

Propriedades Psicométricas do *Action Research Arm Test* - Revisão Sistemática da Literatura -

RESUMO

A parésia do membro superior é um dos sintomas mais prevalentes após a ocorrência de acidente vascular cerebral (AVC) e de extrema relevância para a pessoa afetada. O Action Research Arm Test (ARAT) é um dos instrumentos mais utilizado na avaliação da funcionalidade do membro superior. O objetivo deste estudo é conhecer as propriedades métricas do ARAT aplicado a pessoas com AVC. Foi feita uma revisão sistemática da literatura (RLS), acedeu-se a bases de dados eletrônicas e pesquisou-se artigos no últimos 5 anos. Seis artigos cumpriram os critérios definidos para inclusão no estudo. Todos os artigos incluídos avaliaram uma ou mais propriedades métricas do instrumento. O ARAT apresenta boas propriedades psicométricas, a nível da reprodutibilidade, da validade e da responsividade, em pessoas com alterações motoras do membro superior após um AVC.

Palavras-chave: Acidente Vascular Cerebral, Action Research Arm Test, Validade.

ABSTRACT

Motor deficit of the upper extremity following stroke is a prevalent symptom and considered a major disability for individuals who experience the loss of upper limb functioning. The ARAT is a widely used measure of upper limb extremity motor function after stroke. We performed systematic review of the literature through electronic databases in the last 5 years. Six studies respect the criteria for inclusion. All studies assessed more than one psychometric property. The ARAT has good psychometric properties in stroke patients with upper limb motor deficit and has evidence of reliability, validity and responsiveness.

Keywords: stroke, Action Research Arm Test, Validity.

INTRODUÇÃO

O Acidente Vascular Cerebral (AVC) caracteriza-se pela ocorrência aguda de uma disfunção neurológica decorrente de uma anomalia na circulação cerebral com o surgimento de sintomatologia de acordo com a área do cérebro afetada (Damasceno & Borges, 1988). Duncan *et al.* (2005) referem que um número considerável de pessoas com sequelas de AVC necessitam de intervenções de reabilitação focalizadas, por um período de tempo variável, especialmente ao nível do membro superior afetado. A título de exemplo num primeiro momento pós AVC a função do membro superior é alterada em 73 a 88% dos sobreviventes, em que 55 a 75% destes permanecem com alteração da função durante três a seis meses após a ocorrência da lesão (Cirstea & Levin, 2007).

Para avaliação das condições sensoriais, motoras e funcionais do membro superior foram desenvolvidas diversas escalas, entre as mais referenciadas estão a Escala de Desempenho Físico de Fugl-Meyer (UE-FMA), o Teste de Função Motora de Wolf (WMFT), a escala STREAM (Stroke Rehabilitation Assessment of Movement), bem como o Action Research Arm Test (ARAT), sendo esta última alvo do nosso estudo. Segundo Wade (1992), os instrumentos supracitados fornecem informações sobre o desempenho do membro em causa, evitando a heterogeneidade de medida e a subjetividade do auto-relato. Rosamond *et al.* (2008) reforçam que a capacidade de alcance direcionado, preensão e manipulação são componentes que correspondem ao normal funcionamento do membro superior e que formam a base da capacidade motora necessária à eficaz execução das atividades de vida diárias.

A escala da Função da Extremidade Superior desenvolvida por Carroll, em 1965, no entanto, a sua versão resumida surge posteriormente como, Teste da Ação da Extremidade Superior (Rabadi & Rabadi, 2006). O ARAT foi validado na língua inglesa num estudo com 351 participantes (Koh *et al.*, 2006). Este instrumento “avalia as complexas atividades da extremidade superior sobre o pressuposto de que todas as atividades funcionais podem ser sintetizadas em quatro tipos básicos de função: compressão, preensão, pinçamento e atividades de alcance (função motora grossa)” (Lyle, 1981). A escala

foi concebida com 19 itens, divididos em 4 sub-escalas, em que 6 itens avaliam a capacidade de apertar, 4 itens a capacidade de segurar, 6 itens avaliam a pinça e 3 itens avaliam a motricidade grosseira. A pontuação oscila entre os 0 e os 57 pontos e quanto mais elevado o score melhor o desempenho da pessoa (Koh *et al.*, 2006). Uma vez que a escala está disposta por ordem crescente de dificuldade, a avaliação pode ser interrompida quando uma pontuação de “zero” é atingida, não necessitando o indivíduo de realizar provas mais complexas, as quais não teria capacidade de resolução. Desde que o examinador conheça previamente o quadro funcional da pessoa, apenas alguns itens podem ser suficientes para determinar um score final (Platz, Pinkowisk, Wijck & Johnson, 2005).

O objetivo de estudo consiste em identificar as propriedades métricas do ARAT aplicado em pessoas com alteração da capacidade motora do membro superior na sequência de um AVC, através de uma revisão sistemática da literatura.

METODOLOGIA

Foi feita uma revisão sistemática da literatura (Galvão, Sawada & Trevizan, 2004; Pereira & Bachion, 2006; Santos, Pimenta & Nobre, 2007). Esta investigação foi conduzida pelo método PICO – *Participant* (Tipo de participantes); *Intervention* (Tipo de intervenção); *Comparasion* (Comparação); *Outcomes* (Tipo de resultados), como se pode verificar na tabela nº1.

Tabela nº 1 – Método PICO

Participantes	Pessoa que sofreu AVC com limitação da extremidade superior.
Intervenção	Validação / adaptação do Action Research Arm Test (ARAT)
Comparação	Doentes com diferentes limitações do membro superior após AVC
Resultados	Reprodutividade / Validade / Responsividade

Posto isto, a questão de investigação será “ **quais as propriedades métricas do ARAT na pessoa com comprometimento motor do membro superior após AVC?**”.

A revisão sistemática da literatura partiu em função da questão de investigação e dos descritores/palavras-chave mencionados na tabela 2. Estes foram validados na plataforma MESH e/ou na DeCS, à exceção do descritor “Action Research Arm Test”.

Tabela nº 2 – Método PICO, Palavras-chave e Descritores validados para a pesquisa

	Questões de partida	Palavras-chave	Descritores
Participantes (P)	Pessoa com comprometimento motor do membro superior após AVC		AVC; Stroke.
Intervenção (I)	Validação; adaptação do Action Research Arm Test	“Action Research Arm Test”,	Avaliação; Psychometrics.
Outcomes (O)	Reprodutibilidade; Validade; Responsividade.		Reprodutibilidade; Reliability; Validade.

A investigação foi feita durante o mês de Outubro de 2014, por dois investigadores em períodos distintos, através da consulta das seguintes bases de dados: CINAHL Complete, MEDLINE Complete, Nursing & Allied Health Collection: Comprehensive, Database of Abstracts of Reviews of Effects, Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, Cochrane Methodology Register, Library, Information Science & Technology Abstracts, MedicLatina, Health Technology Assessments, NHS Economic Evaluation Database. EBSCOhost e Biblioteca Virtual em Saúde foram os motores de busca, ou seja, as plataformas informáticas através das quais foram feitos os acessos.

Tabela nº 3 – Critérios de inclusão e exclusão

Crítérios de selecção	Crítérios de inclusão	Crítérios de exclusão
Participantes (P)	Pessoa com alterações da mobilidade do membro superior devido ao AVC.	Pessoa com outra patologia do foro Neurológico.
Intervenção (I)	Reprodutibilidade; Validade; Responsividade.	Não possuir, pelo menos, 1 destes três critérios.
Desenho do estudo (S)	Estudo quantitativo.	Revisão sistemática da literatura; Estudo qualitativo.
Período de publicação	Artigo publicado entre 2009 e 2014.	Qualquer artigo fora deste período de tempo.
Língua em que está publicado	Artigo publicado em Português, Inglês e Espanhol.	Artigo publicado.
Disponibilidade do documento	Artigo completo e de acesso livre.	Artigo incompleto ou que seja necessário pagar a sua aquisição.

Dado o número de estudos/artigos encontrados e a especificidade do tema, apresenta-se na tabela nº 3 os critérios de inclusão e exclusão do presente estudo, com o intuito de direcionar e facilitar a pesquisa e seleccionar os estudos/artigos que posteriormente serão relevantes para a construção deste.

Da pesquisa feita obteve-se os resultados enumerados na tabela 4, através da conjugação de descritores. Cada conjugação foi feita por dois elementos do grupo de trabalho, sendo a seleção dos artigos/estudos também validado pelos dois. Os artigos incluídos foram avaliados quanto à qualidade metodológica através dos critérios do JBI para Estudos descritivos/ Estudos de series de Casos (JBI, 2011, p.181). Em suma, a seguinte tabela sintetiza quantitativamente as pesquisas realizadas e do resultado final constam apenas seis artigos (Figura 1).

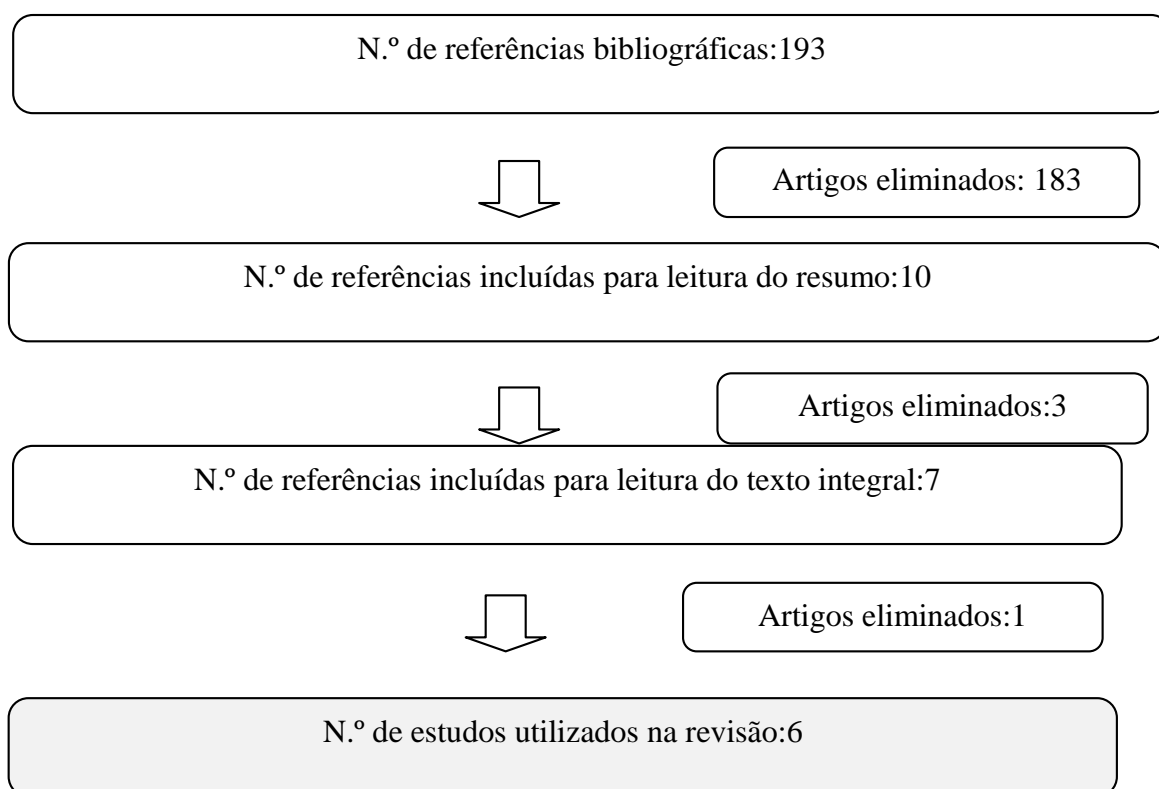
Tabela nº 4 – Conjugação Boleana

Conjugação Boleana	Resultados da pesquisa	
	EBSCOHost	Biblioteca Virtual em Saúde
"Action Research Arm Test" and stroke	63	130
"Action Research Arm Test" and stroke and assessment	45	63
"Action Research Arm Test" and stroke and reliability	11	19

“Action Research Arm Test” and psychometrics	2	8
“Action Research Arm Test” and stroke and validity	11	18
“Action Research Arm Test” and stroke and reproductibility	0	0
Total	132	238

No sentido de facilitar o entendimento da pesquisa realizou-se um fluxograma, que aparece discriminado na figura 1.

Figura nº 1 – Fluxograma da pesquisa realizada



O estudo das propriedades métricas foi realizado com base nos critérios de Validade, Reprodutibilidade e Responsividade (Barbetta & Assis, 2008; Leung, Trevena & Waters, 2012; Marque-Vieira *et al*, 2015; Sousa *et al*, 2015).

DISCUSSÃO DOS RESULTADOS

Foram encontrados 6 artigos que cumpriam os critérios de inclusão definidos para este estudo. Na avaliação da qualidade metodológica, todos os artigos obtiveram um escore \geq a 7, cumprindo assim, pelo menos, 77% dos critérios definidos pela JBI. Atendendo aos critérios de classificação das evidências proposta por Belsey & Snell (2001) os artigos incluídos na RLS apresentam um nível de evidência III, ou seja, evidências de estudos experimentais bem delineados, mas não randomizados, estudos de coorte, séries temporais, estudos de caso-controle combinados.

Efetuuou-se a recolha de informação sobre o ano, país, autor, participantes, intervenções, resultados e níveis de evidência apresentados na tabela nº5 ordenados pela data da publicação dos artigos.

Tabela nº5 - Descrição dos estudos incluídos RLS

Autor, Ano, País	Participantes	Intervenção	Comparação	Resultados	Nível de evidência
Nordin <i>et al</i> , 2014, Suécia	35 participantes com comprometimento do membro superior função após acidente vascular cerebral.	Avaliar a fiabilidade intra e inter-observador do Action Research Arm Test (ARAT).	Action Research Arm Test (ARAT)	Elevada fiabilidade intra- e inter-observadores. Fiabilidade intra-observador ligeiramente superior à inter-observador.	III
Chen <i>et al</i> , 2012, EUA	191 participantes com acidente vascular cerebral com gravidade ligeira a moderada e sem compromisso cognitivo grave	Avaliar a construção interna e validade do Action Research Arm Test (ARAT).	Action Research Arm Test (ARAT)	ARAT possui boas propriedades psicométricas e tem provas de unidimensionalidade, validade preditiva e de fiabilidade inter- e intra-observadores.	III

Arlette, Rodrigo, & Viviana, 2012, Chile	80 Pessoas com alteração da mobilidade ao nível do membro superior após acidente vascular cerebral, ajustada a população chilena	Avaliar a fiabilidade intra-observador e a validade da ARAT.	ARAT	A evidência indica que a ARAT é um instrumento fiável e válido para avaliar a recuperação funcional do membro superior depois de uma lesão cortical em pacientes Chilenos com sequelas de AVC. Boa validade convergente e discriminativa.	III
Nijland, R. <i>et al</i> , 2010, Holanda	40 pessoas com parésia membro superior com pelo menos contracção muscular (Medical Research Council score ≥ 1) após acidente vascular cerebral	Avaliar a validade concorrente do Action Research Arm Test (ARAT) e Wolf Motor Function Test (WMFT) e comparar a sua reprodutibilidade, consistência interna e efeito teto e chão.	Action Research Arm Test (ARAT) e Wolf Motor Function Test (WMFT).	Elevada fiabilidade inter- e intra-observador do ARAT e do WMFT. Elevada validade concorrente entre a ARAT e o WMFT.	III
Lin <i>et al</i> , 2010, China	59 pessoas com alteração da mobilidade ao nível do membro superior após acidente vascular cerebral	Comparar a responsividade e validade entre o Box and Block Test (BBT), o Nine-Hole Peg Test (NHPT) e o Action Research Arm Test (ARAT).	Box and Block Test (BBT), Nine-Hole Peg Test (NHPT) e Action Research Arm Test (ARAT),	O ARAT tem uma boa validade concorrente, semelhante ao BBT, mas superior ao NHPT. Simultaneamente considerando a responsividade e a validade, a BBT e ARAT podem ser considerados mais apropriados para avaliar a destreza manual que o NHPT.	III

Lin, <i>et al.</i> , 2009, China	Participaram 53 pessoas com acidente vascular cerebral dos quais 35 completaram todas as fases do estudo.	Comparar a reprodutibilidade, validade e responsividade da subescala da extremidade superior do Fugl-Meyer Motor Test (FM-ES), da subescala da extremidade superior do Stroke Rehabilitation Assessment of Movement (STREM-ES), do Action Research Arm Test (ARAT) e do Wolf Motor Function Test (WMFT).	Fugl-Meyer Motor Test, Stroke Rehabilitation Assessment of Movement, Action Research Arm Test (ARAT) e Wolf Motor Function Test (WMFT).	Fiabilidade inter- e intra-observador elevada nos 4 instrumentos elevada. Elevada validade concorrente . Os autores afirmam que os resultados são suficientes para suportar a hipótese de que os 4 instrumentos têm validade preditiva. Responsividade moderada dos 4 instrumentos, entre os quais o ARAT foi o mais responsivo.	III
----------------------------------	---	--	---	--	-----

Todos os estudos incluídos na RSL avaliaram uma ou mais propriedades psicométricas da ARAT, os critérios de inclusão e exclusão das pessoas para a formação da amostra são semelhantes e procuram minimizar o enviesamento dos resultados. Alguns dos critérios mais relevantes são: pessoas que sofreram AVC (todos os estudos), primeiro episódio de AVC (Lin *et al.*, 2009; Chen *et al.*, 2012), pessoa com AVC há mais de 6 meses (Arlette, Rodrigo & Viviana, 2012; Lin *et al.*, 2010), parésia do membro superior com pelo menos contração muscular (Nordin, *et al.*, 2014; Nijland *et al.*, 2010), aplicação da escala de Brunnstrom com nível >2 (Chen *et al.*, 2012) e > 4 (Lin *et al.*, 2010) pontuação da Modified Ashworth Scale < 2.5 (Lin *et al.*, 2010) e <3 (Arlette, Rodrigo & Viviana, 2012), capacidade mental para seguir as instruções avaliada através da Mini-mental State Examination (Nordin, *et al.*, 2014; Chen *et al.*, 2012; Arlette, Rodrigo & Viviana, 2012; Lin *et al.*, 2010; Nijland *et al.*, 2010 e Lin *et al.*, 2009); pessoas com limitações no membro superior exclusivas do AVC (Nordin *et al.*, 2014; Chen *et al.*, 2012; Nijland *et al.*, 2010).

A investigação desenvolvida por Nordin, *et al.* (2014) teve como objetivo avaliar a fiabilidade intra e inter-observador do ARAT em pessoas após AVC. Participaram 35 pes-

soas em pós-AVC que apresentavam comprometimento da função do membro superior, pelo menos seis semanas após o início do AVC e 18 ou mais anos de idade, ausência de movimento ativo no braço afetado, não possuísem alterações da função do membro superior não relacionados com o acidente vascular cerebral, e capacidade de seguir as instruções ou entender sueco. Dois fisioterapeutas trabalharam simultaneamente, mas de forma independente, na avaliação das pessoas ao longo dos 19 itens do ARAT, duas vezes por dia, com o papel de observadores do estudo. Para avaliar a fiabilidade inter- e intra-observador utilizaram o coeficiente kappa de Cohen e o ranking de Svensson especialmente direcionado para a análise de divergências em dados ordinais emparelhados (Svensson, 2012). Os métodos estatísticos utilizados neste estudo são pouco utilizados em análises desta natureza, no entanto, os autores defendem que o ranking de Svensson dá informações mais precisas sobre a origem das divergências encontradas e possibilidade de avaliar se as mesmas comprometem a fiabilidade do instrumento.

Na avaliação da fiabilidade intra-observadores, para a pontuação total do ARAT, foi demonstrada a concordância completa em 33 das 35 avaliações de cada um dos avaliadores. Na fiabilidade inter-observadores a concordância completa foi encontrada em 32 e 31 dos 35 pares de avaliações, na primeira e segunda aplicação do teste, respetivamente. Detetaram sistematicamente ligeira discordância nos itens 10,11,14 e 19 intra-observadores e nos itens 1,4,17 e 19 inter-observadores. No entanto, o estudo demonstra que o ARAT é um instrumento fiável para avaliar as funções e actividades do membro superior em pessoas após AVC. A fiabilidade intra-observador foi ligeiramente melhor que a inter-observadores. As ligeiras divergências encontradas intra- e inter- observadores foram originadas por mudanças no desempenho dos participantes ou no uso diferente das categorias de escala entre os avaliadores. Os autores realçam que na utilização do ARAT em ambientes clínicos e de pesquisa, é importante estar ciente que estas pequenas divergências poderão ocorrer.

Chen *et al.* (2012) avaliaram a validade de construto e a validade preditiva do ARAT. Participaram 191 pessoas que sofreram AVC. As pessoas foram divididas em três grupos, o que recebeu treino intensivo com restrição de uso do membro não afetado, o que recebeu treino intensivo de ambos os membros superiores e um terceiro de controlo que recebeu técnicas de neurodesenvolvimento. Três fisioterapeutas treinados avaliaram as

pessoas de cada grupo, sem conhecimento do tipo de técnicas utilizadas na reabilitação, no início do estudo, pré-tratamento, e 3 a 4 semanas depois, pós-tratamento. Selecionaram três instrumentos para examinar a validade preditiva da ARAT: o WMFT, utilizado para avaliar a função motora da extremidade superior; o Motor Activity Log (MAL) utilizado para avaliar a auto-percepção da pessoa na utilização do membro em Atividades de Vida Diárias (AVDs); e do Stroke Impact Scale (SIS), que percebe o estado geral de saúde em pessoas com AVC.

Através da análise Rasch foi demonstrado que o ARAT é um instrumento unidimensional (Análise da Componente Principal de 72,5%), mas apresenta redundância na hierarquização das categorias (escala de 1 a 4), e de acordo com a análise efetuada neste estudo os autores sugerem que os itens do ARAT utilizem uma escala de 3 pontos. Apesar de ter sido aplicada uma análise Rasch na escala reduzida a 3 pontos, deverá ser realizada mais investigação neste sentido antes de ser efetuada a redução da escala, uma vez que os itens "levar da mão atrás da cabeça" e o "levar da mão em topo da de cabeça" induziram a maiores discrepâncias de análise inter-observadores após a redução da escala. O Coeficiente de Pearson (0.94) e o Alfa de Cronbach (0.97) confirmam a elevada fiabilidade do ARAT. Por outro lado, a validade preditiva foi analisada pelo Coeficiente de Correlação de Spearman, e na sua análise entre o ARAT pré-tratamento e o SIS pós-tratamento revelou-se satisfatória (0.45), as correlações entre o ARAT e o WMFT Time (-0.66) e o MAL (0.58) e o SIS-*hand function* (-0.66) revelaram-se moderadas a boas. Já a correlação entre a ARAT e a WMFT- *functional ability scale* (0.76) foi muito boa. Os autores concluem que o ARAT possui boas propriedades psicométricas em pessoas com AVC com alterações motoras de gravidade ligeira a moderada e sem comprometimento cognitivo grave, nomeadamente unidimensionalidade, validade preditiva e fiabilidade. No entanto, apesar de revelar uma validade preditiva satisfatória, é necessária mais investigação em que os instrumentos sejam aplicados em diferentes momentos ao longo da reabilitação.

Também a investigação desenvolvida por Arlette, Rodrigo & Viviana (2012) teve como objectivo determinar as propriedades psicométricas do ARAT e incluiu uma amostra de 80 pessoas com acidente vascular cerebral e idades compreendidas entre os 18 e 80 anos no Chile. Aplicaram o ARAT, a MAL às 80 pessoas em dois períodos, uma medição

inicial e outra passados dois meses. A avaliação da qualidade de vida apenas foi efetuada na segunda avaliação, utilizando a SIS e o General Health Questionnaire (GHQ). As pessoas não receberam qualquer intervenção entre as duas avaliações. Os autores adotaram métodos estatísticos convencionais, para avaliar a consistência interna do ARAT calcularam o alfa de Cronbach no início do estudo e dois meses depois, 0.88 e 0.89, respectivamente, demonstrando ser fiável. A validade convergente foi avaliada através do coeficiente de correlação de Pearson entre a ARAT e a idade (-0.53), a percepção da dor (-0.38), a MAL (entre 0.51 e 0.58 na quantidade e qualidade do movimento, respectivamente), a SIS (variaram entre 0.41 e 0.57 nas diferentes categorias da escala). O instrumento em análise apresenta boa validade convergente, os níveis de correlação com a MAL demonstram que o ARAT avalia tanto as AVDs básicas como instrumentais, já a correlação inversa significativa com a idade e a percepção de dor demonstra que as funções do membro superior são negativamente influenciadas por estes fatores. Os coeficientes de correlação de Pearson entre o ARAT e as funções emoção, memória e comunicação do SIS e para o nível socioeconómico da GHQ não foram significativos (0.08, -0.08, -0.26 e -0.17, respectivamente), contudo o ARAT apresenta validade discriminativa.

A investigação desenvolvida por Nijland *et al.* (2010) teve como objetivo analisar a validade concorrente entre o ARAT e o WMFT e comparar a reprodutibilidade, consistência interna e efeito chão e teto dos dois instrumentos numa amostra de pessoas com acidente vascular cerebral. A amostra foi constituída por 2 grupos de pessoas com acidente vascular cerebral de dois centros de reabilitação na Holanda para a qual os critérios de inclusão foram: hemiparesia do membro superior com alguma contração muscular (Medical Research Council score ≥ 1), não apresentarem défices severos na comunicação, memória e compreensão, pelo que foi aplicado o Mini Mental State Examination (pontuação > 22) e ainda não apresentarem limitações de foro ortopédico no membro superior.

Nijland *et al.* (2010) utilizaram as pessoas do primeiro centro de reabilitação (n=18) para testarem a reprodutibilidade de ambos os instrumentos. Na fiabilidade inter-observador, ambos os observadores aplicaram os instrumentos com menos de uma semana de intervalo para minimizarem os efeitos da recuperação espontânea. No teste da

fiabilidade intra-observador, a mesma amostra (n=18) foi observada duas vezes pelo observador no espaço de 10 dias aproximadamente. Para a análise da consistência interna, validade concorrente e efeitos de chão e teto foram recolhidos dados da amostra total (n=40), sendo que neste caso a aplicação dos instrumentos foi efetuada por um observador experiente e com uma pausa para descanso de 30 min entre cada avaliação para prevenir efeitos da fadiga sobre a avaliação.

Na análise da fiabilidade inter-observador os resultados do Coeficiente de Correlação Intraclases (CCI) do ARAT e do WMFT foram de 0.92 e 0.94, respetivamente. Na análise da fiabilidade intra-observador do ARAT e do WMFT foram de 0.97 e 0.95, respetivamente. Estes resultados indicam que existe uma elevada fiabilidade inter- e intra-observador, e que estão altamente relacionados entre eles.

A análise da consistência interna entre os itens das escalas de ARAT e WMFT foi dada pelo alfa de Cronbach, 0.985 e 0.982 foram os resultados obtidos para o ARAT e o WMFT, respectivamente, e confirmam a elevada consistência interna dos instrumentos, no entanto, estes valores sugerem que pode haver redundância, ou seja, existem itens que estão a medir o mesmo elemento do constructo. Não foram encontrados efeito chão e teto significativos em ambos os instrumentos.

Para analisar a validade concorrente foi calculado o coeficiente de correlação de Spearman (r_s) para a pontuação total do ARAT e do WMFT dividido em quatro grupos de variáveis: Escala de Habilidade Funcional (EHF), pontuação do tempo médio, item 7 e item 14. O coeficiente de correlação de Spearman (r_s) entre o ARAT e EHF foi de 0.86 ($p<0.01$), entre o ARAT e a pontuação do tempo médio foi de -0.89 ($p<0.01$), entre o ARAT e os itens 7 e 14 foram ambos de 0.70 ($p<0.01$). Estes valores indicam uma elevada relação entre os dois instrumentos no que concerne ao constructo que está a ser medido.

O estudo conduzido por Lin *et al.* (2010) pretende comparar a responsividade e validade entre o Box and Block Test (BBT), o Nine-Hole Peg Test (NHPT) e o Action Research Arm Test (ARAT), contando com uma amostra de 59 participantes, selecionados aleatoriamente, que sofreram um AVC, em 3 tipos de tratamentos (terapia intensiva com restrição de uso do membro não afetado, uso simultâneo dos membros superiores e trata-

mento de controlo, durante três semanas. Foram ainda utilizadas três escalas Fugl-Meyer (FMA), Motor Activity Log (MAL), e Stroke Impact Scale (SIS) para testar a validade concorrente. O BBT, NHPT e ARAT demonstraram uma resposta moderada às mudanças e não significativamente diferentes (Resposta Média Standardizada = 0.64 a 0.79). As correlações entre BBT, NHPT e ARAT tiveram resultados moderados a bons em pré-tratamento (coeficiente de correlação de Spearman = -0.55 a -0.80) e pós-tratamento (-0.57 a -0.71). O BBT e ARAT mostraram correlações baixas a moderada com o FMA, MAL, e SIS na função de domínio da mão no pré e pós-tratamento (0.31-0.59), por seu lado o NHPT demonstrou correlações baixas a moderadas com o FMA e MAL (-0.16 a -0.33) e correlações moderadas entre a função do domínio da mão do SIS (-0.58 a -0.66). Os resultados indicam que a BBT, NHPT e ARAT são moderadamente responsivos (standardize response mean = 0.74, 0.64 e 0.79, respetivamente) a mudanças ao longo do tempo, sem diferenças significativas entre os três instrumentos. O ARAT tem uma boa validade de constructo e uma boa validade concorrente, semelhante ao BBT, mas superior ao NHPT. Simultaneamente considerando a responsividade e validade, a BBT e ARAT podem ser considerados mais apropriados para avaliar a destreza manual que o NHPT. No entanto, os autores admitem que estudos com amostras maiores são necessários para validar os resultados encontrados.

O estudo de Lin *et al.* (2009) teve como objetivo comparar a reprodutibilidade (fiabilidade inter- e intra-observador), validade (validade concorrente e validade preditiva) e responsividade da subescala da extremidade superior do Fugl-Meyer Motor Test (FM-ES), da subescala da extremidade superior do Stroke Rehabilitation Assessment of Movement (STREM-ES), do Action Research Arm Test (ARAT) e do Wolf Motor Function Test (WMFT).

O estudo foi dividido em duas fases. Na primeira fase do estudo, a fiabilidade inter-observador, validade e responsividade foram testados em pessoas com AVC no departamento de neurologia do Kaohsiung Medical University Hospital. Participaram 53 pessoas com acidente vascular cerebral, os instrumentos de avaliação foram aplicados em 4 momentos diferentes (14, 30, 90 e 180 dias), completaram a 4 avaliações 35 participantes. Na segunda fase do estudo, a fiabilidade intra-observador foi testada. Para tal foi recrutada uma amostra independente de pessoas com AVC em fase crónica recebendo

intervenção de reabilitação em ambulatório. Os critérios de inclusão foram pessoas com AVC há mais de 1 ano. As escalas FM-ES, ARAT e WMFT tiveram um efeito chão significativo ($\geq 21\%$ dos participantes) na avaliação aos 14 dias depois do AVC e um elevado efeito teto ($\geq 21\%$ dos participantes) aos 30, 90 e 180 dias após o AVC. Foi aplicado o coeficiente de correlação de Spearman entre os 4 testes, os resultados (0.81 a 0.97) mostram uma elevada validade concorrente.

A pontuação dos 4 testes aos 14 dias após o AVC estão moderadamente correlacionados a pontuação do Índice de Barthel (utilizado para avaliar a preditibilidade) aos 180 dias após AVC (0.51 a 0.59). Ainda assim, os autores afirmam que os resultados são suficientes para suportar a hipótese de que os 4 instrumentos têm validade preditiva. Os 4 instrumentos apresentaram uma responsividade moderada (efeito de Cohen ≥ 0.51) na detecção de alterações entre os 14 e 180 dias depois do AVC, sendo que o ARAT foi o instrumento mais responsivo. No entanto, todos os instrumentos se mostraram menos responsivos numa fase mais precoce da recuperação, entre os 14 e 30 dias após AVC. A fiabilidade intra-observador foi bastante elevada ($CCI \geq 0.97$) nos 4 instrumentos, o FM-ES e o ARAT apresentaram Mudança Mínima Detetável (MMD) abaixo dos 10% o que representa um valor satisfatório para erros de medição. A fiabilidade inter-observador também foi elevada ($CCI \geq 0.92$).

Todos os estudos analisados na RLS exploram algumas das propriedades psicométricas do instrumento de ARAT. No que respeita à reprodutibilidade, ou seja, quanto o instrumento esta livre do erro aleatório ou fornece um resultado reprodutível (Barbetta & Assis, 2008), os estudos de Nordin *et al.* (2014), Nijland *et al.* (2010), Chen *et al.* (2010) e Lin *et al.* (2009) analisaram a fiabilidade inter- e intra-observadores, enquanto Arlette, Rodrigo & Viviana (2012) analisaram apenas a fiabilidade intra-observadores. Os 5 estudos utilizaram metodologias diferentes pelo não é possível efetuar uma comparação, apesar da maioria ter recorrido estatisticamente ao alfa de Cronbach, Nordin *et al.* (2014) recorreram a técnicas estatísticas menos usuais nesta área de estudo, o ranking de Svensson. Contudo, todos os seus estudos demonstram uma elevada fiabilidade intra- e inter-observadores do instrumento.

A validade é uma propriedade psicométrica que avalia a capacidade do instrumento medir o que ele se propõe (Barbetta & Assis, 2008). O ARAT demonstrou uma boa validade de constructo segundo Nijland *et al.* (2010) e Chen *et al.* (2010), apesar de, no primeiro estudo, o elevado alfa de Cronbach (0.98) sugerir a existência de redundância nos itens e no segundo identificar-se redundância, em alguns itens, na hierarquização das categorias. Posteriormente, Arlette, Rodrigo & Viviana (2012) analisaram a validade de constructo subdividida em validade convergente e discriminativa. Da sua investigação pode concluir-se que o ARAT apresenta boa validade convergente, na sua correlação com instrumentos que avaliam as AVDs básicas e instrumentais, e uma boa validade discriminativa por não apresentar correlação com variáveis não relacionadas (emoção, comunicação, memória e nível socioeconómico). Três estudos compararam o ARAT com outros instrumentos (Nijland *et al.*, 2010; Lin *et al.*, 2010 e Lin *et al.*, 2009), apresentando boa correlação, em particular com o WMFT. A validade preditiva foi analisada por Chen *et al.* (2012) e Lin *et al.* (2009) e ambos apresentam resultados aceitáveis para se considerar que o instrumento tem validade preditiva, no entanto, é necessária mais investigação sobre esta propriedade, nomeadamente, através da aplicação dos instrumentos ao longo do processo de reabilitação num espaço temporal superior.

A responsividade, capacidade do instrumento de medir alterações num determinado período de tempo (Husted *et al.*, 2000), Nijland *et al.* (2010) não detetaram efeito chão e teto, mas segundo Lin *et al.* (2010) o instrumento tem boa responsividade, o que vai ao encontro dos resultados de Lin *et al.* (2009), que o tinha considerado o mais responsivo entre os 4 instrumentos analisados.

O défice motor é um dos sintomas mais prevalentes após a ocorrência de AVC sendo descrito como uma incapacidade relevante por pessoas com AVC (Feys, 1998). O ARAT foi identificado como um dos instrumentos mais utilizados internacionalmente na avaliação da funcionalidade do membro superior (Cavaco & Alouche, 2010) e verificou-se que apresenta as propriedades psicométricas, reprodutibilidade, validade e responsividade comprovadas. Através da análise destes estudos foi possível dar resposta à questão de partida da RLS.

CONSIDERAÇÕES FINAIS

Existem diversas escalas que avaliam a capacidade funcional do membro superior após a ocorrência de um AVC, sendo o ARAT amplamente utilizado a nível internacional para avaliação do membro em causa.

Os artigos incluídos nesta RSL tiveram como finalidade avaliar uma ou mais propriedades psicométricas do *Action research arm test*. Foi então avaliada a reprodutibilidade (intra e inter-observador), a validade de construto (convergente e discriminativa), concorrente e preditiva, bem como a responsividade (pela análise de efeito teto e chão) do ARAT, através da sua análise isolada, comparada e/ou apoiada noutras escalas/testes.

Os resultados deste estudo permitem concluir que o ARAT demonstrou ser um instrumento válido, fiável e responsivo na avaliação, diagnóstico e tratamento do membro superior lesado após a ocorrência de um AVC, dando assim resposta à questão de partida.

Reconhecemos então a importância da tradução, adaptação e validação do ARAT para a população portuguesa, visando o contributo na prática profissional, com consequentes ganhos na qualidade de vida das pessoas com sequelas de AVC, através da implementação de programas de reabilitação baseados em evidência.

REFERÊNCIAS BIBLIOGRÁFICAS

Barbetta, D. D. C., & Assis, M. R. (2008). Reprodutibilidade, validade e responsividade da escala de Medida de Independência Funcional (MIF) na lesão medular: revisão da literatura; *Acta fisiátrica*, 15 (3), 176-181.

Belsey, J. & Snell, T. (2001). *What is evidence-based medicine?* London: Hayward Medical Communication, (What is...?; 2).

- Cavaco, N. S., & Alouche, S. R. (2010). Instrumentos de avaliação da função de membros superiores após acidente vascular encefálico: uma revisão sistemática. *Fisioterapia e Pesquisa*, 17 (2), 178-183.
- Chen, H., Lin, K., Wu, C., & Chen, C. (2012). Rasch validation and predictive validity of the action research arm test in patients receiving stroke rehabilitation. *Archives of Physical Medicine and Rehabilitation*, 93 (6), 1039–45. doi:10.1016/j.apmr.2011.11.033
- Cirstea, M.C. & Levin, M.F. (2007). Improvement of arm movement patterns and endpoint control depends on type of feedback during practice in stroke survivors. *Neurorehabil Neural Repair*, 21, 398-411.
- Damasceno, B.P. & Borges, G. (1988). Acidentes vasculares cerebrais. *Revista Brasileira de Medicina*, 45 (6),190-196.
- Duncan, P.W., Jorgensen, H.S. & Wade, D.T. (2000). Outcome measures in acute stroke trials: a systematic review and some recommendations to improve the practice. *Stroke*, 31,1429-1438.
- Duncan, P.W., Zorowitz, R., Bates, B., Choi, J.Y., Glasberg, J.J., Graham, G.D., et al. (2005). Management of Adult Stroke Rehabilitation Care: A Clinical Practice Guideline. *Stroke*, 36, 100-143.
- Feys, H.M., De Weerd, W.J., Selz, B.E., Cox Steck, G.A., Spichiger, R., Vereeck, L.E., Putman, K.D. & Van Hoydonck, G.A. (1998). Effect of a therapeutic intervention for the hemiplegic upper limb in the acute phase after stroke: a single-blind, randomized, controlled multicenter trial, *Stroke; A Journal of Cerebral Circulation*, 29, 785-92.
- Galvão, C. M., Sawada, N. O., & Trevizan, M. A. (2004). Revisão sistemática. *Revista Latino-americana de Enfermagem*, 12(3), 549-56.
- Husted, J. A., Cook, R. J., Farewell, V. T., & Gladman, D. D. (2000). Methods for assessing responsiveness: a critical review and recommendations. *Journal of clinical epidemiology*, 53(5), 459-468.

- JBI (2011). *User Manual: Version 5.0 System for the Unified Management. Assessment and Review of Information*. Joanna Briggs Institute's. 123.
- Koh, C.L., Hsueh, I.P., Wang, W.C., Sheu, C.F., Yu, T.Y., Wang, C.H., et al. (2006). Validation of the Action Research Arm Test using item-response theory in patients after stroke. *Journal Rehabilitation Medicine*, 38 (6), 375-380.
- Leung, K., Trevena L. & Waters, D. (2012). Development of an appraisal tool to evaluate strength of an instrument or outcome measure. *Nurse Researcher*, 20 (2), 13-19.
- Li, K.-Y., Lin, K.-C., Wang, T.-N., Wu, C.-Y., Huang, Y.-H., & Ouyang, P. (2012). Ability of three motor measures to predict functional outcomes reported by stroke patients after rehabilitation. *NeuroRehabilitation*, 30(4), 267–75. doi:10.3233/NRE-2012-0755
- Lin, K., Chuang, L., Wu, C., Hsieh, Y., & Chang, W. (2010). Responsiveness and validity of three dexterous function measures in stroke rehabilitation. *The Journal of Rehabilitation Research and Development*, 47 (6), 563. doi:10.1682/JRRD.2009.09.0155
- Linde, K. & Willich, S.N. (2003). How objective are systematic reviews? Differences between reviews on complementary medicine. *Journal of the Royal Society of Medicine*, 96, 17-22.
- Lyle, R.C. (1981). A performance test for assessment of upper limb function in physical rehabilitation treatment and research. *International Journal of Rehabilitation Research*, 4, 483-492.
- Marques-Vieira, C. M. A; Sousa, L. M. M.; Carvalho, M. L; Veludo, F. & José, H. M. G. (2015). Construção, adaptação transcultural e adequação de instrumentos de medida., *Enformação*, 5, 19 - 24.
- Nijland, R., van Wegen, E., Verbunt, J., van Wijk, R., van Kordelaar, J., & Kwakkel, G. (2010). A comparison of two validated tests for upper limb function after stroke: The Wolf Motor Function Test and the Action Research Arm Test. *Journal of Rehabilitation Medicine*, 42 (7), 694–6. doi:10.2340/16501977-0560

- Nordin, Å., Alt Murphy, M., & Danielsson, A. (2014). Intra-rater and inter-rater reliability at the item level of the Action Research Arm Test for patients with stroke. *Journal of Rehabilitation Medicine*, 46 (8), 738–45. doi:10.2340/16501977-1831
- Pereira, Â. L., & Bachion, M. M. (2006). Atualidades em revisão sistemática de literatura, critérios de força e grau de recomendação de evidência. *Revista Gaúcha de Enfermagem*, 27(4), 491.
- Platz, T., Pinkowisk, C., van Wijck, G.J. & Johnson, G. (2005). Arm – Arm Rehabilitation Measurement: Manual for performance and scoring of the Fugl-Meyer test (arm section), Action Research Arm Test, and the Box and Block test. *Deutscher Wissenschafts- Verlag*, 6-105.
- Rabadi, M. H., & Rabadi, F. M. (2006). Comparison of the action research arm test and the Fugl-Meyer assessment as measures of upper-extremity motor weakness after stroke. *Archives of physical medicine and rehabilitation*, 87(7), 962-966.
- Rosamond, W., Flegal, K., Furie, K., Go, A., Greenlund, K., Haase, N., et al. (2008). Heart disease and stroke statistics – 2008 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation*, 117, e25-e46.
- Sandelowski, M.; Barroso, J. (2007). *Handbook for Synthesizing Qualitative Research*. New York: Springer Publishing Company.
- Santos, C., Pimenta, C. & Nobre, M. (2007). A estratégia PICO para a construção da pergunta de pesquisa e busca de evidências. *Revista Latino-americana de Enfermagem*, 15(3), 508-511.
- Svensson, E. (2012). Different ranking approaches defining association and agreement measures of paired ordinal data. *Statistics in medicine*, 31(26), 3104-3117.
- Sousa, L. M. M.; Marques-Vieira, C. M. A; Carvalho, M. L; Veludo, F. & José, H. M. G. (2015). Fidelidade e validade na construção e adequação de instrumentos de medida. *Enfermagem*, 5: 25 – 32

Wade, D.T. (1992). *Measurement in neurological rehabilitation*. Boston: Oxford University Press. Chap. Personal physical disability, 79-82.