



Propriedades psicométricas de questionários de atividade física na adolescência: revisão sistemática

Psychometric properties of physical activity questionnaires in adolescence: systematic review

AUTORES

Maria Cristine Campos¹
Cristiele de Aguiar Felicidade²
Susana da Costa Aguiar¹
Danielle Soares Rocha Vieira²

1 Universidade Federal de Santa Catarina, Programa de Pós-Graduação em Ciências da Reabilitação, Araranguá, Santa Catarina, Brasil.

2 Universidade Federal de Santa Catarina, Departamento de Ciências da Saúde, Araranguá, Santa Catarina, Brasil.

CONTATO

Danielle Soares Rocha Vieira
danielle.vieira@ufsc.br
Universidade Federal de Santa Catarina,
Campus Araranguá. Rodovia Governador
Jorge Lacerda, nº 3201 – Km 35,4. Bairro
Jardim das Avenidas, Araranguá, Santa
Catarina, Brasil.
CEP: 88906-072.

DOI

10.12820/rbafs.23e0045



Este obra está licenciado com uma Licença
Creative Commons Atribuição-NãoComercial-
CompartilhaIgual 4.0 Internacional.

RESUMO

O objetivo desta revisão foi recuperar e sumarizar os estudos que avaliaram as propriedades psicométricas (validade, confiabilidade e responsividade) de questionários utilizados para mensurar a atividade física (AF) em adolescentes. Trata-se de uma revisão sistemática realizada de acordo com o protocolo PRISMA. A busca foi realizada nas bases de dados: MEDLINE (via Ovid), CINAHL, LILACS, Cochrane, SciELO, PsycINFO, Web of Science e SPORTDiscus, utilizando descritores específicos. O risco de viés foi avaliado por meio do Checklist COSMIN. Foram incluídos 26 estudos na revisão e, destes, apenas cinco foram realizados no Brasil. A qualidade dos estudos variou de fraca a boa. Dos 25 questionários que foram avaliados quanto à confiabilidade teste-reteste, 11 apresentaram coeficientes de correlação intraclassa $\geq 0,70$, incluindo o Questionário de Atividade Física para Adolescentes (QFAA). No que se refere à validade, apenas um dos 32 questionários avaliados quanto a esta propriedade apresentou coeficiente de correlação $\geq 0,70$ e cinco apresentaram coeficientes de magnitude moderada (0,56 a 0,69), sendo dois de nacionalidade brasileira. Porém, ambos foram comparados a medidas subjetivas de AF. A responsividade dos questionários não foi avaliada em nenhum dos estudos. Os questionários de AF incluídos nesta revisão apresentaram melhores índices de confiabilidade do que validade e não foram investigados quanto à sua responsividade. Esses resultados reforçam a necessidade de realização de novos estudos com qualidade metodológica mais adequada, que investiguem principalmente a validade e a responsividade desses instrumentos.

Palavras-chave: Atividade motora; Adolescente; Reprodutibilidade dos testes; Validade dos testes.

ABSTRACT

The aim of this review was to retrieve and identify studies that evaluated the psychometric properties (validity, reliability and responsiveness) of questionnaires used to measure physical activity (PA) in adolescents. It is a systematic review carried out according to the PRISMA protocol. Searches were performed in: MEDLINE (via Ovid), CINAHL, LILACS, Cochrane, SciELO, PsycINFO, Web of Science and SPORTDiscus, using specific terms. The risk of bias was assessed through the COSMIN Checklist. Twenty-six studies were included in the review, of which only five were performed in Brazil. The quality of the studies ranged from poor to good. Of the 25 instruments that were evaluated for test-retest reliability, 11 presented intraclass correlation coefficients ≥ 0.70 , including the Physical Activity Questionnaire for Adolescents (QFAA). Regarding validity, only one instrument of the 32 evaluated for this property presented coefficient correlation ≥ 0.70 and five presented coefficients of moderate magnitude (0.56 a 0.69), two of them from Brazil. However, both were compared to subjective criterion measures. The responsiveness of the questionnaires was not assessed in any of the studies. The PA questionnaires included in this review had better reliability indexes than validity and were not investigated for their responsiveness. These results reinforce the need to carry out new studies with more adequate methodological quality that investigate mainly the validity and responsiveness of these instruments.

Keywords: Motor activity; Adolescent; Reproducibility of results; Validity of results.

Introdução

A inatividade física é considerada um dos principais problemas de saúde pública do século XXI¹, além de ser importante fator de risco para o desenvolvimento das doenças crônicas não transmissíveis^{1,2}. Embora grande parte dessas doenças se manifeste na vida adulta, seu desenvolvimento inicia-se na infância e na adolescên-

cia³, fase em que há maior declínio dos níveis de atividade física (AF) quando comparado a qualquer outro ciclo da vida⁴.

A mensuração da AF em adolescentes tem sido realizada com o intuito de monitorar os seus níveis, identificar seus fatores correlatos e determinantes, bem como para avaliar a eficácia de intervenções⁵⁻⁹. Nesse

contexto, medidas de AF por meio de questionários têm sido amplamente realizadas devido a sua praticidade e ao baixo custo¹⁰. Entretanto, devido a diversidade de instrumentos disponíveis na literatura, sua seleção se torna uma tarefa complexa.

Diversos fatores devem ser considerados na escolha de um questionário, como a população em estudo, sua aplicabilidade e suas propriedades psicométricas, com destaque para os níveis confiabilidade, validade e responsividade¹¹⁻¹³. Chinapaw et al.¹¹ investigaram, por meio de uma revisão sistemática, as propriedades psicométricas dos questionários de medida do nível de AF para crianças e adolescentes. Os autores identificaram uma escassez de estudos que reportem as propriedades psicométricas de questionários nesta população, assim como a ausência de uma medida “padrão-ouro” para validação desses instrumentos. Farias Júnior et al.¹⁰ também avaliaram, por meio de uma revisão sistemática, a validade e a confiabilidade dos instrumentos de medida do tipo autorrelato para adolescentes, incluindo questionários. Eles concluíram que os instrumentos apresentaram melhores índices de confiabilidade do que de validade.

Embora essas revisões apresentem um detalhamento dos questionários de avaliação do nível de AF em adolescentes disponíveis na literatura, além do relato de suas propriedades psicométricas, elas abrangem estudos publicados até 2009, o que levanta a importância de uma revisão que possa envolver estudos mais recentes¹⁴⁻²⁴. Nesse contexto, o objetivo deste trabalho foi recuperar e sumarizar os estudos que avaliaram as propriedades psicométricas (confiabilidade, validade e responsividade) dos questionários utilizados para mensurar a AF em adolescentes.

Métodos

Trata-se de uma revisão sistemática, realizada de acordo com os itens do PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses)²⁵.

Nesta revisão, foram incluídos artigos originais que avaliaram a confiabilidade, a validade e/ou a responsividade dos questionários utilizados para avaliar o nível de AF em adolescentes com idade de dez a 19 anos ou com média de idade dentro desse intervalo, publicados nos últimos dez anos em inglês, português e/ou espanhol.

Não foram incluídos estudos que investigaram as propriedades psicométricas de diários ou logs, estudos que avaliaram populações especiais, como pessoas com doenças respiratórias, neurológicas, cardiovasculares,

musculoesqueléticas, dentre outras, assim como estudos de revisão, estudos de caso, monografias, estudos publicados somente como anais de eventos científicos, capítulos de livros e ponto de vista ou opinião de especialista.

Para a avaliação dos itens relacionados à validade de critério, além dos instrumentos “padrão-ouro” como a água duplamente marcada, observação direta e calorimetria indireta, foram considerados também os instrumentos objetivos de medida do nível de AF como critério de referência, a saber, os sensores de movimento (acelerômetros e pedômetros) e monitores de frequência cardíaca^{8,13,26}.

A busca foi realizada no período de janeiro a fevereiro de 2016 nas bases de dados: MEDLINE (via Ovid), CINAHL, LILACS, Cochrane, SciELO, PsycINFO, Web of Science e SPORTDiscus. Os descritores utilizados nas diferentes bases de dados incluíram o grupo populacional, o desfecho “atividade física” e as propriedades psicométricas (confiabilidade, validade e responsividade). Além disso, os operadores booleanos utilizados foram “OR” quando sinônimos e “AND” entre grupo populacional, AF e propriedades psicométricas. Uma síntese dos descritores, assim como as estratégias de busca utilizadas em cada base de dados encontram-se em apêndice.

A seleção dos estudos foi realizada por dois examinadores de forma independente por meio da análise do título e do resumo, obedecendo rigorosamente os critérios de inclusão e exclusão. Todos os artigos potencialmente relevantes foram recuperados na forma de texto para a avaliação independente dos dois revisores. Uma revisão das referências dos artigos incluídos foi realizada a fim de identificar estudos potencialmente relevantes que não tinham sido encontrados nas buscas iniciais.

O checklist proposto pela iniciativa COSMIN (Consensus-based Standards for the Selection of Health Measurement Instruments)²⁷ foi utilizado para avaliar a qualidade dos estudos. Este instrumento permite a avaliação da qualidade de estudos que investigam as propriedades psicométricas a seguir: consistência interna; confiabilidade teste-reteste, intra e interexaminador; validade de conteúdo, de face, de construto e de critério; e responsividade²⁷. Ele possui 12 categorias de avaliação em que dez são usadas para determinar se o estudo satisfaz as normas de boa qualidade e outros dois destinam-se à Teoria de Resposta ao Item (TRI) e generalização dos resultados. Para cada propriedade há uma quantidade de itens a serem avaliados, por exemplo, para a consistência interna há 11 itens; para a confiabilidade há

14; e para a validade de critério há sete itens. Cada item pode ser avaliado como “excelente”, “bom”, “razoável” ou “fraco”. A classificação final do estudo para cada propriedade psicométrica é obtida considerando-se a pior classificação alcançada para um dos itens específicos de cada propriedade. Desta forma, mesmo que algum item de avaliação de determinada propriedade seja “excelente”, se houver outro que seja “fraco” a classificação final para aquela propriedade será “fraco”²⁷.

A avaliação do risco de viés dos estudos incluídos foi realizada de forma independente por dois examinadores e a concordância entre eles foi avaliada por meio do Coeficiente de Correlação Intraclasse (CCI) (two-way random / absolute agreement). Um CCI foi gerado para a avaliação das propriedades psicométricas confiabilidade e consistência interna e outro CCI para a validade. O CCI para a confiabilidade e a consistência interna foi de 0,79 (IC95%: 0,74–0,83) e para validade foi de 0,83 (IC95%: 0,75–0,87). O percentual de itens classificados como “excelente”, “bom”, “razoável” ou “fraco” foi calculado para cada propriedade, por estudo. Para isso, foi realizada a média da classificação de cada um dos examinadores.

A extração do conteúdo dos artigos incluídos na revisão foi realizada por dois examinadores independentes de forma padronizada e foram extraídas informações relativas às características do estudo, à amostra e aos desfechos de confiabilidade, validade e responsividade.

Em todas as etapas da revisão, as divergências, quando existentes, foram resolvidas por um terceiro avaliador.

Uma síntese narrativa dos dados foi utilizada assim como uma discussão sobre o risco de vieses dos estudos foi realizada. Para os resultados relativos à confiabilidade, foi considerado aceitável CCI $\geq 0,70$ e, para a validade, a magnitude dos coeficientes de correlação foi tratada como baixa ($\leq 0,50$), moderada (0,51–0,69) e elevada ($\geq 0,70$)¹⁰.

Resultados

Ao final das buscas nas bases de dados eletrônicas, foram encontrados 11.495 artigos (Figura 1). Após a remoção das duplicatas e dos artigos com título e resumo que não atendiam aos critérios de inclusão, restaram 64 estudos para leitura na íntegra. Destes, 19 foram incluídos, assim como outros sete artigos por meio de buscas manuais, o que resultou em uma síntese baseada nos dados de 26 artigos.

A Tabela 1 apresenta a descrição dos 26 estudos. A

confiabilidade teste-reteste foi investigada em 20^{9,14-22,28-37} deles, a validade de critério em 25^{9,14-19,21-24,28-41} e a consistência interna em 5^{16,21,37,38,40}. Em nenhum dos estudos, foi investigada a responsividade dos questionários. A maioria dos estudos foi realizada na Europa^{9,16,20,23,24,31,35-37,39,41} e apenas cinco^{15,21,28,30,33} deles no Brasil.

Nos artigos incluídos, foram analisados 36 questionários estruturados, sendo seis de nacionalidade brasileira^{15,21,28,30,33}. Desses, foram identificados 27 instrumentos diferentes, uma vez que as propriedades psicométricas do International Physical Activity Questionnaire (IPAQ) versão curta foram avaliadas em quatro estudos^{9,18,28,34}, do International Physical Activity Questionnaire for Adolescents (IPAQ-A) em dois estudos^{23,39}, do Physical Activity Questionnaire for Adolescents (PAQ-A) em quatro estudos^{21,34,37,40} e do Physical Activity Questionnaire for Older Children (PAQ-C) em três estudos^{21,38,40}.

A Tabela 2 apresenta a classificação da qualidade dos estudos para cada uma das propriedades psicométricas de acordo com o Checklist COSMIN. No que diz respeito à confiabilidade teste-reteste, 20 estudos^{9,14-22,28-37} avaliaram esta propriedade e apenas o estudo de Corder et al.³⁶ foi classificado como qualidade “fraco” e os demais estudos foram classificados como qualidade “razoável”. Sobre a consistência interna, cinco estudos^{16,21,37,38,40} avaliaram esta propriedade e os estudos de Janz et al.⁴⁰ e Moore et al.³⁸ foram classificados como “razoável” enquanto os demais estudos^{16,21,37} como “fraco”. Para a validade de critério, 25 estudos^{9,14-19,21-24,28-41} avaliaram esta propriedade. Os estudos de Barbosa et al.³², Corder et al.³⁶, Florindo et al.³⁰, Farias Júnior et al.¹⁵ e Guedes et al.²⁸ obtiveram classificação de qualidade “fraco”, enquanto Lachat et al.³⁴ foi o único com classificação “bom”. Os demais^{9,14,16-19,21-24,29,31,33,35,37-41} apresentaram qualidade “razoável” para esta categoria”.

No material apêndice, encontra-se o Tabela 1, onde são apresentados os percentuais de itens classificados pelos examinadores como “fraco”, “razoável”, “bom” e “excelente” para cada uma das propriedades psicométricas de cada um dos artigos incluídos na revisão. Observa-se que grande parte dos itens foram classificados como “bons” ou “excelentes” para todas as propriedades psicométricas avaliadas nos estudos.

Confiabilidade

Na maioria dos estudos^{9,14-22,28-37}, foi investigada a confiabilidade teste-reteste dos instrumentos e o intervalo de tempo entre as medidas variou de três dias¹⁹ a seis

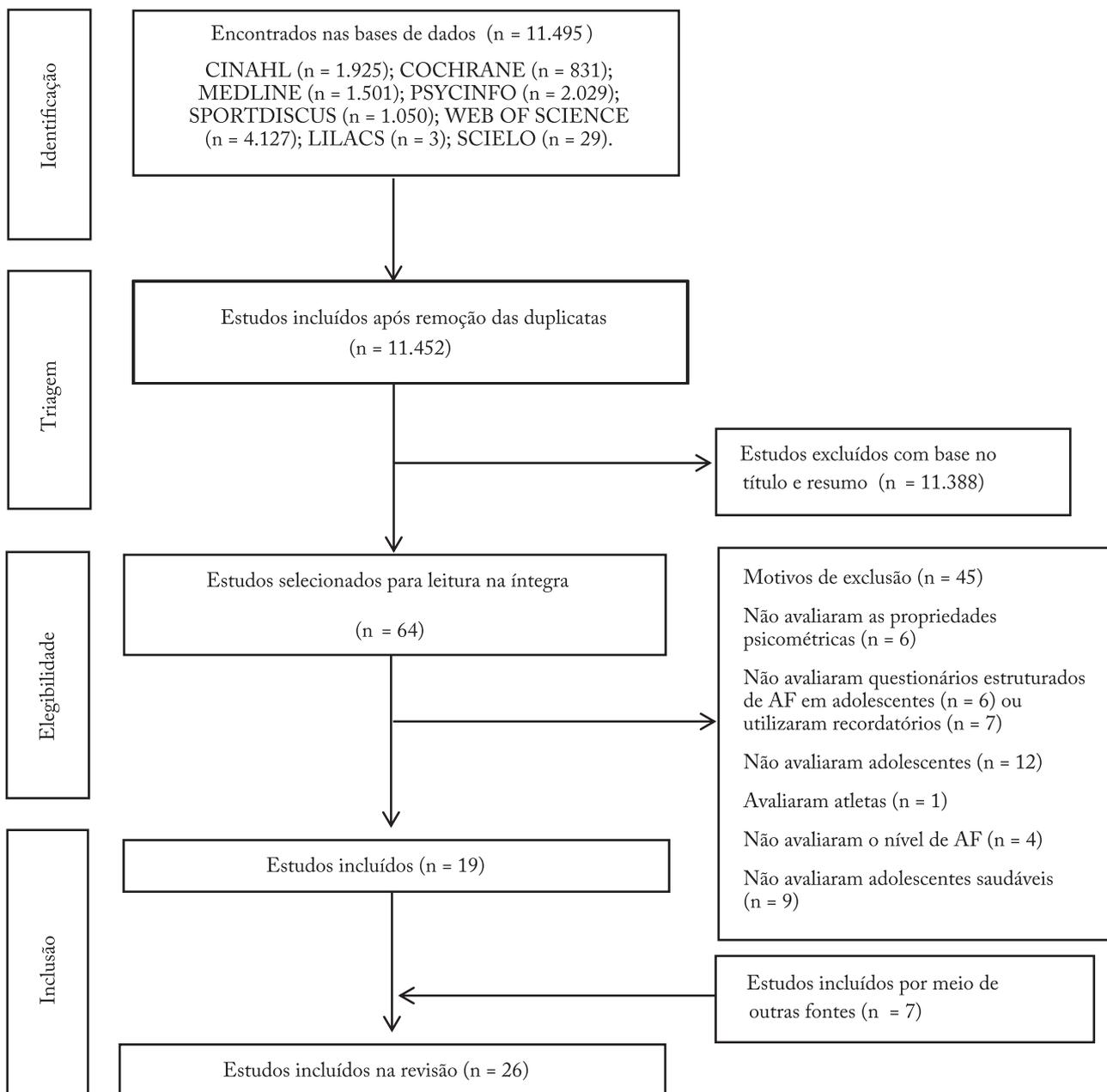


Figura 1 – Fluxograma do processo de seleção dos estudos.

semanas³². O número de participantes avaliados foi de 33³¹ a 857³³ e na maioria dos estudos foi superior a 100^{14,15,17,19-22,28,32-34}.

O CCI foi utilizado na maioria dos estudos^{9,14-22,30,31,35-37} como forma de análise de confiabilidade, com variação de 0,02³⁶ a 1³¹. Dos 25 instrumentos que foram avaliados quanto à confiabilidade teste-reteste, 11 apresentaram CCI $\geq 0,70$ ^{9,15-19,22,29,32,35-37}, considerando-se todos os seus domínios e ambos os sexos e, destes, quatro^{15,16,32,36} alcançaram CCI $> 0,80$. Nos estudos brasileiros, somente o questionário do estudo de Farias

Júnior et al.¹⁵, o Questionário de Atividade Física para Adolescentes (QAFA), alcançou valores de CCI $\geq 0,70$.

A consistência interna foi investigada em apenas cinco estudos^{16,21,37,38,40}. O tamanho amostral variou de 72¹⁶ a 991³⁸ e forma de análise usada foi o alfa de Cronbach, com valores superiores a 0,70 em todos os estudos.

Validade

A validade de critério foi investigada em 25 estudos^{9,14-19,21-24,28-41}. O número de participantes variou de 33³¹ a 2.018²³ e a maioria dos estudos^{17-19,21,23,24,28,34,37-39,41}

Tabela 1 – Características descritivas e propriedades psicométricas analisadas pelos estudos incluídos na revisão.

Autor	Local	Questionários	Faixa etária da amostra	Propriedades psicométricas analisadas		
				Confiabilidade teste-reteste	Consistência interna	Validade de critério
Barbosa et al. ³²	Colômbia	Quantification de L'Activite Physique en Altitude chez les Enfants (QAPACE)	8-16 anos	n = 121 6 semanas CCI = 0,96 (IC95%: 0,95-0,97)	NA	n = 36 QAPACE x VO ₂ pico (Ergoespirometria): r = 0,69 (p < 0,00001) QAPACE x Leger Test: r = 0,56 (p < 0,0001)
Bastos et al. ³³	Brasil	Não apresenta o título do questionário (Será chamado Questionário "versão curta")	10-19 anos	n = 857 14 dias r = 0,62 (p < 0,001) e Kappa = 0,58	NA	n = 92 Questionário "versão curta" x versão longa: r = 0,66 (p < 0,01) e Kappa = 0,63 Questionário "versão curta" x pedômetro: r = 0,26 (p = 0,02) e Kappa = 0,18
Bem Gharbia et al. ¹⁹	Tunísia	Physical activity frequency questionnaire (PAFQ)	10-19 anos	n = 142 3 dias CCI = 0,79-0,93 e Kappa = 0,47-0,78	NA	n = 142 PAFQ x Monitor de FC: r = 0,70 (IC 95%: 0,62-0,76) PAFQ x R-24h: r = 0,81 (IC 95%: 0,77-0,84)
Bobakova et al. ²⁰	República Checa, Eslováquia e Polônia	Health Behaviour in School-aged Children (HBSC)	11-15 anos	n = 693 1 a 4 semanas CCI = 0,41-0,60 (p < 0,001) e Kappa = 0,30-0,5	NA	NA
Corder et al. ³⁶	Reino Unido	Youth Physical Activity Questionnaire (YPAQ) Child Heart and Health Study in England Questionnaire (CHASE) Swedish Adolescent Physical Activity Questionnaire (SWAPAQ)	12-17 anos	n = 41 7 dias YPAQ: CCI = 0,73-0,92 (p < 0,001) CHASE: CCI = 0,64 (p < 0,001) SWAPAQ: CCI = 0,02-0,05	NA	n = 49 YPAQ x acelerômetro: r = 0,11-0,42 (p = 0,04) YPAQ x água duplamente marcada: r = 0,09-0,46 (p = 0,03) SWAPAQ x acelerômetro: r = 0,23 (p = 0,27) SWAPAQ x água duplamente marcada: r = 0,40 (p = 0,04) CHASE x água duplamente marcada: r = 0,45 (p = 0,02) CHASE x acelerômetro: r = 0,12 (p = 0,57)
Farias Júnior et al. ¹⁵	Brasil	Questionário de Atividade Física para Adolescentes (QAFA)	14-19 anos	n = 239 1 semana CCI = 0,88 (IC 95% 0,84-0,91) e Kappa = 0,42-0,58	NA	n = 70 QAFA x R-24h: r = 0,62 (p < 0,001)
Florindo et al. ³⁰	Brasil	Questionário de Atividade Física Habitual	11-16 anos	n = 94 15 dias CCI = 0,61-0,68 #	NA	n = 94 Questionário de Atividade Física Habitual x VO ₂ : r = 0,18 para AF semanal # r = 0,28 (p < 0,01) para AF anual
Gioxari et al. ¹⁶	Grécia	Self-Administered Physical Activity Checklist (SAPAC)	10-13 anos	n = 72 2 semanas CCI = 0,85-0,88 (p < 0,001)	n = 72 α: 0,87-0,89 (p < 0,001)	n = 90 SAPAC x acelerômetro: Kendall's tau-b = 0,31-0,37 (p < 0,01)
Guedes & Guedes ²¹	Brasil	Physical Activity Questionnaire for Older Children – (PAQ-C) Physical Activity Questionnaire for Adolescents – (PAQ-A)	8-18 anos	n = 528 1 semana PAQ-C: CCI = 0,68-0,81 # e PAQ-A: CCI = 0,65-0,85 #	n = 528 PAQ-C: α = 0,71 e PAQ-A: α = 0,76	n = 528 PAQ-C x acelerômetro: r = 0,40 (p < 0,05) e PAQ-A x acelerômetro: r = 0,50 (p < 0,05)
Guedes et al. ²⁸	Brasil	International Physical Activity Questionnaire (IPAQ) versão curta	12-18 anos	n = 161 2 semanas r = 0,49-0,83 (p < 0,05)	NA	n = 161 IPAQ x R-24h: r = 0,09-0,51 (p < 0,05)

Continua...

... continua

Autor	Local	Questionários	Faixa etária da amostra	Propriedades psicométricas analisadas		
				Confiabilidade teste-reteste	Consistência interna	Validade de critério
Hagströmer et al. ³⁹	Europa	International Physical Activity Questionnaire (IPAQ) versão longa modificada	12-17 anos	NA	NA	n = 284 IPAQ _x acelerômetro = 0,17-0,30 (p < 0,05) Tau-b de Kendall = 0,18-0,22 (p < 0,001)
Janz et al. ⁴⁰	EUA	Physical Activity Questionnaire for Adolescents (PAQ-A)	11-13 anos	NA	n = 210 α: 0,72-0,88	n = 49 PAQ-A x acelerômetro = 0,56-0,63 (p < 0,05)
Lachat et al. ³⁴	Vietnã	International Physical Activity Questionnaire (IPAQ) versão curta	Média de 16 anos	2 semanas IPAQ: n = 200 CCI = 0,15-0,40 e r = 0,33-0,52 #	NA	n = 188 IPAQ _x acelerômetro: r = -0,01-0,29 # PAQ-A x acelerômetro: r = 0,18-0,27 #
		Physical Activity Questionnaire for Adolescents (PAQ-A)		PAQ-A: n = 150 CCI = 0,12-0,40 e r = 0,22-0,57 #		
Lubans et al. ³⁵	Reino Unido	Oxford Physical Activity Questionnaire (OPAQ)	11-15 anos	n = 87 1 semana CCI = 0,76-0,91 (IC95%: 0,87-0,95)	NA	n = 51 OPAQ _x acelerômetro: r = 0,01-0,32 (p = 0,02)
Martínez-Gómez et al. (a) ³⁷	Espanha	Physical Activity Questionnaire for Adolescents (PAQ-A)	12-17 anos	n = 78 1 semana CCI = 0,71 #	n = 203 α = 0,65-0,74 #	n = 203 PAQ-A x acelerômetro: r = 0,34-0,39 (p < 0,001)
Martínez-Gómez et al. (b). ⁴¹	Espanha	Pergunta de atividade física enKid	12-17 anos	NA	NA	n = 200 enKid x acelerômetro: r = 0,43(p < 0,001) FITNESSGRAM x acelerômetro: r = 0,36 (p < 0,001) PACE x acelerômetro: r = 0,43 (p < 0,001) Escala comparativa de atividade física x acelerômetro: r = 0,39 (p < 0,001)
		Pergunta de atividade física FITNESSGRAM				
		Questionário de atividade física PACE				
		Escala de atividade física comparativa				
Moore et al. ³⁸	EUA	Physical Activity Questionnaire for Older Children (PAQ-C)	8-14 anos	NA	n = 991 α = 0,70-0,74	n = 991 PAQ-C x Harvard step test: r = 0,08 (p < 0,08)
Murphy et al. ²⁴	República da Irlanda	PACE +	10-18 anos	NA	NA	n = 235 PACE + x acelerômetro: r = 0,27- 0,34 (p < 0,01)
Okuda et al. ¹⁴	Japão	Physical activity questionnaire (PAQ) e o Lifestyle questionnaire (LQ)	10-14 anos	n = 120 1 semana CCI PAQ = 0,13-0,64 e r = 0,33-0,63 (p < 0,001)	NA	n = 96 PAQ _x acelerômetro: r = 0,17-0,73 (p < 0,001) n = 94 LQ _x acelerômetro: r = -0,02-0,75 (p < 0,001)
Ottevaere et al. ²³	Europa	International Physical Activity Questionnaire for Adolescents (IPAQ-A)*	12-17 anos	NA	NA	n = 2018 IPAQ-A x acelerômetro: r = 0,15-0,25 (p < 0,01) n = 1696 IPAQ-A x VO ₂ máx: r = 0,08-0,35 (p < 0,01)

Continua...

... continua

Autor	Local	Questionários	Faixa etária da amostra	Propriedades psicométricas analisadas		
				Confiabilidade teste-reteste	Consistência interna	Validade de critério
Philippaerts et al. ³¹	Bélgica	Flemish Physical Activity Computer Questionnaire (FPACQ)	12-18 anos	n = 33 10 dias CCI = 0,68-1,00 e Kappa = 0,20-1,00	NA	n = 33 FPACQ x acelerômetro r = -0,00-0,78
Rangul et al. ⁹	Noruega	WHO HBSC Physical Activity Questionnaire International Physical Activity Questionnaire (IPAQ) versão curta	11-13 anos	n = 71 8 a 12 dias WHO HBSC: CCI = 0,71-0,73 IPAQ: CCI = 0,10-0,62	NA	n = 71 WHO HBSC x VO ₂ máx: r = 0,33-0,39 (p < 0,01) IPAQ x VO ₂ máx: r = 0,32 (p < 0,01) WHO HBSC x acelerômetro: r = 0,02-0,23 # IPAQ x acelerômetro: r = -0,03-0,09 #
Scott et al. ²²	Austrália	Single-item physical activity measure (SIPAM)	13-15 anos	n = 107 2 semanas CCI = 0,75 (p < 0,001)	NA	n = 96 SIPAM x OPAQ: r = 0,50 (p < 0,001) n = 72 SIPAM x acelerômetro: r = 0,44 (p < 0,001)
Teo et al. ¹⁷	China	New computer-based physical activity questionnaire (cPAQ)	11-13 anos	n = 203 2 semanas CCI = 0,72 (p < 0,001)	NA	n = 222 cPAQ x monitor de FC: r = 0,69 (p < 0,001) cPAQ x diário: r = 0,63 (p < 0,001)
Treuth et al. ²⁹	EUA	Fels physical activity questionnaire (PAQ)	7-19 anos	6 dias Elementary school (n = 70) CCI = 0,70 # Middle school (n = 81) CCI = 0,62 # High school (n = 78) CCI = 0,71 #	NA	PAQ x acelerômetro Elementary school (n = 70) r = 0,34 (p = 0,004) Middle school (n = 81) r = 0,11 (p = 0,31) High school 1 (n = 78) r = 0,21 (p = 0,006)
Wang et al. ¹⁸	China	International Physical Activity Questionnaire (IPAQ) versão curta	12-18 anos	n = 92 1 semana CCI = 0,74 (p < 0,001)	NA	n = 1.021 IPAQ x acelerômetro: r = 0,18-0,31 #

nível de significância da relação não apresentado; CCI = Coeficiente de Correlação Intraclasse; r = coeficiente de correlação de Spearman ou de Pearson; n = número da amostra; R-24h = recordatório de 24 horas; α = alfa de cronbach; AF = atividade física; IC95% = intervalo de confiança de 95%; VO₂max = consumo máximo de oxigênio; VO₂pico = consumo de oxigênio pico; p = valor p; NA = Não avaliado; FC = frequência cardíaca.

Nota: Questionários citados na Tabela – CHASE = Child Heart and Health Study in England Questionnaire; cPAQ = New Computer-Based Physical Activity Questionnaire; Fels PAQ = Fels Physical Activity Questionnaire; FPACQ = Flemish Physical Activity Computer Questionnaire; HBSC = Health Behaviour in School-aged Children; IPAQ = International Physical Activity Questionnaire; IPAQ-A = International Physical Activity Questionnaire for Adolescents; IPAQ-SF = International Physical Activity Questionnaire-Short Form; LQ = Lifestyle Questionnaire; OPAQ = Oxford Physical Activity Questionnaire; PAFQ = Physical Activity Frequency Questionnaire; PAQ = Physical Activity Questionnaire; PAQA = Physical Activity Questionnaire For Adolescents; PAQ-A = Questionário de Atividade Física para Adolescentes Espanhóis; PAQ-C = Physical Activity Questionnaire for Older Children; QAFA = Questionário de Atividade Física para Adolescentes; QAPACE = Quantification De L'activite Physique En Altitude Chez Les Enfants; SAPAC = Self-Administered Physical Activity Checklist; SWAPAQ = Swedish Adolescent Physical Activity Questionnaire; WHO HBSC = World Health Organization Health Behaviour in School-aged Children; YPAQ = Youth Physical Activity Questionnaire; Single-item physical activity measure (SIPAM); *Trata-se do mesmo questionário utilizado no estudo de Hagströmer et al.³⁹

envolveu um número superior a 100 adolescentes.

A medida objetiva mais utilizada como critério para a validação dos questionários foi o acelerômetro^{9,14,16,18,21-24,29,31,34-37,39-41}. A água duplamente marca-

da³⁶, assim como o pedômetro³³ foram utilizados em apenas um estudo cada. O teste de esforço cardiopulmonar associado à análise de gases^{9,32} e os monitores de frequência cardíaca^{17,19,38} também foram utilizados

Tabela 2 – Classificação da qualidade dos estudos para cada uma das propriedades psicométricas de acordo com o Checklist COSMIN.

Autor	Confiabilidade teste-reteste	Confiabilidade consistência interna	Validade de critério
Barbosa et al. ³²	Razoável	NA	Fraco
Bastos et al. ³³	Razoável	NA	Razoável
Bem Gharbia et al. ¹⁹	Razoável	NA	Razoável
Bobakova et al. ²⁰	Razoável	NA	NA
Corder et al. ³⁶	Fraco	NA	Fraco
Florindo et al. ³⁰	Razoável	NA	Fraco
Farias Júnior et al. ¹⁵	Razoável	NA	Fraco
Gioxari et al. ¹⁶	Razoável	Fraco	Razoável
Guedes et al. ²⁸	Razoável	NA	Fraco
Guedes & Guedes ²¹	Razoável	Fraco	Razoável
Hagströmer et al. ³⁹	NA	NA	Razoável
Janz et al. ⁴⁰	NA	Razoável	Razoável
Lachat et al. ³⁴	Razoável	NA	Bom
Lubans et al. ³⁵	Razoável	NA	Razoável
Martínez-Gómez et al. (a) ³⁷	Razoável	Fraco	Razoável
Martínez-Gómez et al. (b). ⁴¹	NA	NA	Razoável
Moore et al. ³⁸	NA	Razoável	Razoável
Murphy et al. ²⁴	NA	NA	Razoável
Okuda et al. ¹⁴	Razoável	NA	Razoável
Ottevaere et al. ²³	NA	NA	Razoável
Philippaerts et al. ³¹	Razoável	NA	Razoável
Rangul et al. ⁹	Razoável	NA	Razoável
Scott et al. ²²	Razoável	NA	Razoável
Teo et al. ¹⁷	Razoável	NA	Razoável
Treuth et al. ²⁹	Razoável	NA	Razoável
Wang et al. ¹⁸	Razoável	NA	Razoável

NA = propriedade psicométrica não avaliada no estudo. No Checklist COSMIN, para cada propriedade psicométrica há uma quantidade de itens a serem avaliados: 14 para a confiabilidade teste-reteste, 11 para a consistência interna e 7 para a validade de critério. Cada item pode ser avaliado como “excelente”, “bom”, “razoável” ou “fraco”. A classificação final do estudo para cada propriedade psicométrica, a qual é apresentada nesta tabela, é obtida considerando-se a pior classificação alcançada para um dos itens específicos de cada propriedade.

como medidas de comparação. O Shuttle Run Test foi empregado em dois estudos^{23,30}. Por outro lado, também se observou o uso de medidas subjetivas, como os recordatórios de 24 horas^{15,19,28} e diários de autorrelato¹⁷.

Nos estudos brasileiros, o pedômetro³³, o Shuttle Run Test³⁰, o recordatório de 24 horas^{15,28} e o acelerômetro²¹ foram os critérios adotados para a validação dos questionários.

Em relação ao método de análise dos dados para validação, observou-se o predomínio dos coeficientes de correlação de Spearman^{9,14,15,17-19,21,23,24,28-30,33-37,39-41} ou Pearson^{19,22,31,32,38}. Esses coeficientes variaram de 0,00³¹ a 0,82¹⁹. Dentre os 32 instrumentos que foram avaliados quanto à sua validade, apenas o Physical activity frequency questionnaire (PAFQ)¹⁹ apresentou coeficiente

de correlação $\geq 0,70$. Este instrumento foi comparado a uma medida subjetiva (recordatório de 24 horas) e a uma medida objetiva (monitor de frequência cardíaca).

Cinco instrumentos^{15,17,32,33,40} apresentaram coeficientes de correlação de magnitude moderada que variaram entre 0,56^{32,40} a 0,69¹⁷, sendo dois de nacionalidade brasileira^{15,33}. Dentre os estudos brasileiros, o questionário investigado por Bastos et al.³³ apresentou coeficiente de magnitude moderada quando comparado a medida de critério subjetiva enquanto que quando comparado ao pedômetro, o coeficiente obtido foi de baixa magnitude. Ademais, no estudo de Farias Júnior et al.¹⁵ foi usado como critério de referência apenas uma medida de caráter subjetivo.

A maioria dos instrumentos^{9,14,16,18,21-24,28-31,34-39,41}

apresentou medidas de validade $\leq 0,50$. Verificou-se que grande parte deles foram comparados com ao menos uma medida objetiva^{9,14,16,18,21-24,29-31,34-39,41}, que em sua maioria, foi o acelerômetro^{9,14,16,18,21,22-24,29,31,34-37,39-41}.

Discussão

Os resultados do presente estudo mostraram a diversidade de questionários disponíveis na literatura para a mensuração do nível de AF em adolescentes de diferentes nacionalidades. A confiabilidade do tipo teste-reteste e a validade de critério foram as propriedades mais investigadas e nenhum estudo avaliou a responsividade dos instrumentos.

Em relação à qualidade dos estudos, a maioria dos estudos foram classificados como razoável ou fraco. Com base no Checklist COSMIN, observou-se que as principais limitações apresentadas pelos estudos no que diz respeito à confiabilidade foram: a não descrição do tipo de modelo de CCI empregado para a análise dos dados e a ausência de informações relativas aos procedimentos de coleta (avaliadores envolvidos e estabilidade dos sujeitos sob análise). Em relação à validade, destacam-se as limitações das medidas usadas como critérios de referência. Ademais, tanto para confiabilidade como para a validade os autores não descreveram o percentual de itens missing após a aplicação dos questionários e como esses itens foram abordados. O único estudo que descreveu esse percentual foi o de Scott et al.²² que mesmo assim recebeu uma classificação “razoável”, pois descreveu a quantidade de itens perdidos, mas não como a essa perda foi abordada.

No que diz respeito à confiabilidade teste-reteste, 11 instrumentos^{9,15-19,22,29,32,35-37} apresentaram índices iguais ou superiores a 0,70, sendo apenas um de nacionalidade brasileira (QAFA)¹⁵. Em todos os estudos, o intervalo de tempo entre as medidas foi indicado. No entanto, foi observada grande variação deste intervalo bem como a ausência de descrição se os sujeitos sob investigação se encontravam estáveis neste período. No que diz respeito à análise dos dados, a maioria dos estudos^{9,14-22,29-32,34-37} utilizou o CCI, o qual é considerado o mais apropriado para tal fim^{42,43}. Entretanto, coeficientes de correlação de Spearman^{28,33} ou Pearson²² e o índice Kappa^{15,19,20,31,33} também foram empregados, embora a literatura não as recomende como adequadas para estimar a concordância entre as medidas¹⁰.

Nos estudos de validade, os coeficientes de correlação de Spearman & Pearson foram os mais utilizados. Observou-se que a maior parte dos estudos^{14,15,17,19,22,23,28,32,33,36,39}

que empregaram Spearman ou Pearson também utilizaram estatística Kappa ou o método de Bland-Altman. Os coeficientes de correlação apresentaram grande variação (0,00³¹ a 0,82¹⁹) e apenas seis questionários^{15,17,19,32,33,40} apresentaram índices de correlação de magnitude moderada ou elevada, sendo dois brasileiros^{15,33}. O fato das medidas obtidas pelos questionários dependerem da capacidade dos adolescentes de recordar e quantificar a AF praticada, as diferenças de operacionalização da medida de AF entre os instrumentos de medida bem como o desconhecimento da validade das medidas subjetivas usadas como critérios de referência podem contribuir para essas discrepâncias¹⁰.

A água duplamente marcada foi usada como medida de critério em apenas um dos estudos³⁶. Segundo Sirard & Pate²⁶, este método, assim como a calorimetria indireta e a observação direta são considerados padrão-ouro para os estudos de validação dos instrumentos que mensuram o nível de AF. Contudo, possuem limitações quanto ao seu uso em decorrência da logística e do alto custo^{19,44,45}, o que pode justificar a posição única deste estudo.

Na presente revisão, observou-se que a grande maioria dos instrumentos que alcançaram coeficientes de magnitude baixa ($\leq 0,50$) foram comparados com ao menos uma medida objetiva^{9,14,16,18,21-24,29-31,34-39,41}, principalmente o acelerômetro^{9,14,16,18,21,22-24,29,31,34-37,39-41}. Esses resultados são semelhantes aos relatados nos estudos de Farias Júnior et al.¹⁰ e Chinapaw et al.¹¹ É importante destacar que embora neste estudo os métodos objetivos tenham sido considerados como medidas de referência para os estudos de validação, todos os instrumentos utilizados na mensuração da AF apresentaram alguma limitação relacionada ao custo, à logística e aos diferentes pontos de corte disponíveis. Além disso, cada instrumento mensura a AF em sua própria unidade de medida^{10,45}.

Em relação à aplicabilidade dos instrumentos, a maioria dos questionários apresentados avaliaram diferentes domínios e mensuraram diferentes períodos de tempo relacionados à AF, o que também foi encontrado nos questionários apresentados no estudo de Farias Júnior et al.¹⁰. Em sua grande maioria, os estudos apresentaram a descrição do instrumento bem como quais domínios e o período de tempo que é avaliado, o que garante uma melhor aplicabilidade do mesmo.

Este estudo possui algumas limitações, dentre as quais se destacam a baixa qualidade metodológica dos estudos, o que requer cautela na generalização dos re-

sultados. Contudo, destaca-se que a ferramenta utilizada nesta pesquisa classifica os estudos a partir da pior contagem dos itens avaliados e que na maioria dos estudos houve grande percentual de itens avaliados como bons ou excelentes, o que pode minimizar este risco. Além disso, a grande variedade de métodos e critérios utilizados nos estudos dificultou comparações mais diretas entre os instrumentos.

Como implicações clínicas, os achados desta revisão apontam para os instrumentos disponíveis na literatura com suas características psicométricas que estão disponíveis para a mensuração da AF em adolescentes. Desta forma, embora muitos instrumentos estejam disponíveis, pesquisadores ou práticos da área da saúde poderão selecionar o instrumento mais adequado tendo em vista a apresentação de todas essas importantes características. Ademais, a escolha de um instrumento depende de vários fatores e cabe ao avaliador escolher o melhor instrumento para cada situação. Além disso, deve-se considerar que a combinação de métodos objetivos e subjetivos resulta em uma avaliação mais ampla da AF⁴⁵.

Em conclusão, 36 questionários de diferentes nacionalidades tiveram suas propriedades psicométricas relatadas, dentre as quais se destacam a confiabilidade teste-reteste e a validade de critério. Foi observado baixo número de instrumentos com bons índices de confiabilidade e validade, principalmente para esta última propriedade psicométrica. Além disso, a responsividade não foi investigada em nenhum dos estudos incluídos na revisão. No Brasil, este cenário mostra-se ainda mais desafiador, pois dos seis questionários investigados, apenas um apresentou nível adequado para a confiabilidade (QAFA) e dois apresentaram níveis moderados para a validade (questionário proposto por Bastos et al.³³ e o QAFA). No entanto, ambos foram comparados a medidas de critério subjetivas. Os resultados desta revisão reforçam a importância da realização de estudos com qualidade metodológica mais adequada e com padronização das medidas usadas como critérios de referência.

Conflito de interesses

Os autores declaram não haver conflito de interesses.

Contribuição dos autores

MC Campos, participou da concepção, de todas as fases do processo da revisão sistemática, redação e revisão crítica. CA Felicidade, participou das fases do processo da revisão sistemática e redação. SC Aguiar, participou das fases do processo da revisão sistemática e redação. DSR

Vieira, participou da concepção, de todas as fases do processo da revisão sistemática, redação e revisão crítica.

Referências

1. Blair SN. Physical inactivity: the biggest public health problem of the 21st century. *Br J Sports Med.* 2009;43(1):1-2.
2. Dahlöf B. Cardiovascular disease risk factors: epidemiology and risk assessment. *Am J Cardiol.* 2010;105(Suppl 1):3A-9A.
3. Twisk JW. Physical activity guidelines for children and adolescents: a critical review. *Sports Med.* 2001;31(8):617-27.
4. Dumith SC, Gigante DP, Domingues MR, Kohl HW. Physical activity change during adolescence: a systematic review and a pooled analysis. *Int J Epidemiol.* 2011;40(3):685-98.
5. Currie C, Ziglio E, Barnekow V, Negru L. Inequalities in young people's health Key findings from the Health Behaviour in School-aged Children (HBSC) 2005/2006 survey. *Health Policy for Children and Adolescents.* Copenhagen: World Health Organization Regional Office for Europe; 2008.
6. Carson V, Ridgers ND, Howard BJ, Winkler EA, Healy GN, Owen N, et al. Light-intensity physical activity and cardiometabolic biomarkers in US adolescents. *PLoS One.* 2013;8(8):e71417.
7. Dias DF, Loch MR, Ronque ER. Perceived barriers to leisure-time physical activity and associated factors in adolescents. *Ciênc Saúde Colet.* 2015;20(11):3339-50.
8. Corder K, Ekelund U, Steele RM, Wareham NJ, Brage S. Assessment of physical activity in youth. *J Appl Physiol.* 2008;105(3):977-87.
9. Rangul V, Holmen TL, Kurtze N, Cuyppers K, Midthjell K. Reliability and validity of two frequently used self-administered physical activity questionnaires in adolescents. *BMC Med Res Methodol.* 2008;8(47).
10. Farias Júnior JC, Lopes AS, Florindo AA, Hallal PC. Validity and reliability of self-report instruments for measuring physical activity in adolescents: a systematic review. *Cad Saude Publica.* 2010;26(9):1669-91.
11. Chinapaw MJ, Mokkink LB, van Poppel MN, van Mechelen W, Terwee CB. Physical activity questionnaires for youth: a systematic review of measurement properties. *Sports Med.* 2010;40(7):539-63.
12. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res.* 2010;19(4):539-49.
13. Dollman J, Okely AD, Hardy L, Timperio A, Salmon J, Hills AP. A hitchhiker's guide to assessing young people's physical activity: Deciding what method to use. *J Sci Med Sport.* 2009;12(5):518-25.
14. Okuda M, Yoshitake N, Tanaka S, Kunitsugu I, Tan N, Uechi H, et al. Validity and reliability of physical activity questionnaire for Japanese students. *Pediatr Int.* 2011;53(6):956-63.
15. Farias Júnior JC, Lopes AS, Mota J, Santos MP, Ribeiro JC, Hallal PC. Validade e reprodutibilidade de um questionário para medida de atividade física em adolescentes: uma adaptação do Self-Administered Physical Activity Checklist. *Rev Bras Epidemiol.* 2012;15(1):198-210.
16. Giouxari A, Kavouras SA, Tambalis KD, Maraki M, Kollia M, Sidossis LS. Reliability and criterion validity of the Self-Administered Physical Activity Checklist in Greek children. *Eur J Sport Sci.* 2013;13(1):105-11.

17. Teo PS, Nurul-Fadhilah A, Foo LH. Development of a new computer-based physical activity questionnaire to estimate habitual physical activity level in Malaysian adolescents. *J Sci Med Sport*. 2013;16(4):327-31.
18. Wang C, Chen P, Zhuang J. Validity and reliability of International Physical Activity Questionnaire-Short Form in Chinese youth. *Res Q Exerc Sport*. 2013;84(2):S80-6.
19. Ben Gharbia H, Gartner A, Traissac P, Delpuech F, Maire B, El Ati J. A frequency questionnaire to estimate free-living physical activity among Tunisian preadolescent and adolescent children. *Public Health Nutr*. 2014;17(10):2253-62.
20. Bobakova D, Hamrik Z, Badura P, Sigmundova D, Nalecz H, Kalman M. Test-retest reliability of selected physical activity and sedentary behaviour HBSC items in the Czech Republic, Slovakia and Poland. *Int J Public Health*. 2015;60(1):59-67.
21. Guedes DP, Guedes JERP. Medida da Atividade Física em Jovens Brasileiros: Reprodutibilidade e Validade do PAQ-C e do PAQ-A. *Rev Bras Med Esporte*. 2015;21(6):425-32.
22. Scott JJ, Morgan PJ, Plotnikoff RC, Lubans DR. Reliability and validity of a single-item physical activity measure for adolescents. *J Paediatr Child Health*. 2015;51(8):787-93.
23. Ottevaere C, Huybrechts I, De Bourdeaudhuij I, Sjöström M, Ruiz JR, Ortega FB, et al. Comparison of the IPAQ-A and actigraph in relation to VO₂max among European adolescents: the HELENA study. *J Sci Med Sport*. 2011;14(4):317-24.
24. Murphy MH, Rowe DA, Belton S, Woods CB. Validity of a two-item physical activity questionnaire for assessing attainment of physical activity guidelines in youth. *BMC Public Health*. 2015;5:1080.
25. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev*. 2015;4(1).
26. Sirard JR, Pate RR. Physical activity assessment in children and adolescents. *Sports Med*. 2001;31(6):439-54.
27. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. COSMIN checklist manual. EMGO Institute for Health and Care Research. Amsterdam, Netherlands; 2012. Disponível em: <http://www.cosmin.nl/images/upload/files/COSMIN%20checklist%20manual%20v9.pdf>.
28. Guedes DP, Lopes CC, Guedes JERP. Reprodutibilidade e validade do Questionário Internacional de Atividade Física em adolescentes. *Rev Bras Med Esporte*. 2005;11(2):151-58.
29. Treuth MS, Hou N, Young DR, Maynard LM. Validity and reliability of the Fels physical activity questionnaire for children. *Med Sci Sports Exerc*. 2005;37(3):488-95.
30. Florindo AA, Romero A, Peres SV, Silva MV, Slater B. Desenvolvimento e validação de um questionário de avaliação da atividade física para adolescentes. *Rev Saúde Pública*. 2006;40(5):802-9.
31. Philippaerts RM, Matton L, Wijndaele K, Balduck AL, De Bourdeaudhuij I, Lefevre J. Validity of a physical activity computer questionnaire in 12- to 18-year-old boys and girls. *Int J Sports Med*. 2006;27(2):131-6.
32. Barbosa N, Sanchez CE, Vera JA, Perez W, Thalabard JC, Rieu M. A Physical Activity Questionnaire: Reproducibility and Validity. *J Sports Sci Med*. 2007;6(4):505-18.
33. Bastos JP, Araujo CL, Hallal PC. Prevalence of insufficient physical activity and associated factors in Brazilian adolescents. *J Phys Act Health*. 2008;5(4):777-94.
34. Lachat CK, Verstraeten R, Khanh LNB, Hagströmer M, Khan NC, Van NDA, et al. Validity of two physical activity questionnaires (IPAQ and PAQA) for Vietnamese adolescents in rural and urban areas. *Int J Behav Nutr Phys Act*. 2008;5:37-44.
35. Lubans DR, Sylva K, Osborn Z. Convergent validity and test-retest reliability of the Oxford Physical Activity Questionnaire for secondary school students. *Behav Change*. 2008;25(1):23-34.
36. Corder K, van Sluijs EMF, Wright A, Whincup P, Wareham NJ, Ekelund U. Is it possible to assess free-living physical activity and energy expenditure in young people by self-report? *Am J Clin Nutr*. 2009;89(3):862-70.
37. Martínez-Gómez D, Martínez-de-Haro V, Pozo T, Welk GJ, Villagra A, Calle ME, et al. Fiabilidad y validez del cuestionario de actividad física PAQ-A em adolescentes españoles. *Rev Esp Salud Publica*. 2009;83(3):427-39.
38. Moore JB, Hanes Júnior JC, Barbeau P, Gutin B, Treviño RP, Yin Z. Validation of the Physical Activity Questionnaire for Older Children in children of different races. *Pediatr Exerc Sci*. 2007;19(1):6-19.
39. Hagströmer M, Bergman P, De Bourdeaudhuij I, Ortega FB, Ruiz JR, Manios Y, et al. Concurrent validity of a modified version of the International Physical Activity Questionnaire (IPAQ-A) in European adolescents: The HELENA Study. *Int J Obes (Lond)*. 2008;32(Suppl5):S42-8.
40. Janz KF, Lutuchy EM, Wenthe P, Levy SM. Measuring activity in children and adolescents using self-report: PAQ-C and PAQ-A. *Med Sci Sports Exerc*. 2008;40(4):767-72.
41. Martínez-Gómez D, Martínez-De-Haro V, Del-Campo J, Zapatera B, Welk GJ, Villagra A, et al. Validez de cuatro cuestionarios para valorar la actividad física en adolescentes españoles. *Gac Sanit*. 2009;23(6):512-17.
42. Baumgartner TA, Chung H. Confidence limits for intraclass reliability coefficients. *Meas Phys Educ Exerc Sci*. 2001;5(3):179-88.
43. Sallis JF, Saelens BE. Assessment of physical activity by self-report: status, limitations, and future directions. *Res Q Exerc Sport*. 2000;71(2):S1-14.
44. Westerterp KR. Assessment of physical activity: a critical appraisal. *Eur J Appl Physiol*. 2009;105(6):823-8.
45. Cafruni CB, Valadão RDCD, Mello EDD. Como Avaliar a Atividade Física? *Rev Bras Ciênc Saúde*. 2012;10(33):61-71.

Recebido: 01/08/2017

Aprovado: 12/01/2019

Como citar este artigo:

Campos MC, Felicidade CA, Aguiar SC, Vieira DSR. Propriedades psicométricas de questionários de atividade física na adolescência: revisão sistemática. *Rev Bras Ativ Fis Saúde*. 2018;23:e0045. DOI: 10.12820/rbafs.23e0045

Apêndice

Estratégias de busca utilizadas nas diferentes bases de dados

MEDLINE (via Ovid)

(“adolescent” [Mesh]) OR (“students” [Mesh]) OR (“young people” [tiab] OR “youths” [tiab] OR “school-age” [tiab] OR “teen” [tiab] OR “teenager” [tiab] OR “high school” [tiab] OR “adolescence” [tiab] OR “young” [tiab]) AND (“motor activity” [Mesh]) OR (“motor activities” [Mesh]) OR (“exercises” [Mesh]) OR (“exercis*” [tiab] OR “physical activity” [tiab] OR “physical exercise” [tiab] OR “physical inactivity” [tiab]) AND (“reproducibility of results” [Mesh]) OR (“validity” [tiab] OR “reliability” [tiab] OR “reproducibility” [tiab] OR “test-retest” [tiab] OR “validation” [tiab] OR “psychometric analysis” [tiab] OR “responsiveness” [tiab])).

Web of Science

TS = (adolescent OR students OR young people OR youths OR school-age OR teen OR teenager OR high school OR adolescence OR young) AND TS=(motor activity OR motor activities OR exercis* OR physical activity OR physical exercise OR physical inactivity) AND TS = (reproducibility of results OR validity OR reliability OR reproducibility OR test-retest OR validation OR psychometric analysis OR responsiveness).

PsycINFO

(AnyField : (reproducibility of results) OR AnyField : (validity) OR AnyField : (reliability) OR AnyField : (reproducibility) OR AnyField : (test-retest) OR AnyField : (validation) OR AnyField : (psychometric analysis) OR AnyField : (responsiveness) AND PublicationYear : [2005 To 2016]) AND (AnyField : (motor activity) OR AnyField : (motor activities) OR AnyField : (exercis*) OR AnyField : (physical activity) OR AnyField : (physical exercise) OR AnyField : (physical inactivity) AND Publication Year : [2005 To 2016]) AND (AnyField : (adolescent) OR AnyField : (students) OR AnyField : (young people) OR AnyField : (youths) OR AnyField : (school-age) OR AnyField : (teen) OR AnyField : (teenager) OR AnyField : (high school) OR AnyField : (adolescence) OR AnyField : (young)).

CINAHL

(adolescent OR students OR young people OR youths

OR school-age OR teen OR teenager OR high school OR adolescence OR young) AND (motor activity OR motor activities OR exercis* OR physical activity OR physical exercise OR physical inactivity) AND (reproducibility of results OR validity OR reliability OR reproducibility OR test-retest OR validation OR psychometric analysis OR responsiveness)

SciELO

(adolescent) OR (students) OR (young people) OR (youths) OR (school-age) OR (teen) OR (teenager) OR (high school) OR (adolescence) OR (young) AND (motor activity) OR (motor activities) OR (exercis*) OR (physical activity) OR (physical exercise) OR (physical inactivity) AND (reproducibility of results) OR (validity) OR (reliability) OR (reproducibility) OR (test-retest) OR (validation) OR (psychometric analysis) OR (responsiveness)

LILACS

(mh:(adolescente)) OR (mh:(juventude)) OR (tw:(jovens)) OR (tw:(estudante)) AND (mh:(atividade motora)) OR (mh:(exercício)) OR (tw:(atividade física)) OR (tw:(exercício físico)) AND (mh:(reprodutibilidade dos resultados)) OR (mh:(validade de testes)) OR (mh:(estudos de validação)) OR (tw:(confiabilidade)) OR (tw:(reprodutibilidade)) OR (tw:(responsividade))

SPORTDiscus

(adolescent OR students OR young people OR youths OR school-age OR teen OR teenager OR high school OR adolescence OR young) AND (motor activity OR motor activities OR exercis* OR physical activity OR physical exercise OR physical inactivity) AND (reproducibility of results OR validity OR reliability OR reproducibility OR test-retest OR validation OR psychometric analysis OR responsiveness)

Cochrane

#1 = “adolescent” or “students” or “young people” or “youths” or “school-age” or “teen” or “teenager” or “high school” or “adolescence” or “young”

#2 = “motor activity” or “motor activities” or exercises or “physical activity” or “physical exercise” or “physical inactivity”

#3 = “reproducibility of results” or “validity” or “reliability” or “reproducibility” or “test-retest” or “validation” or “psychometric analysis” or “responsiveness”

#4 = #1 and #2 and #3

Tabela 1 – Percentual de classificação dos itens dos questionários para cada uma das propriedades psicométricas

Autor/Ano	Confiabilidade teste-reteste (% de itens)				Consistência interna (% de itens)				Validade de critério (% de itens)			
	Fraco	Razoável	Bom	Excelente	Fraco	Razoável	Bom	Excelente	Fraco	Razoável	Bom	Excelente
Barbosa et al. ³² 2007	0	31,8	36,4	31,8	—	—	—	—	8,3	33,3	25,0	33,3
Bastos et al. ³³ 2008	0	18,2	45,5	36,4	—	—	—	—	0	14,3	28,6	57,1
Ben Gharbia et al. ¹⁹ 2013	0	20,8	29,2	50,0	—	—	—	—	8,3	16,7	16,7	58,3
Bobakova et al. ²⁰ 2014	0	29,2	29,2	41,7	—	—	—	—	—	—	—	—
Corder et al. ³⁶ 2009	9,1	18,2	31,8	40,9	—	—	—	—	16,7	16,7	16,7	50,0
Farias Júnior et al. ¹⁵ 2012	0	12,5	25,0	62,5	—	—	—	—	16,7	16,7	33,3	33,3
Florindo et al. ³⁰ 2006	0	18,2	40,9	40,9	—	—	—	—	16,7	16,7	33,3	33,3
Gioxari et al. ¹⁶ 2013	0	18,2	36,4	45,5	25	18,75	25	31,25	0	16,7	33,3	50,0
Guedes & Guedes ²¹ 2015	0	13,6	18,2	68,2	—	—	—	—	8,3	16,7	16,7	50,0
Guedes et al. ²⁸ 2005	0	22,7	27,3	50,0	16,7	16,7	16,7	50,0	0	8,3	8,3	83,3
Hagströmer et al. ³⁹ 2009	—	—	—	—	—	—	—	—	0	16,7	25,0	58,3
Janz et al. ⁴⁰ 2008	—	—	—	—	0	25	16,7	58,3	0	33,3	16,7	50,0
Lachat et al. ³⁴ 2008	0	13,6	45,5	40,9	—	—	—	—	0	16,7	16,7	66,7
Lubans et al. ³⁵ 2008	0	22,7	31,8	45,5	—	—	—	—	0	16,7	33,3	50,0
Martínez-Gomez et al. (a) ³⁷ 2009	0	9,1	45,5	45,5	18,75	18,75	25	37,5	0	16,7	16,7	66,7
Martínez-Gomez et al. (b) ⁴¹ 2009	—	—	—	—	—	—	—	—	0	8,3	16,7	75,0
Moore et al. ³⁸ 2007	—	—	—	—	0	11,1	11,1	77,8	8,3	16,7	16,7	58,3
Murphy et al. ²⁴ 2015	—	—	—	—	—	—	—	—	0	7,1	14,3	78,6
Okuda et al. ¹⁴ 2011	0	36,4	27,3	36,4	—	—	—	—	0	25,0	25,0	50,0
Ottevaere et al. ²³ 2011	—	—	—	—	—	—	—	—	0	16,7	16,7	66,7
Philippaerts et al. ³¹ 2006	0	29,2	20,8	50,0	—	—	—	—	0	33,3	16,7	50,0
Rangul et al. ⁹ 2008	0	18,2	36,4	45,5	—	—	—	—	0	16,7	25,0	58,3
Scott et al. ²² 2015	0	22,7	18,2	59,1	—	—	—	—	0	16,7	25,0	58,3
Teo et al. ¹⁷ 2013	0	22,7	31,8	45,5	—	—	—	—	8,3	25,0	16,7	50,0
Treuth et al. ²⁹ 2005	0	18,2	50,0	31,8	—	—	—	—	0	16,7	33,3	50,0
Wang et al. ¹⁸ 2013	0	13,6	40,9	45,5	—	—	—	—	0	8,3	8,3	83,3

Valores apresentados em percentual (%).