

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

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

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

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

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.

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.

* Professor Coordenador na Escola Superior de Enfermagem de Coimbra. Mestre em Gestão e Economia da Saúde, 3001 – 091, Coimbra, Portugal [amaral@esenfc.pt].

Morada: Escola Superior de Enfermagem de Coimbra Av. Bissaya Barreto 3001 - 091 Coimbra, Portugal.

** Professor Associado com Agregação na Faculdade de Economia da Universidade de Coimbra. Doutorando em Engenharia Industrial – Teoria da Decisão/Sistemas em Saúde, Centro de Estudos e Investigação em Saúde da Universidade de Coimbra, Faculdade de Economia da Universidade Coimbra Avenida Dias da Silva, 165 3004-512, Coimbra, Portugal [pedrolf@fe.uc.pt].

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

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.

Recebido para publicação em: 20.11.13

Aceite para publicação em: 10.02.14

Introdução

A medição de resultados em saúde, como forma de produzir indicadores de efetividade, de melhorar a qualidade dos cuidados e de garantir a prestação de contas (*accountability*) de prestadores e das unidades, tornou-se uma prática essencial nos diferentes contextos em que os cuidados de saúde são produzidos e está muito ligada aos processos de financiamento das organizações e a uma tendência internacional de redução de custos. Isto é muito evidente nos cuidados de enfermagem, uma vez que os serviços prestados pelos enfermeiros representam uma fatia importante dos cuidados de saúde e os gestores exigem a justificação dos gastos que lhe estão associados e, sobretudo, a sua ligação com os resultados que se obtêm nos doentes (Donabedian, 1992).

Neste contexto a medição dos resultados tem que ser um objetivo comum que una os interesses e as atividades de todas as partes interessadas, porque tem repercussões profissionais, financeiras e políticas, sendo por isso necessário encontrar instrumentos de medida que sejam fiáveis e válidos (Porter, 2010).

Para analisar o efeito dos cuidados de enfermagem na obtenção dos melhores resultados para os pacientes é necessário avaliar os resultados sensíveis aos cuidados de enfermagem, considerados por Doran, Sidani, Watt-Watson, Laschinger e Hall (2001) como um estado, um comportamento ou percepção de um doente ou família, mensurável ao longo de um *continuum*, que ocorre em resposta a uma intervenção de enfermagem.

As organizações e os sistemas de saúde têm enfrentado grandes desafios motivados pelo envelhecimento da população com correspondente aumento das doenças crónico-degenerativas e com recursos cada vez mais limitados. De acordo com o Instituto Nacional de Estatística (2013), o Índice de Envelhecimento da População passou de 103 para 128 entre o ano de 2001 e 2011. Isto constitui um desafio porque as pessoas idosas, assim como as que têm incapacidades, estão mais suscetíveis à ocorrência de problemas de saúde que exigem cuidados multidisciplinares (Hirdes et al., 2008). A fragilidade da pessoa e a complexidade do estado de saúde fazem aumentar o risco de resultados adversos associados ao internamento hospitalar como, por exemplo, as quedas e as úlceras de pressão aos quais se associam

mais custos socioeconómicos, nomeadamente pelo aumento no tempo de internamento e despesas, pela taxa de mortalidade mais elevada, pelo declínio após a alta, pelas altas taxas de readmissão e pela perda de capacidades para realizar as atividades de vida diária, entre outras (Brand et al., 2011).

A complexidade e multidimensionalidade das situações exigem a utilização de instrumentos compreensivos que meçam mais do que uma dimensão da pessoa e, principalmente, espelhem uma visão holística dos cuidados (Hirdes et al., 2008). A sua utilização coloca, no entanto, múltiplos desafios aquando da validação, uma vez que não se podem utilizar, do mesmo modo, os testes habituais usados nos instrumentos que medem apenas uma dimensão da pessoa (Wellens, Milisen, Flamaing, & Moons, 2010).

Validação e fiabilidade

Os processos de validação de instrumentos têm-se desenvolvido de acordo com a evolução do conceito de validade, que passou de uma propriedade estática de uma medida, em que um teste era válido ou não de acordo com a correlação existente entre a medida e outra medida de critério externo, para uma outra visão, em que a validade foi categorizada em tipos específicos: validade de conteúdo; validade do constructo; e validade de critério (que inclui a validade concorrente e a validade preditiva). Mais tarde foi apresentada uma abordagem multi-método que introduziu os tipos de validade convergente e discriminante fazendo parte da validade de constructo. Nos anos 80 emergiu uma nova conceptualização de validade, liderada por Cronbach (1980) e Messick (1980), que enfatizou as inferências e as decisões a partir dos valores obtidos pelos testes. Este processo levou à standardização dos testes psicológicos e, partir daí, a validade passou a ser definida como a adequação, significado e utilidade das inferências feitas pelos testes sendo essa acumulação de evidência que constitui a base científica para a interpretação dos *scores* (Goodwin & Leech, 2003). Assim, os últimos padrões para a validação de instrumentos indicam cinco linhas de força que revelam a evidência de validade, sendo elas: evidência baseada no conteúdo; responsividade (processo de respostas); estrutura interna; relação com outras variáveis; e consequências dos testes (*Idem*).

Quanto à fiabilidade, que se refere à capacidade de um qualquer instrumento medir de forma

consistente e confiável um qualquer atributo populacional, independentemente do momento em que se mede e de quem mede, os instrumentos compreensivos também colocam vários desafios (Wellens, Milisen, Flamaing, & Moons, 2011; Streiner & Norman, 2008). Desse modo os autores advogam a necessidade de medir a evidência de fiabilidade baseada na equivalência das medidas efetuadas por dois observadores diferentes, a evidência baseada na estabilidade e na consistência interna. A evidência baseada na fidelidade é muito importante para se verificar se uma medida mede apenas uma ideia e se as variáveis que compõe a escala têm consistência interna (Streiner & Norman, 2008).

InterRAI-AC

O InterRAI é um grupo de investigação internacional que desenvolve instrumentos de avaliação compreensiva orientados para pessoas que necessitam de cuidados, sobretudo pessoas idosas (InterRAI, 2012). Os Instrumentos de Avaliação de Residentes (RAI) são utilizados internacionalmente em vários contextos e com diferentes fins, como planeamento de cuidados, medição de resultados e indicadores de qualidade (Glenny & Stolee, 2009).

Atualmente existem 18 instrumentos de avaliação InterRAI para diferentes contextos de prestação de cuidados (cuidados paliativos, cuidados domiciliários, cuidados em lares, cuidados pós-agudos, cuidados em saúde mental comunitária, entre outros). O InterRAI AC, apresentado em 2006, foi desenvolvido para ser utilizado em hospitais de cuidados agudos onde as pessoas podem beneficiar de uma avaliação abrangente por múltiplos especialistas (Gray et al., 2008).

O instrumento InterRAI-AC foi desenvolvido para ser aplicado em pessoas com limitações funcionais em atividades pessoais, como as Atividades de Vida Diária (AVD) e em Atividades Instrumentais de Vida Diária (AIVD); em pessoas com limitações cognitivas; em pessoas com doenças que sejam acompanhadas de perda da capacidade funcional para a realização das atividades pessoais, ou em pessoas com mais de 75 anos por existir maior probabilidade de desenvolvimento destes problemas; e fornece uma informação holística da complexidade e fragilidade das pessoas internadas. (Gray et al., 2006; Gray et al., 2008).

Os instrumentos InterRAI são compostos por um conjunto de itens, que são comuns a todos os

instrumentos (cerca de 70 ex.: AVD e Dor); cerca de 100 itens estão presentes em quase todos os instrumentos e alguns são específicos para contextos especiais e variam consoante os instrumentos (Wellens et al., 2012).

O InterRAI-AC é composto por várias dimensões que avaliam diferentes domínios clínicos, em quatro momentos: na Pré-Admissão, ou seja, no período de três dias antes do aparecimento da situação que causou o internamento (neste caso os informantes podem ser os familiares se os doentes não forem capazes de o fazer); na Admissão, ou seja, no período de 24 horas seguintes ao internamento; ao 14º dia de internamento; e na Alta, sendo que a avaliação se reporta às 24 horas anteriores à saída. Assim, a particularidade deste instrumento, ao nível da sua extensão e estrutura, é a possibilidade de avaliar todo o episódio de internamento da pessoa. A avaliação da situação anterior à doença não só permite estabelecer uma ligação com o estado atual, como pode servir de referência para a reabilitação e para o tratamento. Para além disto, a medição de resultados exige um ponto de referência e, como referem Sidani, Doran e Mitchell (2004), o estado do doente antes da situação que o levou a recorrer ao hospital influencia o resultado dos cuidados. Porque se trata de uma medida essencialmente destinada a apoiar a prática clínica, após o seu preenchimento total não é possível determinar um valor final, pelo que se calculam os valores obtidos em cada uma das escalas que refletem a severidade da doença e a sua evolução em domínios selecionados (Wellens et al., 2010; Hirdes et al., 2008). O InterRAI Cuidados Agudos versão portuguesa (InterRAI-AC-PT) é constituído por 16 secções: Identificação; Admissão e História Inicial; Datas de Avaliação; Cognição; Comunicação, Visão e Audição; Comportamento e Ânimo; Estado Funcional; Continência; Diagnósticos; Condições de Saúde; Estado Nutricional e Oral; Estado da Pele; Tratamentos e Procedimentos; Potencial da Alta; Alta; Informação da Avaliação.

Das secções do instrumento é possível produzir, de acordo com o Manual, através de somatórios ou de algoritmos que combinam itens, um conjunto de escalas que medem: Desempenho Cognitivo; Comunicação; Dor; Atividades Instrumentais de Vida Diária; Delírio; Hierárquica das Atividades de Vida Diária; Versão Curta desta Escala; Escala Curta de Classificação da Depressão; Índice de Massa Corporal).

Escala do Desempenho Cognitivo (DES_COG), construída através de um algoritmo que combina as variáveis “capacidades cognitivas para tomar decisões no dia-a-dia”, “capacidade para memorizar/recordar - boa memória de curto prazo”, “capacidade para memorizar/recordar - boa memória processual”, “fazer-se entender” e “auto desempenho nas Atividades de Vida Diária – alimentação”. Este algoritmo origina um conjunto de 6 níveis de desempenho, do nível 0 (pessoa Independente) ao nível 6 (incapacidade muito grave).

Escala da Comunicação (COMUNIC) avalia a capacidade que as pessoas têm para comunicar e é construída através do somatório das variáveis “fazer-se entender” e “capacidade para compreender os outros”. O somatório destas duas variáveis origina um conjunto de oito níveis de capacidade para comunicar de 0 (Independente) a 8 (Incapacidade muito grave). Escala da Dor (DOR) avalia a dor que a pessoa refere e é construída através de um algoritmo que combina códigos das variáveis “sintomas de dor - frequência com que a pessoa se queixa ou mostra sinais de dor” e “sintomas de dor - nível mais elevado de intensidade da dor atual”. Este algoritmo origina um intervalo de 0 (sem dor) a 4 (dor diária insuportável) com valores intermédios 1 (dor não diária), 2 (dor diária mas não severa) e 3 (dor diária severa).

Escala das Atividades Instrumentais de Vida Diária (AIVD) avalia o nível de dependência das pessoas na realização de atividades, como preparação das refeições, as tarefas domésticas ou gestão das finanças pessoais e é construída através da recodificação e somatório das variáveis auto desempenho nas AIVD e capacidade na “preparação das refeições”, “tarefas domésticas”, “gerir as finanças pessoais”, “gerir a medicação”, “uso do telefone”, “escadas”, “compras” e “transporte”. Esta escala origina uma classificação num intervalo de 0 a 48, cujos valores mais altos representam maior dependência e os valores mais baixos independência ou fraca dependência na realização das AIVD.

Escala do Delírio (DELIR) avalia a presença de sintomas de delírio e é construída através da recodificação e somatório das variáveis “perturbações periódicas do pensamento ou consciência - facilmente distraído”, “perturbações periódicas do pensamento ou consciência - momentos de discurso desorganizado”, “perturbações periódicas do pensamento ou consciência - função mental varia ao longo do dia”

e “alteração aguda do estado mental em relação ao habitual”. Da recodificação e somatório destas variáveis resulta um intervalo de 0 a 4, cujos valores mais altos indicam alta probabilidade de delírio.

Escala Hierárquica das Atividades de Vida Diária (AVD_H). Avalia o nível de dependência das pessoas na realização de atividades do dia-a-dia, como usar o WC, caminhar ou comer, e é construída através de um algoritmo que combina as variáveis “auto desempenho na higiene pessoal”, “auto desempenho no caminhar”, “auto desempenho no uso da casa de banho”, “auto desempenho na alimentação”. A construção desta escala é baseada no conceito de perda “precoce”, “média” e “tardia” no desempenho das AVD. As variáveis são recodificadas sendo que o desempenho na higiene pessoal recodificado corresponde à variável “perda precoce”, o desempenho a andar recodificado corresponde à variável “perda média”, o desempenho na alimentação recodificado corresponde à variável “perda tardia” e o desempenho no uso do WC recodificado corresponde à variável “uso do WC”. Este processo origina uma classificação em sete níveis de desempenho distintos: “Independente”, “com supervisão”, “incapacidade diminuta”, “assistência alargada – 1”, “assistência alargada – 2”, “dependente”, “dependência total”.

Versão Curta da Escala das Atividades de Vida Diária (AVD_BR). Utiliza as variáveis: “auto desempenho na higiene pessoal”, “auto desempenho no caminhar”, “auto desempenho no uso da casa de banho” e “auto desempenho na alimentação” que são recodificadas e somadas resultando um intervalo de 0 a 16 e cujos valores mais altos representam maior dependência e os mais baixos menor dependência na realização das Atividades de Vida Diária. Esta escala é mais útil na avaliação da função na globalidade enquanto a escala hierárquica é mais útil quando se deseja conhecer exatamente o padrão de desempenho de uma pessoa. Escala Curta de Classificação da Depressão (DEPRESS) avalia a presença de sintomas de depressão e é construída através da recodificação e somatório das variáveis “ânimo auto reportado - com pouco interesse ou prazer em fazer as coisas de que normalmente gosta”, “ânimo auto reportado - ansioso, impaciente ou inquieto” e “ânimo auto reportado - triste, deprimido ou desesperado”. O somatório destas variáveis recodificadas cria um intervalo de 0 (sem sintomas de depressão) a 3 “todos os sintomas presentes nos últimos 3 dias/24 horas),

com os valores intermédios 1 (“um sintoma presente algumas vezes”, 2 (pelo menos 2 sintomas presentes nos últimos 3 dias/24 horas).

Índice de Massa Corporal (IMC) é uma medida nutricional que relaciona o peso e a altura da pessoa, originando um valor que se for alto representa obesidade e se for baixo representa fraqueza e potencial desnutrição. A fórmula para calcular o IMC é $(\text{Peso}/\text{altura}^2)$.

Tendo em conta as exigências metodológicas e a complexidade deste instrumento de medição, os objetivos do presente estudo são os seguintes: (i) criar a versão portuguesa do InterRAI-AC, por nós denominada InterRAI-AC-PT, o seu processo de tradução e os métodos para assegurar a validação para a população Portuguesa; e (ii) validar o instrumento InterRAI-AC-PT.

Metodologia

Foi solicitada à organização InterRAI a autorização para utilizar o instrumento InterRAI-AC no contexto de um projeto de pesquisa, cujo objetivo é medir a efetividade dos cuidados de enfermagem, tendo esta sido concedida. Como este instrumento foi concebido para ser utilizado sobretudo com pessoas idosas e considerando que os doentes internados nos hospitais são maioritariamente idosos, considerámos que a sua utilização seria não só pertinente como útil porquanto a análise dos dados que produz tem o potencial de melhorar a qualidade pela possibilidade que permite de avaliar resultados, análise de casos, *benchmarking* e planeamento (Gray et al., 2008). Para a sua utilização procedemos não apenas à tradução do instrumento, mas também do seu manual.

Validação cultural e linguística

A validação linguística foi realizada através de um processo de tradução/retro tradução com análise de peritos. A validação do conteúdo das componentes do instrumento, nas vertentes da sua relevância, clareza do enunciado e a sua importância clínica para dar resposta ao objetivo do estudo, bem como a tradução do seu manual e a forma de codificação, foram realizadas com a colaboração de um grupo de peritos constituído por enfermeiros e docentes de enfermagem, tal como propõem Goodwin e Leech (2003).

Para avaliar a clareza dos itens e a sua equivalência cultural foi realizada uma análise da percentagem de não respostas por variável ou de respostas inválidas por codificação errada. Segundo Wellens et al. (2010), a existência de muitos dados omissos pode resultar do facto dos itens não serem suficientemente claros ou culturalmente adaptados, os dados serem difíceis de recolher ou o constructo ser irrelevante, por outro lado se as respostas tiverem muitos códigos inválidos é porque é preciso melhorar a sua redação ou porque a forma de codificação não é clara.

Validade da estrutura interna

Tendo em conta a característica deste instrumento é importante verificar em que medida existe coerência entre as respostas a questões que pertencem a diferentes variáveis, mas que estão relacionadas entre si (“*twin scores*”). Torna-se então necessário analisar as inter-relações entre variáveis distintas, mas que avaliam uma parte específica do mesmo constructo, pelo que a codificação de algumas variáveis tem que ser consistente com a codificação de outras (Goodwin & Leech, 2003; Wellens et al., 2010).

Assim avaliámos a consistência das respostas através da análise das codificações das variáveis que se encontram inter-relacionadas. Para isso formulámos 25 hipóteses, que pretendem espelhar essas inter-relações e que se encontram listadas na Tabela 1.

Fiabilidade

Foi avaliada a fiabilidade interna do InterRAI AC com a análise de dois aspetos chave: a fiabilidade inter-juizes e a consistência interna (Wellens et al., 2011).

A evidência de fiabilidade baseada na equivalência foi testada através da fiabilidade inter-juizes. Assim, foram efetuadas avaliações duplas a um conjunto de 31 doentes por enfermeiros dos serviços e pela equipa de investigadores. Estas foram feitas de forma aleatória, num intervalo que tinha que ser inferior a 48 horas e sem que o 2º avaliador tivesse conhecimento da codificação anterior. Para analisarmos o grau de concordância na classificação dos dados utilizou-se o coeficiente K de Cohen. De acordo com Landis e Koch (1977), valores de K abaixo de 0 significam que não há acordo; entre 0 e 0.20 considera-se uma concordância ligeira; entre 0.21 e 0.40 concordância razoável; entre 0.41 e 0.60 concordância moderada; entre 0.61 e 0.80 é considerado uma concordância substancial; acima de 0.81 considera-se excelente ou quase perfeita,

ou seja, os dois avaliadores codificaram de forma igual ou quase perfeita as mesmas questões. Para a avaliação dos valores de K foram considerados os descritores de força e direção indicados por Landis e

Koch (1977), no entanto devido ao facto de existirem poucos valores correspondentes ao primeiro (< 0), segundo ($0 - 0.20$) e terceiro ($0.21 - 0.40$) grupos, agregaram-se estes 3 num só.

TABELA 1 – Evidência baseada na estrutura interna - Análise das inter-relações entre variáveis

Nº	Se	Então
R1	Proveniência = Unidade de Cuidados Continuados	Formas de coabitação = Com não familiares
R2	Acamada por motivos médicos	Principal modo de locomoção = Acamado
R3	Principal modo de locomoção = Acamado	Recusou-se a fazer o teste de caminhada de quatro metros cronometrada, ou não foi testada
R4	Continência urinária controlada com cateter ou ostomia	Aparelho de recolha de urina
R5	Não percorreu qualquer distância	Recusou-se a fazer o teste de caminhada de quatro metros cronometrada, ou não foi testada
R6	Incapaz de começar qualquer atividade por fadiga	Assistência máxima, dependência total ou não houve atividade relativamente a preparação das refeições
R7	Incapaz de começar qualquer atividade por fadiga	Assistência máxima, dependência total ou não houve atividade relativamente a tarefas domésticas
R8	Incapaz de começar qualquer atividade por fadiga	Assistência máxima, dependência total ou não houve atividade relativamente à gestão das finanças pessoais
R9	Incapaz de começar qualquer atividade por fadiga	Assistência máxima, dependência total ou não houve atividade relativamente à gestão da medicação
R10	Incapaz de começar qualquer atividade por fadiga	Assistência máxima, dependência total ou não houve atividade relativamente à utilização do telefone
R11	Incapaz de começar qualquer atividade por fadiga	Assistência máxima, dependência total ou não houve atividade relativamente à utilização de escadas
R12	Incapaz de começar qualquer atividade por fadiga	Assistência máxima, dependência total ou não houve atividade relativamente à ida às compras
R13	Incapaz de começar qualquer atividade por fadiga	Assistência máxima, dependência total ou não houve atividade relativamente à utilização de transportes
R14	Incapaz de começar qualquer atividade por fadiga	Assistência máxima, dependência total ou não houve atividade relativamente tomar banho
R15	Incapaz de começar qualquer atividade por fadiga	Assistência máxima, dependência total ou não houve atividade relativamente à higiene pessoal
R16	Incapaz de começar qualquer atividade por fadiga	Assistência máxima, dependência total ou não houve atividade relativamente a caminhar
R17	Incapaz de começar qualquer atividade por fadiga	Assistência máxima, dependência total ou não houve atividade relativamente a transferências para a sanita
R18	Incapaz de começar qualquer atividade por fadiga	Assistência máxima, dependência total ou não houve atividade relativamente à utilização da casa de banho
R19	Incapaz de começar qualquer atividade por fadiga	Assistência máxima, dependência total ou não houve atividade relativamente à mobilidade na cama
R20	Incapaz de começar qualquer atividade por fadiga	Assistência máxima, dependência total ou não houve atividade relativamente a comer
R21	Não mostra sinais de dor	Intensidade da dor = Sem dor
R22	Intensidade da dor = Sem dor	Persistência da dor = Sem dor
R23	Recebe apoio da comunidade para a realização das tarefas domésticas	Tarefas domésticas no mínimo e assistência limitada
R24	Recebe apoio da comunidade em serviço de refeições	Preparação das refeições no mínimo e assistência limitada
R25	Sem atividade na alimentação	Sem atividade no modo de ingestão nutricional

Uma forma para verificar a consistência interna de um instrumento é através da utilização do α de Cronbach. No entanto, esta técnica não pode ser utilizada para avaliar a fidelidade do InterRAI-AC na sua globalidade

dado o seu carácter multidimensional. As variáveis que o constituem não devem ser somadas para formar um valor global e o valor do α de Cronbach pode ser fortemente afetado por uma grande quantidade de

variáveis, como é o caso. Contudo, podemos utilizar esta técnica para testar a consistência interna das várias escalas que o constituem, mas apenas nas que resultam do somatório de variáveis, pois as que são constituídas por algoritmos não garantem que cada variável contribua igualmente para a pontuação final (Wellens et al., 2011).

Tendo em conta a forma de constituição de cada escala, foi testada a consistência interna medida com o α Cronbach da Escala da Comunicação, da Escala das Atividades Instrumentais da Vida Diária, da Escala do Delírio, da Escala Curta de Classificação da Depressão e da Versão Curta da Escala das Atividades de Vida Diária, nos três momentos (Pré-Admissão, Admissão e Alta) (*Idem*). Outra forma de testar a consistência interna das escalas é verificar as correlações existentes entre as variáveis que as compõe e o *score* total, tendo sido utilizados os coeficientes de correlação *r* de Pearson, ρ de Spearman e o *V* de Cramer. Espera-se que as variáveis tenham uma correlação significativa e forte com a Escala à qual pertencem (Goodwin & Leech, 2003; Streiner & Norman, 2008).

Desenho do estudo e amostra

Este estudo é do tipo longitudinal descritivo correlacional. A colheita dos dados decorreu entre março e julho de 2012, em quatro hospitais de cuidados agudos da Região Centro de Portugal, em serviços de medicina e cirurgia gerais.

Foi solicitada a colaboração dos enfermeiros que constituíam as equipas dos serviços para a recolha dos dados. Em cada serviço, o enfermeiro chefe indicou dois enfermeiros para coordenarem o trabalho de recolha de informação, junto dos utentes e familiares, e para se articularem com a equipa de investigadores. O registo dos dados colhidos poderia ser efetuado em papel ou em suporte magnético. Para esta última hipótese, foi construído um aplicativo informático que foi utilizado nos serviços que o pretenderam.

Foram realizadas formações a todos os enfermeiros sobre o instrumento InterRAIAC e seu preenchimento. Para cada serviço foi destinado um elemento da equipa de investigação que esclarecia dúvidas e seletivamente recolhia dados de doentes para avaliar a consistência dos registos.

A amostra foi constituída por todos os doentes internados nos serviços durante o período da recolha dos dados, por um período superior ou igual a três dias.

Procedimentos éticos

Foi solicitada a autorização para a realização do estudo aos conselhos de administração dos hospitais participantes que, após parecer positivo das respetivas comissões de ética, a concederam.

A todos os doentes ou familiares foi solicitada a autorização para a participação no estudo através de uma declaração de consentimento. Esta autorização foi precedida de um esclarecimento sobre o estudo e os seus objetivos, assim como da confidencialidade dos dados e, ainda, da não existência de quaisquer custos ou prejuízos para os participantes.

Tratamento estatístico

As análises foram realizadas através do programa SPSS versão 19.0.

Foram utilizadas medidas de frequência e medidas descritivas como a média e desvio padrão. De acordo com o nível de mensuração das variáveis foram utilizados o teste de qui quadrado para avaliar a independência entre variáveis qualitativas, a análise da variância com um fator para a comparação de médias, os coeficientes de correlação *r* de Pearson e ρ de Spearman para a análise da relação entre variáveis, respetivamente, numéricas e categóricas, e o *V* de Cramer.

Resultados

Equivalência cultural e linguística

Deste processo que foi executado por um grupo de professores de enfermagem, todos doutorados e por um grupo de enfermeiros com experiência profissional de pelo menos 5 anos, em que todos eles falavam inglês, resultou um instrumento com ligeiras alterações em relação ao original, nomeadamente a exclusão do 14º dia de internamento como momento de registo, tendo em conta que o tempo médio de internamento global, nos hospitais portugueses com a mesma tipologia da amostra, é de 7,64 dias, tal como está expresso no Relatório Final do Grupo Técnico para a Reforma Hospitalar criado pelo Despacho N.º 10.601/2011 do Ministério da Saúde (Ministério da Saúde. Direção-Geral da Saúde, 2011).

Também se decidiu retirar a componente que diz respeito à medicação, dado que a análise desta informação não é útil para o estudo. Na secção relativa à avaliação do estado da pele, decidimos incluir mais

uma questão referente ao risco de adquirir úlcera de pressão. Esta variável pode constituir uma base de evidência da efetividade dos cuidados de enfermagem, na medida em que permite estabelecer uma relação entre o diagnóstico de risco e a ocorrência ou não de úlcera de pressão.

Por não serem consideradas relevantes para o estudo, foram retiradas da secção de identificação as seguintes questões: identificação da etnia/raça, do nome do médico responsável pelo doente, número de identificação nacional, da instituição/agência fornecedora do número e fontes de financiamento do internamento. Da secção relativa à admissão e história inicial, foi retirado o código correspondente à Residência de Saúde Mental dado que este tipo de valência não existe no nosso sistema de saúde.

Características da amostra de dados

Foram recolhidos um total de 1823 questionários. Depois de uma revisão em que se eliminaram os questionários muito incompletos ou que não cumpriam os critérios de inclusão, reunimos um total de 1764 questionários. Comparadas as amostras dos 4 hospitais verificámos, através do teste *qui-quadrado*, que nelas não existem diferenças significativas na proporção de homens e mulheres ($\chi^2 = 6,626$; $p = 0,085$). Em relação à idade verifica-se que a média global é de 70,78 anos e que através da ANOVA se verifica-se que não existem diferenças significativas entre as médias de idades nos quatro hospitais ($F = 0,604$; $p = 0,612$).

Pela análise da Tabela 2 é ainda possível verificar que a maioria dos doentes provém dos serviços de medicina com 64% e que existe uma proporção de 12,3% de indivíduos que vivem sozinhos.

TABELA 2 – Distribuição da amostra

Variável	Valores	Nº	%
Género	Masculino	849	50,0%
	Feminino	850	50,0%
Estado civil	Casados	944	55,0%
	Viúvos	495	28,8%
	Solteiros	194	11,2%
Idade	Média	70,8	
	Desvio padrão	17,0	
Língua materna	Português	1731	98%
Proveniência	Casa própria ou arrendada	1430	81,9%
	Lar residencial	194	11,1%
Formas de coabitação	Com o cônjuge/parceiro	530	30%
	Com cônjuge/parceiro e outros	377	21,4%
	Com filhos	268	15,4%
	Com não familiares	233	13,4%
	Sozinhos	214	12,3%
Enfermaria	Medicina	1129	64%
	Cirurgia	635	36,0%

Fiabilidade inter-juízes

Foram avaliados 151 itens distribuídos por 13 secções do InterRAI CA-PT, recolhidos em cada um dos três momentos - Pré-admissão, Admissão e Alta. A proporção de concordância foi quase perfeita em 44% (66) dos itens, foi substancial em 41% (62) dos itens, foi moderada em 14% (22) dos itens e foi pobre em 1% (1) dos itens, como podemos verificar na Tabela 3. Os itens com maior interesse são aqueles em que o avaliador interpreta os dados que recolhe do utente, como os referentes ao desempenho cognitivo, ao humor, às atividades de vida diária, às atividades

instrumentais de vida diária e à dor (Gray et al., 2008). Na Secção “Comportamento e Ânimo”, que inclui os itens referentes ao humor, a proporção de concordância em 80% dos itens foi superior a substancial, ou seja, os valores de K foram superiores a 0.61. Na secção “Estado Funcional”, que inclui os itens referentes às AVDs, e às AIVDs, a proporção de concordância em 94% dos itens foi superior a substancial. Na Secção Condições de Saúde, que inclui os itens referentes à dor, a proporção de concordância em 67% dos itens foi superior a substancial.

TABELA 3 – Concordância inter-juízes

Secção	Nº de variáveis por proporção de concordância (K)				Total variáveis testadas por secção
	Pobre <0,40	Moderado 0,41 – 0,60	Substancial 0,61 – 0,80	Excelente >0,80	
Identificação	0	0	0	4	4
Admissão e história inicial	0	0	1	5	6
Cognição	0	0	5	14	19
Comunicação, visão e audição	0	2	5	3	10
Comportamento e ânimo	0	1	6	3	10
Estado funcional	0	3	23	21	47
Continência	0	3	2	7	12
Condições de saúde	0	6	11	1	18
Estado nutricional e oral	0	1	1	1	3
Estado da pele	0	2	3	0	5
Tratamentos e procedimentos	0	0	3	1	4
Potencial da alta	1	4	2	5	12
Alta	0	0	0	1	1
Total de variáveis por proporção de concordância	1	22	62	66	151

Consistência Interna

O coeficiente de α de Cronbach para cada uma das escalas foi superior a 0.80, como podemos verificar na Tabela 4, e não aumentou com a eliminação de qualquer das variáveis constituintes das escalas.

Foi testada a correlação entre os itens que compõe a Escala Curta de Classificação da Depressão e a própria escala e os valores das correlações foram todos

superiores a 0.8. Quando testámos as correlações entre os itens que compõe a Escala da Comunicação e a própria escala, verificámos que as correlações são bastante altas, superiores a 0.90. Nas Escalas do Delírio, das Atividades Instrumentais de Vida Diária e a Versão Curta da Escala das Atividades de Vida Diária, em todas elas a força das correlações é superior a 0.8.

TABELA 4 – Correlação entre itens e o total da escala à qual pertencem e fiabilidade das escalas

Escala/Subescalas	r de Pearson	α de Cronbach		
		Pré-Admissão	Admissão	Alta
Curta de Classificação da Depressão		0,863	0,833	0,870
Com pouco interesse ou prazer em fazer as coisas de que gosta	0,864			
Ansioso, impaciente ou inquieto	0,856			
Triste, deprimido ou desesperado	0,877			
Comunicação		0,971	0,963	0,963
Ser entendido	0,986			
Compreender	0,986			
Delírio			0,946	0,956
Facilmente distraído	0,893			
Momentos de discurso desorganizado	0,914			
Função mental varia ao longo do dia	0,903			
Atividades Instrumentais de Vida Diária		0,963		0,968
Refeições	0,868			
Tarefas	0,870			
Finanças	0,918			
Medicação	0,918			
Telefone	0,885			
Escadas	0,901			
Compras	0,933			
Transporte	0,921			
Atividades de Vida Diária - Versão curta		0,977	0,948	0,971
Desempenho na higiene	0,938			
Desempenho no andar	0,943			
Uso de sanitário	0,961			
Desempenho na alimentação	0,876			

Validade de conteúdo

As percentagens de dados omissos por variável são irrelevantes, como podemos verificar na Tabela 5, ou seja, 13 variáveis (7%) não têm dados omissos, 104 (57%) têm percentagens de dados omissos inferiores a 1%, 57 variáveis (31%) têm percentagens de dados omissos entre 1% e 2% e apenas 9 variáveis (5%) têm percentagens de dados omissos superiores a

2%. As variáveis Altura e Peso foram as que tiveram percentagens mais altas de dados omissos, 12,9% e 13%, respetivamente, mas isto pode-se dever ao facto de estes dados serem difíceis de obter em situações de dependência funcional. Quanto aos três momentos de avaliação do InterRAI AC, o preenchimento foi idêntico e as percentagens de dados omissos não variam muito nos três momentos.

TABELA 5 – Percentagens de dados omissos em cada secção do InterRAI AC

Secção	% de Dados Omissos				Total de variáveis por secção
	0%	< 1%	1% - 2%	> 2%	
Identificação	3	0	1	3	7
Admissão e história inicial	0	2	4	0	6
Cognição	0	15	5	0	20
Comunicação, visão e audição	0	9	1	0	10
Comportamento e ânimo	9	0	3	0	12
Estado funcional	1	42	7	0	50
Continência	0	11	1	0	12
Condições de saúde	0	20	3	0	23
Estado nutricional e oral	0	1	1	3	5
Estado da pele	0	3	1	1	5
Tratamentos e procedimentos	0	0	18	1	19
Potencial da alta	0	0	12	1	13
Alta	0	1	0	0	1
Total de variáveis por % de dados omissos	13	104	57	9	183

Analisando a quantidade de códigos inválidos por variável, verificámos que das 214 variáveis analisadas, 67 (31%) tinham códigos inválidos, sendo que 31 destas tinham apenas um código inválido num só questionário preenchido, e nas restantes variáveis a quantidade de códigos inválidos foi considerada irrelevante.

Validade de constructo

Para testar a estrutura interna da versão portuguesa do InterRAI-AC, foram analisadas as relações correspondentes às 25 hipóteses por nós formuladas e apresentadas no Tabela 1. Como resultado, obtivemos a confirmação total da consistência (100% de consistência) com todas as relações verificadas.

Validade concorrente

Para além das relações entre os códigos das variáveis foram também analisadas sete correlações entre variáveis que se espera terem correlação com base na evidência da literatura, como se pode verificar na Tabela 6.

As sete correlações demonstraram ser significativas: estado funcional com desempenho cognitivo; estado funcional com Índice de Massa Corporal; estado funcional com atividade física; estado funcional com a visão (Stuck et al. como citado por Wellens et al., 2010); continência fecal com úlceras de pressão; úlceras de pressão com mobilidade na cama (Allman, 1997, *apud* Wellens et al., 2010) e úlceras de pressão com alterações da consciência (Maklebust e Magnan, 1994, *apud* Wellens et al., 2010).

TABELA 6 – Validade baseada na correlação entre variáveis

Variável 1	Variável 2	N	r	Valor de p
Desempenho cognitivo	Estado funcional	1732	0,759	0,000
Índice de Massa Corporal	Estado funcional	1507	-0,087	0,001
Atividade física	Estado funcional	1743	0,473	0,000
Visão	Estado funcional	1740	0,471	0,000
Continência fecal	Úlcera de pressão	1751	0,442	0,000
Mobilidade na cama	Úlcera de pressão	1749	0,429	0,000
Úlcera de pressão mais grave	Alterações da consciência	1703	0,247	0,000

Discussão

A tradução, validação e adaptação cultural de um instrumento num idioma diferente daquele em que é produzido revela-se importante, uma vez que ajuda a desenvolver os métodos de validação e o próprio instrumento, pois é pela comparação dos resultados da utilização que se vai construindo a validação do instrumento em relação à sua pertinência e utilidade (Wellens et al., 2010). Por outro lado, é sempre uma tarefa desafiante que exige uma metodologia adequada para que a sua utilização seja segura e a possibilidade de erro seja minimizada. Por outro lado, uma vez que os contextos são variáveis, os processos de adaptação e validação são essenciais para comparar resultados entre instituições, entre sistemas de prestação de cuidados e entre países.

No caso dos instrumentos da organização InterRAI, particularmente do InterRAI AC, esse processo é ainda mais desafiante e complexo, uma vez que se trata de instrumentos compreensivos, multidimensionais que necessitam de ser avaliados como um todo e transversalmente a todos os domínios, em todos os momentos, e em cada item (Hirdes et al., 2008).

Tendo em conta o carácter do instrumento, seguiu-se a opinião e as recomendações de Wellens et al. (2010, 2011) quer para a avaliação da validade, quer para a da fiabilidade.

A avaliação da validade do constructo é habitualmente feita com ajuda da análise fatorial, cujo objetivo é a redução de um conjunto grande de variáveis para um conjunto mais pequeno com características comuns ou a criação de dimensões subjacentes. Uma vez que o InterRAI AC tem as características de um resumo mínimo de dados, em que cada domínio já está identificado, tem já um número mínimo de itens e não é calculado um *score* global, Wellens et al. (2010) consideram que não é apropriado na validação fazer

a redução de variáveis através das técnicas de análise fatorial.

Os resultados que obtivemos nesse processo permitem dizer que a versão que resultou é mais um passo na consolidação do InterRAI AC como um instrumento capaz de ajudar os profissionais de saúde na obtenção de dados sobre os doentes que permitam a produção de informação e conhecimento para a qualidade de cuidados e para a determinação de políticas que permitam uma organização mais centrada na pessoa e nos resultados das intervenções, para além da monitorização e validação de dados produzidos por outros meios e instrumentos.

Quanto aos dados obtidos verificámos que a média de idade da amostra é de 70.78 anos, o que corresponde a uma população idosa, a maioria a viver na sua residência.

No que diz respeito à validação de conteúdo e à adaptação cultural, todos os peritos foram unânimes quanto ao valor deste instrumento no seu conjunto, quanto aos itens que compõem e quanto à sua utilidade.

No que toca à equivalência da informação, quando foram obtidos e registados dados acerca do mesmo doente por juizes diferentes, registaram-se concordâncias substanciais na maioria dos casos. Nas situações em que é necessária uma capacidade interpretativa maior, por parte do juiz, verificaram-se níveis de concordância em 80% dos itens superiores a substanciais (> 0.61) no Comportamento e ânimo. No que toca aos itens relacionados com as AVDs e com as AIVDs, o nível de concordância inter-juizes foi também superior a substancial em 94% dos itens. Estes dados estão em linha com os obtidos por Hirdes et al. (2008).

No que diz respeito à consistência interna das escalas que produzimos, em todas elas foi encontrado um valor superior a 0,80, sendo que a maioria apresentava

valores superiores a 0,90, o que significa uma forte consistência interna das escalas e o que contribui para uma boa fiabilidade do instrumento.

Em relação à validade baseada na estrutura interna, analisada a partir da relação entre codificações de variáveis relacionadas (Wellens et al., 2010), que corresponde a 25 hipóteses que formulámos nesse sentido, verificámos em todas elas 100% de consistência. Este resultado é, na perspectiva das autoras, essencial para a validação de instrumentos complexos como o InterRAI AC.

Obtivemos também valores de correlação significativos entre variáveis que a teoria diz estarem relacionadas. Em relação a estas correlações deve referir-se que a sua significância estatística se deve em grande medida ao tamanho da amostra, porquanto os valores de correlação são em alguns casos <70 o que significa que o grau de variação de uma variável que é explicada pela variação de outra variável nem sempre é grande. Exemplo do exposto são as correlações entre: índice de massa corporal e estado funcional ($r=-0.087$); atividade física e estado funcional ($r=0.473$); visão e estado funcional ($r=0.471$). Apesar disso devemos considerar o contributo e a necessidade de se ter intervenção nesses atributos para obter ganhos na atividade física, esse é um dado a reter nesta validação.

Limitações

O tempo de demora com o preenchimento de cada questionário (em média 15 minutos) revelou-se um fator limitativo deste estudo, pois em circunstâncias em que o tempo já é escasso para os cuidados, a disponibilidade dos enfermeiros para o seu preenchimento é menor, o que tem repercussões no tamanho da amostra.

Conclusão

As análises realizadas no processo de validação do InterRAI AC, versão portuguesa, demonstram que este instrumento é válido e fiável, sendo portanto uma excelente ferramenta para ser utilizada nos contextos de prestação de cuidados em serviços de cuidados agudos.

Este instrumento pode ser muito útil na obtenção de dados para diversos fins como indicadores de qualidade, planeamento de cuidados, base de

evidência para a tomada de decisão, entre outros.

Para além da sua utilização pontual no âmbito de algum trabalho de investigação, o InterRAI Cuidados Agudos, versão portuguesa, poderia ser integrado como ferramenta fundamental no processo de recolha de informação sobre os pacientes, pois este instrumento fornece uma visão holística da história e da evolução do paciente, possibilitando assim prestar cuidados mais documentados.

O principal obstáculo na aplicação do InterRAI Cuidados Agudos, versão portuguesa, é o tempo necessário para o seu preenchimento. Serão necessários mais estudos para avaliar se o tempo gasto na aplicação deste instrumento pode ter como retorno a melhoria da efetividade e qualidade dos cuidados.

Artigo científico produzido no âmbito do projeto de investigação “Resultados dos Cuidados de Enfermagem: Qualidade e Efetividade” (FCT-PTPC/CS-SOC/113519/2009) inscrito na Unidade de Investigação em Ciências da Saúde – Enfermagem.

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, Canadá: 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. Retirado de <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. Direcção 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*. Retirado de 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* (4ª 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.

