronbach's α for each scale was higher than 0.80 (Table 4), and it did not increase with the elimination of any of the variables which compose the scales. Correlation between the items which compose the Depression Rating Scale – short version and the scale itself were tested, and the correlation values were

all higher than 0.8. When correlations between the items which compose the Communication Scale and the scale itself were tested, quite high correlations, above 0.90, were found. In the Scales of Delirium, Instrumental Activities of Daily Living Scale and Short ADL scale, correlations were higher than 0.8.

TABLE 4 – Item-total correlation and scale reliability.

Scale/Subscales	Pearson's r	Cronbach's α		
		Pre-Admission	Admission	Discharge
Depression rating – short version		0.863	0.833	0.870
Little interest or pleasure in doing things s/he normally enjoys	0.864			
Anxious, restless or uneasy	0.856			
Sad, depressed or hopeless	0.877			
Communication		0.971	0.963	0.963
Be understood	0.986			
Understand	0.986			
Delirium			0.946	0.956
Easily distracted	0.893			
Episodes of disorganized speech	0.914			
Mental function varies over the course of the day	0.903			
Instrumental activities of daily living		0.963		0.968
Meals	0.868			
Ordinary housework	0.870			
Finances	0.918			
Medications	0.918			
Telephone	0.885			
Stairs	0.901			
Shopping	0.933			
Transportation	0.921			
Activities of Daily Living – short version		0.977	0.948	0.971
Performance in hygiene	0.938			
Performance in mobility	0.943			
Toilet use	0.961			
Performance in eating	0.876			

Content validity

The percentages of missing data by variable are irrelevant, as can be seen in Table 5, i.e. 13 variables (7%) have no missing data, 104 (57%) have less than 1% of missing data, 57 variables (31%) have between 1% and 2% of missing data, and only 9 (5%) have more than 2% of missing data. Height and weight variables

had the highest percentage of missing data, 12.9% and 13% respectively, but this can be due to the fact that data may be difficult to obtain as a consequence of functional dependence. At the three moments of assessment of the InterRAI-AC, the completion of the instrument was identical and the percentage of missing data was almost similar.

TABLE 5 – Percentages of missing data in each section of the InterRAI AC.

Section	% of Missing Data				Total of variables by section
	0%	< 1%	1% - 2%	> 2%	
Identification	3	0	1	3	7
Intake and initial history	0	2	4	0	6
Cognition	0	15	5	0	20
Communication, vision and hearing	0	9	1	0	10
Mood and behaviour	9	0	3	0	12
Functional Status	1	42	7	0	50
Continence	0	11	1	0	12
Health status	0	20	3	0	23
Oral and nutritional status	0	1	1	3	5
Skin condition	0	3	1	1	5
Treatments and procedures	0	0	18	1	19
Discharge potential	0	0	12	1	13
Discharge	0	1	0	0	1
Total of variables by % of missing data	13	104	57	9	183

When analysing the quantity of invalid codes by variable, 67 (31%) of the 214 analysed variables were considered invalid. From these, 31 had only one invalid code in a single completed questionnaire, and in the remaining variables the quantity of invalid codes was considered as irrelevant.

Construct validity

In order to test the internal structure of the InterRAI-AC-PT, correlations between the 25 hypotheses formulated and presented in Table 1 were analysed. Results confirmed total consistency (100% of consistency) in all verified correlations.

Concurrent validity

In addition to the correlations between the variable codes, seven inter-variable correlations were also analysed. According to evidence from literature, these variables are expected to correlate, as can be seen in Table 6.

These seven correlations proved to be significant: functional status with cognitive performance; functional status with Body Mass Index; functional status with physical activity; functional status with vision (Stuck et al., cited by Wellens et al., 2010); bowel continence with pressure ulcers; pressure ulcers with bed mobility (Allman, cited by Wellens et al., 2010); and pressure ulcers and awareness changes (Maklebust & Magnan, cited by Wellens et al., 2010).

TABLE 6 – Validity based on inter-variable correlation.

Variable 1	Variable 2	N	r	P value
Cognitive performance	Functional Status	1732	0.759	0.000
Body Mass Index	Functional Status	1507	-0.087	0.001
Physical activity	Functional Status	1743	0.473	0.000
Vision	Functional Status	1740	0.471	0.000
Bowel continence	Pressure ulcer	1751	0.442	0.000
Bed mobility	Pressure ulcer	1749	0.429	0.000
More severe pressure ulcer	Awareness changes	1703	0.247	0.000

Discussion

The translation, validation and cultural adaptation of an instrument into a language other than the

original are an important part of the process. They help develop validation methods and the instrument itself as it is by comparing the results of its use that the instrument is validated in terms of relevance and

usefulness (Wellens et al., 2010). Furthermore, it is a challenge which requires an adequate methodology so as to be safely applied with minimum probability of error. Since settings vary, adaptation and validation processes are essential to compare results between institutions, healthcare provision systems and countries.

In the case of instruments from the InterRAI organization, particularly the InterRAI-AC, this process is even more challenging and complex, as they are comprehensive, multidimensional instruments which must be assessed as a whole, but also crosswise to all the domains, at all moments, and by item (Hirdes et al., 2008).

Considering the nature of the instrument, the recommendations and point of view of Wellens et al. (2010, 2011) were followed, regarding both validity and reliability assessment.

The assessment of the construct validity is usually performed using factor analysis, whose aim is to reduce a large set of variables into a smaller set with common characteristics, or the creation of underlying dimensions. Since the InterRAI-AC has the characteristics of a minimum data set, where each domain is previously identified, it already has a minimum number of items and a total score is not calculated. Wellens et al. (2010) consider that, during the validation process, a decrease in the number of variables through factor analysis techniques is not appropriate.

The obtained results confirm that the Portuguese version is another step towards the consolidation of the InterRAI-AC as an instrument capable of helping health professionals to collect data on their patients. The InterRAI-AC will facilitate the production of information and knowledge to improve the quality of care and help establish policies for a more person-centred and outcome-centred organization, in addition to the monitoring and validation of data produced by other means and instruments.

Results show that the mean age of the sample was 70.78 years, corresponding to an elderly population mostly living in their own homes. Concerning content validity and cultural adaptation, all experts were unanimous as to the value of this instrument as a whole, of its items and of its usefulness.

As to equivalence, when data on the same patient was collected and recorded by different raters, there was a substantial agreement in most cases. In situations

which require more interpretation from the rater, such as "Mood and Behaviour", in 80% of the cases the levels of agreement were higher than substantial (> 0.61). In items related to ADLs and IADLs, inter-rater reliability was also higher than substantial in 94% of the items. These data are in line with those obtained by Hirdes et al. (2008).

In terms of the scales' internal consistency, they all presented a value higher than 0.80, and most of them had values higher than 0.90, which indicates a strong internal consistency of the scales and contributes to good reliability.

In what concerns validity based on internal structure, which was analysed through the correlation between the codings of the related variables (Wellens et al., 2010), corresponding to the 25 formulated hypotheses, a 100% consistency was verified. According to the authors, this result is essential to validate complex instruments such as the InterRAI-AC.

Significant correlations between variables which theory refers as being related were also achieved. Regarding these correlations, it must be said that their statistical significance is mainly due to sample size, as the correlation scores in some of the cases are < 70 , which means that the variance of a variable explained by the variance of another variable is not always high. Examples of this are the correlations between: BMI and functional status ($r = -0.087$); physical activity and functional status ($r = 0.473$); and vision and functional status ($r = 0.471$). However, the contribution and the need to intervene in these aspects to achieve physical activity gains is an important data for this validation and should be considered.

Limitations

The time taken to fill out each questionnaire (about 15 minutes) turned out to be a limitation in this study. In circumstances where time is already short for the provision of care, nurses' availability for the completion of the questionnaire is lower, which affects the sample size.

Conclusion

The analyses carried out to validate the InterRAI-AC-PT confirm that the instrument is valid and reliable.

It is therefore an excellent tool to be used in acute care settings. This instrument can be very useful in collecting data for various purposes, such as quality indicators, care planning, evidence base for decision making, and others.

In addition to its occasional use in research, the InterRAI-AC-PT could be used as an essential tool in the process of patient data collection since this instrument provides a holistic perspective of the patient's history and evolution, thus enabling the provision of more documented care.

The main obstacle to the application of the InterRAI-AC-PT is the time required to fill it out. Further studies are necessary to assess whether the time spent in the application of this instrument may have the improvement of effectiveness and quality of care as an outcome.

This paper was written as part of the research project "Nursing Outcomes: Quality and Effectiveness" (FCT-PTPC/CS-SOC/113519/2009), which is integrated in the Health Sciences Research Unit: Nursing.

Referências bibliográficas

- Brand, C., Martin-Khan, M., Wright, O., Jones, R., Morris, J., Travers, C.,...Gray, L. (2011). Development of quality indicators for monitoring outcomes of frail elderly hospitalized in acute care health settings: Study Protocol. *BMC Health Services Research*, 11 (281), 1-8.
- Cronbach, L. (1980). Validity on parole: How can we go straight. In W. B. Schrader (Ed.), *New directions for testing and measurement: Measuring achievement, progress over a decade* (Vol. 5, pp. 99-108). San Francisco, CA: Jossey-Bass.
- Donabedian, A. (1992). The role of outcomes in quality assessment and assurance. *Quality Review Bulletin*, 18 (11), 356-360.
- Doran, D., Sidani, S., Watt-Watson, J., Laschinger, H., & Hall, L. (2001). A methodological review of the literature on nursing-sensitive outcomes. In Ministry of Health and Long-term Care, University of Toronto, The Nursing and Effectiveness, Utilization and Outcomes Research Unit, *Invitational Symposium: Nursing and Health Outcomes Project*. Toronto, Canada: Ministry of Health and Long Term Care.
- Glenny, C., & Stolee, P. (2009). Comparing the functional independence measure and the interRAI/MDS for use in the functional assessment of older adults: A review of the literature. *BMC Geriatrics*, 9, 52.
- Goodwin, L. D., & Leech, N. L. (2003). The meaning of validity in the new standards for educational and psychological testing: Implications for measurement courses. *Measurement and Evaluation in Counseling and Development*, 36, 181-191.
- Gray, L., Berg, K., Bernabei, R., Blasco, S., Carpenter, G., Finne-Soveri, H.,... Topinkova, E. (2006). *InterRAI Acute Care. Guide for use of the InterRAI AC assessment form* (Version 09). Ann Arbor, MI: InterRAI.
- Gray, L., Bernabei, R., Berg, K., Finne-Soveri, H., Brant, F., Hirdes, J.,...Ariño-Blasco, S. (2008). Standardizing assessment of elderly people in acute care: The interRAI Acute Care instrument. *Journal of the American Geriatrics Society*, 56, 536-541.
- Hirdes, J., Ijunggren, G., Morris, J., Fijters, D., Soveri, H., Gray, L.,...Gilgen, R. (2008). Reliability of the InterRAI suite of assessment instruments: A 12-country study of an integrated health information system. *BMC Health Services Research*, 8, 277.
- Instituto Nacional de Estatística. (2013). *Estatísticas demográficas 2011*. Lisboa, Portugal: Autor.
- InterRAI (2012). *Acute and Post-Acute Care (AC and PAC)*. Ann Arbor, MI: InterRAI. Retrieved from <http://www.interrai.org/welcome.html>
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33 (1), 159-174.
- Messick, S. (1980). Test validity and ethics of assessment. *American Psychologist*, 35, 1012-1027.
- Ministério da Saúde. Direção Geral da Saúde. (2011). *Relatório final do grupo técnico para a reforma hospitalar. Os cidadãos no centro do sistema. Os profissionais no centro da mudança*. Retrieved from http://www.portaldasaude.pt/NR/rdonlyres/84FCFCE2-3C84-4ABE-8E5F-AD4DDB0B46F4/0/RelatorioGTRH_Nov2011.pdf
- Porter, M. E. (2010). What is value in health care? *The New England Journal of Medicine*, 363, 2477-2481.
- Sidani, S., Doran, D., & Mitchell, P. (2004). A theory-driven approach to evaluating quality of nursing care. *Journal of Nursing Scholarship*, 36 (1), 60-65.
- Streiner, D. L., & Norman, G. R. (2008). *Health measurement scales: A practical guide to their development and use* (4th ed). Oxford, England: Oxford University Press.
- Wellens, N., Milisen, K., Flamaing, J., & Moons, P. (2010). Methods to assess the validity of the InterRAI Acute Care: A framework to guide clinimetric testing. *Journal Evaluation in Clinical Practice*, 18 (2), 296-306.
- Wellens, N., Milisen, K., Flamaing, J., & Moons, P. (2011). Methods to assess the validity of the interRAI Acute Care: A framework to guide clinimetric testing. Part II. *Journal Evaluation in Clinical Practice*, 18 (4), 822-827.
- Wellens, N., Lancker, A., Flamaing, J., Gray, L., Moons, P., Verbeke, G.,...Milisen, K. (2012). Interrater reliability of the interRAI Acute Care (interRAI AC). *Archives of Gerontology and Geriatrics*, 55 (1), 165-172.