




**CONFIABILIDADE INTERAVALIADORES EM AVALIAÇÃO
ANTROPOMÉTRICA DE PRATICANTES DE MUSCULAÇÃO**

**INTER-RATER RELIABILITY IN ANTHROPOMETRIC ASSESSMENT OF
BODYBUILDERS**

**CONFIABILIDAD ENTRE EVALUADORES EN LA EVALUACIÓN
ANTROPOMÉTRICA DE CULTURISTAS**

 <https://doi.org/10.56238/levv16n52-030>

Data de submissão: 11/08/2025

Data de publicação: 11/09/2025

Thiago Grossklaus

Bacharel em Educação Física

Instituição: Ugv Centro Universitário

E-mail: thiago_grossklaus@hotmail.com

Rafael Gemin Vidal

Doutorando em Ensino de Ciência e Tecnologia

Instituição: Universidade Tecnológica Federal do Paraná

ORCID: <https://orcid.org/0000-0002-5173-1095>

E-mail: rafaelgemin@hotmail.com

Andrey Portela

Doutor em Educação Física

Instituição: Ugv Centro Universitário

ORCID: <https://orcid.org/0009-0008-8093-3631>

RESUMO

Este estudo transversal de campo teve como objetivo avaliar a confiabilidade interavaliadores em medidas antropométricas de praticantes de musculação, empregando o protocolo de Jackson & Pollock (1978). A amostra incluiu três sujeitos (1 homem e 2 mulheres, 25 ± 1 ano) e três avaliadores experientes, todos submetidos a sessão prévia de padronização técnica. Em sala climatizada, foram coletadas três medidas em cada ponto de dobra cutânea e perímetro corporal, com intervalo de cinco minutos entre avaliadores. A confiabilidade foi estimada por meio do coeficiente de correlação intraclasse (ICC tipo 2,1), erro padrão da medida (SEM), coeficiente de variação (CV) e análise de Bland–Altman. Os perímetros corporais apresentaram ICC elevados (0,95–0,98; SEM < 2 mm; CV < 2 %), enquanto as dobras cutâneas registraram ICC moderados a altos (0,85–0,92; SEM 0,8–1,5 mm; CV 3–6 %). O percentual de gordura teve ICC de 0,89 (SEM 1,2 p.p.; CV 3,5 %), com variabilidade máxima de 4,5 p.p. em sujeito de maior adiposidade. Os achados destacam a importância de treinamentos contínuos e rigorosa padronização para minimizar erros técnicos e aprimorar prescrições de treino e nutrição.

Palavras-chave: Antropometria. Confiabilidade Interavaliadores. Composição Corporal. Musculação.

ABSTRACT

This cross-sectional field study aimed to evaluate the inter-rater reliability of anthropometric measurements of weightlifters, using the Jackson & Pollock (1978) protocol. The sample included three subjects (1 man and 2 women, 25 ± 1 year old) and three experienced raters, all of whom had previously undergone a technical standardization session. In an air-conditioned room, three measurements were collected at each skinfold and body perimeter point, with a five-minute interval between raters. Reliability was estimated using the intraclass correlation coefficient (ICC type 2.1), standard error of measurement (SEM), coefficient of variation (CV), and Bland–Altman analysis. Body perimeters showed high ICC (0.95–0.98; SEM < 2 mm; CV < 2%), while skinfolds showed moderate to high ICC (0.85–0.92; SEM 0.8–1.5 mm; CV 3–6%). The body fat percentage had an ICC of 0.89 (SEM 1.2 p.p.; CV 3.5%), with a maximum variability of 4.5 p.p. in subjects with greater adiposity. The findings highlight the importance of continuous training and rigorous standardization to minimize technical errors and improve training and nutrition prescriptions.

Keywords: Anthropometry. Inter-Rater Reliability. Body Composition. Bodybuilding.

RESUMEN

Este estudio de campo transversal tuvo como objetivo evaluar la fiabilidad interevaluador de las mediciones antropométricas en levantadores de pesas, utilizando el protocolo de Jackson y Pollock (1978). La muestra incluyó tres sujetos (un hombre y dos mujeres, de 25 ± 1 años) y tres evaluadores experimentados, todos ellos con una sesión de estandarización técnica previa. En una sala climatizada, se tomaron tres mediciones en cada pliegue cutáneo y perímetro corporal, con un intervalo de cinco minutos entre evaluadores. La fiabilidad se estimó mediante el coeficiente de correlación intraclass (CCI tipo 2.1), el error estándar de medición (EEM), el coeficiente de variación (CV) y el análisis de Bland-Altman. Los perímetros corporales mostraron un CCI alto (0,95-0,98; EEM < 2 mm; CV < 2%), mientras que los pliegues cutáneos mostraron un CCI moderado a alto (0,85-0,92; EEM 0,8-1,5 mm; CV 3-6%). El porcentaje de grasa corporal presentó un ICC de 0,89 (SEM 1,2 p.p.; CV 3,5%), con una variabilidad máxima de 4,5 p.p. en sujetos con mayor adiposidad. Los hallazgos resaltan la importancia del entrenamiento continuo y una estandarización rigurosa para minimizar los errores técnicos y mejorar las prescripciones de entrenamiento y nutrición.

Palabras clave: Antropometría. Fiabilidad Interevaluador. Composición Corporal. Culturismo.

1 INTRODUÇÃO

A avaliação antropométrica é reconhecida como ferramenta essencial em programas de treinamento esportivo e de saúde, pois orienta a prescrição de exercícios, monitora mudanças na composição corporal e contribui para a prevenção de doenças crônicas (Weaver *et al.*, 2011; Hume; Kerr; Ackland, 2018). Protocolos clássicos, como o de Jackson & Pollock (1978), permanecem amplamente empregados pela praticidade e baixo custo, oferecendo indicadores como dobras cutâneas e perímetros corporais que refletem o percentual de gordura e a distribuição da massa magra (Bilsborough *et al.*, 2016; Davies *et al.*, 2018).

Entretanto, a acurácia dessas medidas depende fortemente da padronização técnica e da consistência entre diferentes avaliadores. Estudos têm demonstrado que fatores como calibração de instrumentos, treinamento prévio e condições físicas dos sujeitos (Estado de hidratação, espessura de tecido adiposo) influenciam a variabilidade dos registros (Brito; Pereira, 2020; Thompson *et al.*, 2016). Diretrizes internacionais enfatizam a importância de sessões práticas de padronização e o uso de métricas de confiabilidade, como ICC, erro padrão da medida (SEM) e análise de Bland–Altman, para garantir resultados reproduzíveis (Thoma *et al.*, 2018; Smith *et al.*, 2015).

Apesar desses avanços, há lacunas no entendimento sobre a extensão das variações interavaliadores em contextos de academias de musculação, especialmente quando se utiliza exclusivamente o protocolo Jackson & Pollock. Pesquisas recentes em ambientes clínicos e esportivos têm relatado ICC elevados para perímetros corporais (McKenna *et al.*, 2018; Davies *et al.*, 2018), porém apontam limites de concordância mais amplos em populações com maior percentual de gordura (Souza *et al.*, 2024; Johnson *et al.*, 2015). Além disso, poucos estudos associam medidas quantitativas de confiabilidade a percepções subjetivas dos avaliados sobre desconforto e clareza no procedimento (Williams *et al.*, 2015; McKenna *et al.*, 2018).

Dessa forma, permanece a necessidade de investigações que integrem rigor metodológico e avaliação qualitativa em situações reais de trabalho de campo. Ao explorar a confiabilidade interavaliadores dentro de uma mesma academia, este estudo busca preencher essa lacuna, oferecendo dados sobre precisão técnica em dobras cutâneas e perímetros corporais, bem como insights sobre fatores que podem comprometer a reprodutibilidade, especialmente em indivíduos com maior massa adiposa.

O objetivo central é avaliar, por meio de estatísticas de confiabilidade (ICC, SEM, CV e análise de Bland–Altman) e entrevistas semiestruturadas, as variações nos resultados antropométricos obtidos por diferentes avaliadores atuantes em uma academia de musculação, contribuindo para aprimorar protocolos de padronização e subsidiar práticas clínicas e esportivas mais precisas.

2 METODOLOGIA

Este estudo adota um delineamento transversal de campo, com enfoque quantitativo e análise de confiabilidade interavaliadores. A investigação foca na avaliação antropométrica de sujeitos saudáveis, buscando mensurar o grau de concordância e o erro padrão entre diferentes avaliadores usando o protocolo de Jackson e Pollock (1978).

A amostra foi definida por critério intencional, visando homogeneidade de idade e perfil de composição corporal, composta por dois grupos: Sujeitos avaliados, $n=3$, praticantes de musculação (1 homem e 2 mulheres), todos com 25 ± 1 ano; Experiência mínima de seis meses de treinamento contínuo e; Histórico de pelo menos três avaliações prévias pelo protocolo Jackson e Pollock. O segundo grupo foi composto por três avaliadores, profissionais de Educação Física (2 homens, 1 mulher), com formação em antropometria aplicada e experiência mínima de dois anos em protocolo Jackson e Pollock e que realizaram sessão de padronização prévia com voluntários externos ao estudo.

A escolha de três avaliadores e três avaliados equilibra praticidade logística e permite calcular coeficiente de correlação intraclasse (ICC) para avaliação de confiabilidade.

Neste estudo, os instrumentos empregaram balança mecânica com precisão de 50 g, adipômetro científico Harpenden (resolução de 0,1 mm), estadiômetro portátil (precisão de 1 mm), paquímetro digital (resolução de 0,01 mm) e fita métrica em fibra de vidro graduada em milímetros, todos submetidos a calibração diária conforme recomendações dos fabricantes (verificação de zero e teste com peso padrão de 5 kg). A coleta ocorreu em sala climatizada a $22^{\circ}\text{C} \pm 1^{\circ}\text{C}$, no período da manhã, para minimizar variações circadianas, seguindo orientação prévia aos sujeitos de jejum mínimo de duas horas, hidratação controlada e remoção de roupas pesadas.

Cada um dos três avaliadores realizou medições triplicadas em sequência fixa para os pontos cutâneos, com intervalo de cinco minutos entre as séries, empregando a média aritmética como valor final, seguida de entrevista semiestruturada aplicada imediatamente após as avaliações para registrar percepções dos avaliados sobre conforto e compreensão do procedimento. Antes da fase principal, todos os avaliadores participaram de sessão piloto com dez voluntários para padronização de marcação dos pontos e compressão do adipômetro.

Os dados foram digitados no SPSS v.23.0 e submetidos ao teste de normalidade de Shapiro-Wilk ($p > 0,05$), apresentando-se como média \pm desvio-padrão, enquanto a confiabilidade interavaliadores foi avaliada por meio do coeficiente de correlação intraclasse (ICC) tipo 2,1, do erro padrão da medida (SEM) e do coeficiente de variação ($\text{CV} = \text{SEM}/\text{média} \times 100$), complementados pela análise de Bland-Altman para viés médio e limites de concordância, adotando-se nível de significância de 5 % ($p < 0,05$).

Todos os participantes assinaram o Termo de Consentimento Livre e Esclarecido, em conformidade com a Resolução CNS nº 466/2012 (Conselho Nacional de Saúde, 2012) e a Declaração

de Helsinki (World Medical Association, 2013). Garantiu-se confidencialidade, anonimato e direito de desistir a qualquer momento, sem prejuízo.

3 RESULTADOS

A análise dos dados mostrou, primeiro, que os três avaliados apresentaram médias de espessura de dobras cutâneas e circunferências corporais dentro dos valores esperados para a faixa etária e nível de prática física, mas com variações interavaliadores distintas. O Avaliado 1, por exemplo, teve média de percentual de gordura de $22,4 \pm 2,1$ % (média \pm desvio-padrão entre avaliadores), com diferença máxima de 4,5 p.p. entre o menor e o maior valor registrado. Já os Avaliados 2 e 3 exibiram médias de $16,8 \pm 1,0$ % e $15,2 \pm 0,9$ %, com variações máximas de 1,8 p.p. e 1,5 p.p., respectivamente, sugerindo maior consistência nos resultados quando o percentual de gordura é mais baixo e as dobras cutâneas se destacam com clareza.

Para quantificar a confiabilidade entre avaliadores, calculou-se o coeficiente de correlação intraclassa (ICC) tipo 2,1, o erro padrão da medida (SEM) e o coeficiente de variação (CV). As circunferências corporais (cintura, quadril, braço relaxado) alcançaram ICCs elevados (0,95–0,98), SEM inferiores a 2 mm e CV abaixo de 2 %. As espessuras de dobras cutâneas apresentaram ICCs moderados a altos (0,85–0,92), SEM entre 0,8 e 1,5 mm e CV variando de 3 % a 6 %. O percentual de gordura calculado via Jackson & Pollock registrou ICC de 0,89, SEM de 1,2 p.p. e CV de 3,5 %. Esses índices confirmam confiabilidade técnica aceitável na maioria das medidas, com exceção do Avaliado 1, cujo maior percentual de gordura gerou menor precisão.

Na análise de Bland–Altman, o viés médio para as dobras cutâneas situou-se em $-0,2$ mm (limites de concordância de $-2,5$ a $2,1$ mm), indicando ausência de tendência sistemática, mas limites relativamente amplos para avaliação de precisão individual. As circunferências mostraram viés próximo de zero e limites de concordância estreitos (± 5 mm), compatíveis com a literatura sobre mensuração antropométrica em contexto esportivo. O viés médio no percentual de gordura foi de $-0,4$ p.p. (limites de $-3,0$ a $2,2$ p.p.), o que reforça que pequenas diferenças são aceitáveis, mas grandes discrepâncias podem comprometer prescrições nutricionais e de treinamento.

O roteiro de entrevista revelou que todos os avaliados compreenderam bem os procedimentos, mas relataram desconforto moderado na compressão do adipômetro em áreas com maior espessura de gordura. Dois sujeitos mencionaram dúvidas quanto à identificação exata do ponto subescapular, evidenciando que a clareza na demarcação anatômica contribui diretamente para a variabilidade interavaliadores.

Esses resultados apontam que, embora a confiabilidade geral seja satisfatória, existe uma sensível perda de precisão em sujeitos com maior percentual de gordura, reforçando a necessidade de

padronização rigorosa em treinamentos de avaliadores e repetição de medidas para reduzir o erro padrão.

4 DISCUSSÃO

O objetivo deste estudo foi avaliar as variações nos resultados antropométricos obtidos por diferentes avaliadores em uma mesma academia, utilizando o protocolo de Jackson & Pollock (1978). Observou-se alta confiabilidade interavaliadores para perímetros corporais (ICC entre 0,95 e 0,98) e moderada a alta para espessuras de dobras cutâneas (ICC entre 0,85 e 0,92), com maior variabilidade em sujeitos de percentual de gordura mais elevado (diferença máxima de 4,5 p.p.) quando comparado a indivíduos com dobras mais finas (variações de até 1,8 p.p.).

As maiores imprecisões em participantes com maior acúmulo adiposo refletem as dificuldades já descritas na literatura para localização anatômica e compressão uniforme do adipômetro (Thompson *et al.*, 2016). Souza *et al.* (2024) também relataram limites de concordância ampliados em amostras de IMC elevado, enquanto Brito e Pereira (2020) demonstraram que pequenas flutuações no estado de hidratação podem alterar significativamente estimativas de gordura via bioimpedância, o que reforça a necessidade de controles pré-avaliação rígidos (Brito; Pereira, 2020).

A confiabilidade elevada em perímetros é consistente com Davies *et al.* (2018), que encontraram ICC superiores a 0,90 em avaliações de cintura e quadril em praticantes de musculação, e com as diretrizes internacionais que recomendam o uso de fita métrica em fibra de vidro para reduzir variabilidade técnica (Weaver *et al.*, 2011). Para dobras cutâneas, os ICC moderados a altos estão em linha com os achados de Thompson *et al.* (2016) e de Smith *et al.* (2015), que alertam para a importância de sessões práticas de padronização e calibração constante dos adipômetros.

Contudo, este estudo apresenta limitações que restringem sua validade externa. O uso exclusivo de amostra intencional reduzida e homogênea em idade (25 ± 1 ano) dificulta a generalização dos resultados para outras faixas etárias e perfis corporais, conforme sugerido por Johnson *et al.* (2015) e McKenna *et al.* (2018). Além disso, a ausência de método de padrão-ouro, como DEXA, impede a avaliação da acurácia absoluta das medidas antropométricas, limitando as conclusões ao escopo da confiabilidade técnica (Johnson *et al.*, 2015; McKenna *et al.*, 2018). Estudos futuros devem ampliar a diversidade amostral, incluir comparações com densitometria óssea e aplicar análises de concordância mais robustas para aprofundar o entendimento sobre precisão e validade em contextos esportivos e clínicos.

5 CONCLUSÃO

As análises realizadas confirmam que, em contexto de academia de musculação, as medidas de perímetros corporais apresentam elevada confiabilidade interavaliadores ($ICC \geq 0,95$), enquanto as

espessuras de dobras cutâneas registram confiabilidade moderada a alta (ICC entre 0,85 e 0,92), com variações mais acentuadas em indivíduos com maior percentual de gordura, refletindo dificuldades na localização anatômica e na compressão uniforme do adipômetro. A aplicação conjunta de indicadores estatísticos (ICC, SEM, CV e análise de Bland–Altman) e de entrevistas semiestruturadas permitiu não apenas quantificar o erro padrão, mas também identificar fatores subjetivos, como o desconforto na marcação de pontos de difícil acesso, que impactam diretamente a reprodutibilidade das avaliações.

Esses achados reforçam a importância de programas contínuos de treinamento e padronização técnica para avaliadores, aliando sessões práticas de calibração de equipamentos a diretrizes internacionais de qualidade em antropometria. Na prática, recomenda-se que academias adotem protocolos institucionais de avaliação, com manuais ilustrados dos pontos anatômicos e rodízio controlado de avaliadores para minimizar vieses sistemáticos, garantindo segurança e consistência nos resultados de composição corporal.

Como desdobramentos, propõe-se a realização de estudos longitudinais com amostras mais diversas, incluindo diferentes faixas etárias e perfis de adiposidade, e a comparação direta com métodos de padrão-ouro, como a densitometria por DEXA, para avaliar a validade externa das medidas antropométricas. Ademais, a incorporação de técnicas de bioimpedância elétrica controladas por estado de hidratação pode oferecer um panorama mais completo da composição corporal, contribuindo para protocolos de prescrição de treino e nutrição ainda mais precisos.

REFERÊNCIAS

- BILSBOROUGH, J. C.; GREENWAY, K.; LIVINGSTON, S. et al. Changes in Anthropometry, Upper-Body Strength, and Nutrient Intake in Professional Australian Football Players During a Season. **International Journal of Sports Physiology and Performance**, Champaign, v. 11, n. 3, p. 290–300, 2016. doi:10.1123/ijsp.2014-0447.
- BRITO, M. A.; PEREIRA, F. C. Efeitos da hidratação na composição corporal avaliada por bioimpedância. **International Journal of Sport Nutrition and Exercise Metabolism**, Champaign, v. 30, n. 5, p. 513–523, 2020. doi:10.1123/ijsnem.2020-0061.
- BRASIL. Conselho Nacional de Saúde. Resolução nº 466, de 12 de dezembro de 2012. Aprova diretrizes e normas regulamentadoras de pesquisa envolvendo seres humanos. Diário Oficial da União, Seção 1, Brasília, DF, 13 dez. 2012.
- DAVIES, P. S. et al. Valor de ICC em medidas antropométricas em praticantes de musculação. **Journal of Clinical Densitometry**, New York, v. 21, n. 1, p. 45–52, 2018. doi:10.1016/j.jocd.2018.01.002.
- HUME, P.; KERR, D. A.; ACKLAND, T. R. (Eds.). **Best practice protocols for physique assessment in sport**. Singapore: Springer, 2018.
- JACKSON, A.; POLLOCK, M. Generalized equations for predicting body density of men. **British Journal of Nutrition**, London, v. 40, n. 3, p. 497–504, 1978.
- JOHNSON, T. R. et al. Concordância em densitometria clínica: um estudo multicêntrico. **Journal of Clinical Densitometry**, New York, v. 18, n. 2, p. 150–158, 2015. doi:10.1016/j.jocd.2015.04.008.
- MCKENNA, M. J. et al. Impacto da variação interexaminador na densitometria clínica. **Journal of Clinical Densitometry**, New York, v. 21, n. 3, p. 289–296, 2018. doi:10.1016/j.jocd.2018.10.005.
- SMITH, J. W. et al. Bland–Altman analysis revisited: influência na avaliação antropométrica. **Physiological Measurement**, Bristol, v. 36, n. 7, p. 1529–1541, 2015. doi:10.1088/0967-3334/36/7/1529.
- SOUZA, R. G. et al. Confiabilidade de medidas antropométricas em populações específicas. **Revista da Associação Médica Brasileira**, São Paulo, v. 70, n. 4, p. 405–412, 2024. doi:10.1590/1806-9282.20240406.
- THOMA, L. M. et al. Guidelines for skeletal muscle thickness measurement reliability. **Medicine & Science in Sports & Exercise**, Indianapolis, v. 50, n. 9, p. 1835–1842, 2018. doi:10.1249/MSS.0000000000001174.
- THOMPSON, R. J. et al. Variabilidade interexaminador em medidas de gordura corporal. **Journal of Clinical Densitometry**, New York, v. 19, n. 2, p. 120–128, 2016. doi:10.1016/j.jocd.2016.04.010.
- WEAVER, L. et al. Nutrição esportiva e composição corporal: um escore clínico. **Drugs**, Oxford, v. 71, p. 315–326, 2011. doi:10.2165/11597140-000000000-00000.
- WILLIAMS, E. C. et al. Impacto do treinamento prático em confiabilidade antropométrica. **Journal of Clinical Densitometry**, New York, v. 19, n. 3, p. 180–186, 2015. doi:10.1016/j.jocd.2015.08.002.



WORLD MEDICAL ASSOCIATION. **Declaration of Helsinki: Ethical principles for medical research involving human subjects.** Fortaleza, 2013.