

IMPACTO DA EXPERIÊNCIA EM RADIOLOGIA NA MENSURAÇÃO DE
ÁREA E DENSIDADE DE MÚSCULO ESQUELÉTICO EM IMAGENS
TOMOGRÁFICAS PARA A AVALIAÇÃO DE SARCOPENIA

Dissertação de Mestrado

Guilherme Pucci Stangler

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS DA SAÚDE: CARDIOLOGIA
E CIÊNCIAS CARDIOVASCULARES**

**IMPACTO DA EXPERIÊNCIA EM RADIOLOGIA NA MENSURAÇÃO
DE ÁREA E DENSIDADE DE MÚSCULO ESQUELÉTICO EM
IMAGENS TOMOGRÁFICAS PARA A AVALIAÇÃO DE SARCOPENIA**

Autor: Guilherme Pucci Stangler

Orientador: Prof. Dr. Felipe Soares Torres

Co-orientador: Dr. Murilo Foppa

*Dissertação submetida como requisito para
obtenção do grau de Mestre ao Programa de
Pós-Graduação em Ciências da Saúde, Área de
Concentração: Cardiologia e Ciências
Cardiovasculares, da Universidade Federal do
Rio Grande do Sul*

Porto Alegre
2019

Agradecimentos

Ao meu orientador Prof. Dr. Felipe Soares Torres, por compartilhar seus conhecimentos e principalmente pelo exemplo profissional e pessoal durante toda a minha formação como radiologista.

Ao meu co-orientador Dr. Murilo Foppa, por ter oportunizado e viabilizado o desenvolvimento deste projeto e pelo constante estímulo ao meu desenvolvimento acadêmico.

Ao Dr. Raúl San José Estépar, professor adjunto de radiologia da Faculdade de Medicina de Harvard, por ter criado as ferramentas necessárias para que este trabalho fosse possível.

Ao meu colega Eduardo Neumann Tavares e ao aluno de medicina Edson Kenzo Mizushima, por terem acreditado neste projeto e dedicado seu tempo para a realização do mesmo.

À Dr^a Cristiane Melere, por ter confiado a mim grande parte de seus dados, facilitando a execução deste projeto.

Aos idealizadores do ELSA no Rio Grande do Sul, Prof. Dr. Bruce Bartholow Duncan e Prof^a. Dr^a. Maria Inês Schmidt, por tornar esse trabalho viável e pela dedicação constante ao progresso do conhecimento científico em nosso país.

Ao Hospital de Clínicas de Porto Alegre, ao Serviço de Radiologia e ao Programa de pós-graduação em Cardiologia e Ciências Cardiovasculares da Universidade Federal do Rio Grande do Sul, serviços de excelência e prestígio em suas áreas, cujos profissionais são incansáveis na melhoria da saúde de nossa população.

Aos meus pais e meu irmão, que mesmo à distância contribuíram com seu carinho e apoio.

À minha companheira Juliana, por ser a fonte constante de alegria em minha vida.

Sumário

1. Introdução	7
2. Revisão da literatura	8
2.1. Estratégias para localizar e selecionar informações	8
2.2. Sarcopenia	8
2.3. Diagnóstico de sarcopenia	13
2.4. Métodos seccionais de imagem como método de mensuração	16
3. Justificativa	19
4. Objetivos	20
5. Referências bibliográficas da revisão de literatura	21
6. Artigo original	25
7. Conclusões e considerações finais	45

Lista de abreviaturas e siglas

BIA	Análise de bioimpedância
BMI	<i>Body mass index</i>
CCI	Coeficiente de correlação intraclasse
CCS	<i>Coronary calcium score</i>
CID-10	Classificação Internacional de Doenças e Problemas relacionados à Saúde
CIP	<i>Chest imaging platform</i>
CT	<i>Computed tomography</i>
CV	<i>Coefficient of variation</i>
DEXA	Densitometria por dupla emissão de raios-X
ELSA	Estudo Longitudinal da Saúde do Adulto
HU	<i>Hounsfield units</i>
ICC	<i>Intraclass correlation coefficient</i>
IM	Índice musculoesquelético
L3	Terceira vértebra lombar
L3-L4	Transição entre a terceira e quarta vértebras lombares
MD	<i>Mean difference</i>
PACS	<i>Picture archiving and communication system</i>
RM	Ressonância magnética
SD	<i>Standard deviation</i>
SMI	<i>Skeletal muscle index</i>
SMV	<i>Skeletal muscle volume</i>
TC	Tomografia computadorizada
UH	Unidades Hounsfield
LOA	Limits of agreement
MRI	Magnetic resonance imaging

Resumo

Introdução: Medidas da área seccional muscular em imagens tomográficas do abdômen são utilizadas em estudos de sarcopenia, mas há pouca evidência sobre o impacto da experiência prévia em radiologia na obtenção dessas medidas, sendo este o objetivo deste estudo. **Métodos:** Imagens tomográficas de 195 pacientes foram analisadas por dois residentes de radiologia e um estudante de medicina. Foram calculados o coeficiente de correlação intraclasse (CCI) e o coeficiente de variância para as variáveis de área e densidade musculoesqueléticas, cujas medidas também foram comparadas utilizando gráficos de Bland-Altman. **Resultados:** A concordância intraobservador e interobservador encontrada foi excelente para todas as comparações, sendo que fatores como sexo, idade e índice de massa corporal não influenciaram de maneira significativa a mensuração das variáveis estudadas. **Conclusões:** Nosso estudo não encontrou impacto significativo da experiência prévia em radiologia na mensuração da área e da densidade musculoesqueléticas para a avaliação da sarcopenia.

Palavras-chave: concordância intraobservador, concordância interobservador, sarcopenia, índice musculoesquelético, tomografia computadorizada.

Abstract

Introduction: Measurements of cross-sectional muscle area in abdominal computed tomography (CT) scans are widely used in sarcopenia studies, but there is little evidence on whether previous experience in radiology have any influence in these measurements, and that is the objective of this study. **Methods:** CT images of 195 patients were analyzed by two radiology residents and one medical student. Intra and Inter-rater agreement for skeletal muscle area and density was assessed using the intraclass correlation coefficient (ICC), coefficient of variance and compared by Bland-Altman plots. **Results:** Both Intra and Inter-rater agreement were excellent for skeletal muscle area and density for all comparisons. Factors like age, sex and body mass index did not have a significant impact on these measurements. **Conclusion:** We found no significant impact of previous experience in CT imaging when performing measurements of skeletal muscle area and density for the assessment of sarcopenia.

Keywords: intra-rater agreement, inter-rater agreement, sarcopenia, skeletal muscle index, computed tomography.

1. Introdução

Nos últimos anos, observou-se um aumento importante no número de estudos publicados sobre sarcopenia. Este fato é evidente tanto pelo número crescente de publicações sobre o tema como pelo surgimento de periódicos quase exclusivamente focados no estudo das doenças de consumo muscular (ex: *Journal of Cachexia, Sarcopenia and Muscle*). O envelhecimento da população e consequente aumento na incidência de fragilidade e de doenças crônicas tem levado a um importante crescimento no número de pacientes sarcopênicos em nosso meio, com consequências importantes para a sociedade como um todo (1,2,3).

A sarcopenia continua sendo uma doença pouco conhecida e diagnosticada no Brasil, muitas vezes considerada simplesmente como parte do processo de senescência humana. A utilização de estratégias de rastreamento e aplicação de critérios diagnósticos bem estabelecidos é essencial para que os pacientes sob risco sejam identificados e adequadamente tratados, como forma de prevenir danos à sua saúde. Esta ação é especialmente útil para aqueles pacientes que iniciam muito cedo a perda de massa muscular, geralmente de forma secundária a doenças crônicas que cursam com o consumo exagerado ou que gerem imobilidade prolongada (1,4).

Como grande parte do diagnóstico depende da adequada detecção de uma baixa massa ou qualidade muscular, é essencial que os métodos de avaliação empregados sejam fáceis de aplicar, rápidos e confiáveis, capazes de ser reproduzidos em diferentes centros de pesquisa e por diferentes profissionais (1).

Este estudo tem como objetivo avaliar a concordância intraobservador e interobservador das medidas de área e densidade do músculo esquelético obtidas em cortes tomográficos do abdômen realizados como parte da avaliação complementar de uma coorte de pacientes adultos no Rio Grande do Sul. Com isso, se espera agregar evidências à literatura para que essas medidas sejam cada vez mais utilizadas na avaliação e diagnóstico dos pacientes sob risco de sarcopenia.

2. Revisão da literatura

2.1. Estratégias para localizar e selecionar informações

A revisão de literatura buscou dados e estudos tratando sobre as definições, epidemiologia e diagnóstico da sarcopenia e síndrome sarcopênica, bem como sobre métodos para a mensuração da massa muscular corporal. A estratégia de busca envolveu as bases Scielo, PubMed e o banco de teses do CAPES, no período de 1990 até 2018. As buscas foram realizadas através dos descritores “sarcopenia”; “sarcopenic syndrome”; “computed tomography”; “low muscle mass” e suas combinações.

2.2. Sarcopenia

A sarcopenia é definida pelo atual Consenso Europeu como uma “condição patológica musculoesquelética progressiva e generalizada, associada com risco aumentado de efeitos adversos incluindo quedas, fraturas, deficiência motora e morte” (1). O tema tem sido objeto de numerosas publicações nos últimos anos e houve crescimento do interesse da comunidade científica no avanço do conhecimento sobre esta doença.

Apesar de não ter sido o primeiro grupo de pesquisadores a trabalhar com o tema, a publicação do primeiro consenso europeu no ano de 2010 (2) foi extremamente importante ao sintetizar de forma sistemática as evidências disponíveis sobre definição, diagnóstico e possíveis tratamentos da síndrome sarcopênica. Desde então, diversos grupos de pesquisadores de outras partes do mundo publicaram seus próprios consensos (3,4,5). Essas publicações, apesar de apresentarem discretas variações de terminologia e definições, concordam em reforçar um aspecto principal: a sarcopenia deve ser considerada uma doença por si só e não deve ser encarada somente como um achado derivado de outras patologias ou como parte do processo natural do envelhecimento.

A prevalência estimada de sarcopenia é bastante variável (2,4), especialmente quando se leva em consideração a grande quantidade de métodos e critérios que podem ser utilizados para o seu diagnóstico. Uma metanálise recente que incluiu 35 estudos e 58.404 pacientes estimou a prevalência desta doença em aproximadamente 10%, tanto em homens quanto mulheres (6). Os autores deste trabalho destacam, contudo, a importante heterogeneidade dos métodos utilizados para a mensuração da massa muscular e o possível impacto nos resultados. O grupo de estudos asiático em sarcopenia, em seu Consenso (4), cita ainda que a prevalência de sarcopenia seja de aproximadamente 20% nos pacientes acima de 65 anos e 50% acima dos 80 anos de idade. O Consenso asiático também demonstra preocupação com o número absoluto de pessoas que estão atingindo a faixa etária de maior risco (acima dos 65 anos) em populações que apresentaram crescimento explosivo na primeira metade do Século XX.

A sarcopenia pode ser dividida em primária ou secundária de acordo com as bases fisiopatológicas identificadas. O Consenso Europeu define como sarcopenia primária quando não são identificadas outras causas ou fatores de risco, quando passa a ser considerada secundária (**Figura 1**).

Envelhecimento (Sarcopenia primária)

- Perda de massa muscular relacionada ao envelhecimento

Doenças crônicas

- Doenças inflamatórias
- Neoplasias
- Osteoartrose/Artrites
- Doenças neurológicas

Falta de atividade física

- Comportamento sedentário
- Restrições por deficiência

Má nutrição

- Pouca ingestão ou má-absorção
- Anorexia medicamentosa
- Obesidade

Figura 1. Fatores que influenciam na patogênese da sarcopenia. Fonte: adaptado de Cruz-Jentoft et al (2018).

O estudo da sarcopenia antigamente era restrito a implicações em perda de mobilidade e risco de quedas, mas ganhou força a partir de estudos que encontraram associações de sarcopenia com desfechos desfavoráveis, especialmente em pacientes com doenças crônicas que cursam com consumo excessivo das reservas metabólicas. Em um estudo que avaliou 1473 pacientes com neoplasias sólidas do trato gastrointestinal e respiratório, se observou menor sobrevida em pacientes com baixo Índice musculoesquelético (IM), com menor densidade muscular e naqueles com perda ponderal superior a 8% (7), mesmo quando controlado para outros fatores. Neste mesmo estudo, também foi encontrada associação entre baixa densidade muscular e *Diabetes mellitus*, baixa atividade física e distrofias musculares, achado que poderia estar relacionado a uma perda da qualidade muscular, bem como sua depleção volumétrica. Outros estudos correlacionaram sarcopenia com pior prognóstico em pacientes com carcinoma hepatocelular (8), tratados ou não com sorafenibe (9), carcinoma urotelial (10) e outras neoplasias sólidas (11). Estes achados são importantes pois caracterizam a sarcopenia não somente como um limitador físico, mas também como um fator prognóstico independente.

Também foi demonstrada associação da sarcopenia com mortalidade em pacientes com hepatopatia crônica. Um estudo avaliou 298 pacientes cirróticos que haviam realizado tomografias de abdome para pesquisa de hepatocarcinoma, avaliando a influência da massa muscular na predição de mortalidade quando associada ao Model for End-Stage Liver Disease (MELD), obtendo um novo escore denominado MELD-sarcopenia. O estudo concluiu que a inclusão da sarcopenia ao escore era superior ao escore MELD clássico na predição de mortalidade nos pacientes cirróticos. Este efeito era mais perceptível nos pacientes com baixa pontuação no escore original (12).

Desfechos pós-operatórios também foram correlacionados com baixa massa muscular. Um estudo que incluiu 144 pacientes submetidos a hepatectomia demonstrou maior taxa de mortalidade em pacientes sarcopênicos (9,7% no grupo sarcopênico e 1,4% no grupo sem sarcopenia, $p = 0,021$). Apesar disso, a taxa de complicações diretamente relacionadas a cirurgia, com necessidade de reintervenção, foi similar nos dois grupos (13).

Atuando como fator de risco para obesidade e como estado pró-inflamatório, a sarcopenia tem um efeito deletério sobre o sistema cardiovascular. Em um estudo sobre composição corporal de 2943 pacientes coreanos por meio de densitometria

(14), encontrou-se relação entre a sarcopenia e fatores de risco para doença cardiovascular, tais como resistência insulínica e síndrome metabólica. Fatores de risco foram mais prevalente em pacientes que apresentaram obesidade sarcopênica, definida pela coexistência de sarcopenia e obesidade.

Um grupo de estudos canadense, por sua vez (15), estudou o efeito isolado da sarcopenia nos fatores de risco cardiovascular em mulheres obesas após a menopausa. Os resultados, embora limitados pela pequena amostra, sugerem aumento no risco cardiovascular das pacientes com obesidade sarcopênica em relação àquelas simplesmente obesas.

Outro estudo, publicado em 2014 (16), acompanhou 4252 pacientes classificados através de medidas antropométricas em quatro grupo: pacientes eutróficos, sarcopênicos, obesos e obesos sarcopênicos. Neste estudo, tanto a sarcopenia quanto a obesidade aumentaram a mortalidade cardiovascular e geral. Porém, não houve diferença significativa na mortalidade cardiovascular para os pacientes obesos sarcopênicos em relação aos pacientes isoladamente obesos ou sarcopênicos.

A sarcopenia também apresenta associação com comorbidades psiquiátricas, presumivelmente pelo aumento nos índices de dependência e redução das atividades diárias de um modo geral. Uma metanálise recente (17) estudou a associação entre depressão e sarcopenia, identificando que as duas patologias estão associadas, mesmo após análise ajustada (OR 1.82, 95% CI 1.16 - 2.86). Como limitação do estudo, destaca-se a utilização apenas de estudos incluindo análises de bioimpedância (BIA) e densitometria por dupla emissão de raios-x (DEXA), sendo a associação com depressão maior nos estudos que utilizaram BIA como método diagnóstico. De acordo com diversos estudos recentes (1), a BIA pode ser utilizada como critério diagnóstico mas seu uso é desencorajado por maior imprecisão nas medidas e falta de validação para algumas populações.

Outro estudo (18) aplicou um questionário de qualidade de vida a 387 pacientes que foram classificados como sarcopênicos ou não de acordo com diversos critérios clínicos, inclusive por aquele proposto pelo Consenso Europeu (2). Para esta comparação, observou-se uma menor média de pontuação no questionário para os pacientes sarcopênicos em relação aos pacientes com massa e função muscular

normais (56.3 ± 13.4 vs 68.0 ± 15.2 , $p < 0.001$). Estes achados apoiam a noção de que os pacientes com sarcopenia apresentam menor qualidade de vidas, índices de dependência e sintomas depressivos.

O impacto econômico da sarcopenia também já foi demonstrado. Em uma análise de custos hospitalares os pacientes sarcopênicos tiveram risco até cinco vezes maior de apresentar custos elevados em relação a pacientes sem sarcopenia (19). Outro estudo, de forma similar, concluiu que a sarcopenia é um fator independente relacionado a aumento dos custos de internação hospitalar, com incremento estimado em cerca de 58% para pacientes abaixo de 65 anos de idade e 34% para pacientes com idade superior ou igual a 65 anos (20). Esses estudos foram capazes de estimar custos reais em um ambiente hospitalar controlado, porém não foram encontradas pesquisas que estimem o custo do paciente sarcopênico para a sociedade em termos de perda de capacidade laboral, custos ambulatoriais ao sistema de saúde e órgãos de previdência.

A importância desta patologia do ponto de vista social, econômico e de saúde levou recentemente os organizadores da Classificação Internacional de Doenças e Problemas Relacionados à Saúde (CID-10) a classificar a sarcopenia como doença, sob o código M62.84 (21). Esta inclusão no CID-10 foi considerada importante por muitos pesquisadores, pois reflete a preocupação da comunidade científica e médica com o tema. A mudança, segundo alguns autores, também irá facilitar o diagnóstico e tratamento pois os pacientes poderão utilizar planos e seguradoras de saúde (22) para a terapêutica necessária.

2.3. Diagnóstico de Sarcopenia

Os critérios diagnósticos instituídos pelo Consenso Europeu de 2010 (2) foram amplamente utilizados em pesquisas nos últimos anos, porém o grupo responsável decidiu revisar estes critérios em publicação de 2018 (1). A maior mudança diz respeito à obrigatoriedade da avaliação da massa muscular para o início do tratamento. Na última edição, conforme demonstrado na **Figura 2**, alteração isolada nos testes de força muscular já é considerada motivo suficiente para determinar um diagnóstico de “sarcopenia provável” e desencadear intervenções. O principal motivador desta mudança foi a dificuldade em comprovar perda da massa ou da qualidade muscular de pacientes a nível ambulatorial. Os autores argumentam que, embora diversos métodos tenham sido validados para mensuração da massa muscular, a maior parte dos pacientes não pode ou não deseja ser submetida a testes adicionais e os autores não veem razão em adiar o início do tratamento sob risco de agravos da saúde do paciente.

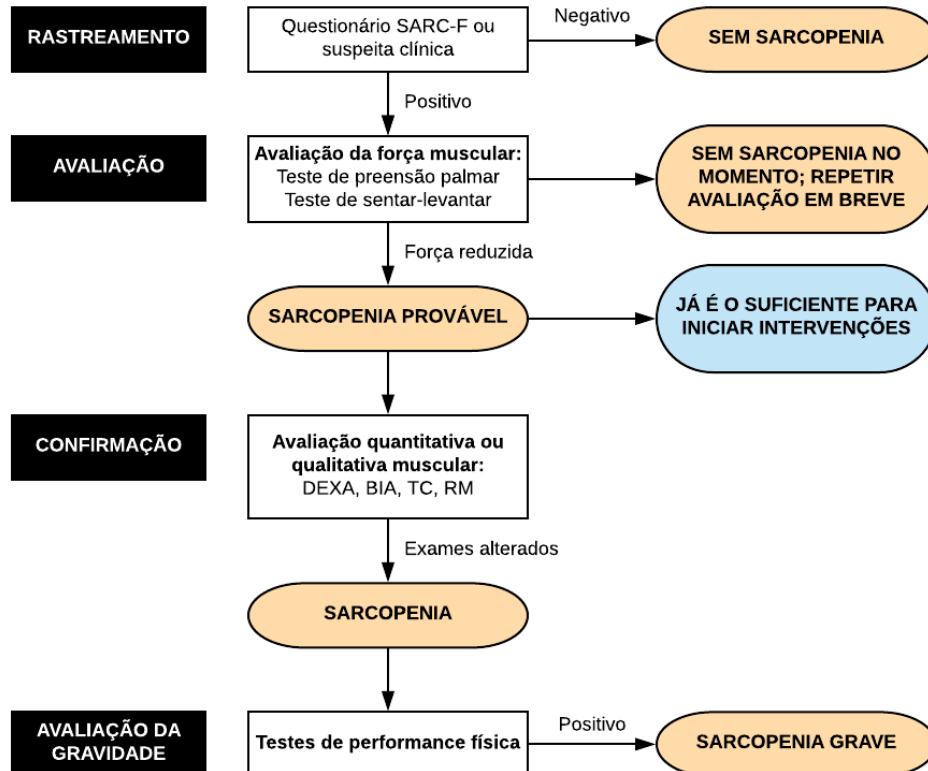


Figura 2. Algoritmo diagnóstico da sarcopenia. Fonte: Adaptado de Cruz-Jentoft et al (2018). BIA, Análise de bioimpedância; DEXA, Densitometria por dupla emissão de raios-x; RM, Ressonância magnética; TC, Tomografia computadorizada.

A suspeita clínica e testes validados devem ser o primeiro passo para identificar paciente em risco de sarcopenia. Para a avaliação rotineira, pelo menos duas escalas foram validadas para aplicação, sendo que a mais utilizada e recomendada pelo Consenso é a SARC-F (**Tabela 1**). O questionário é de fácil aplicação e consiste em perguntas sobre a vida diária, tais como dificuldade em carregar objetos minimamente pesados, necessidade de assistência para locomoção e troca de decúbito, bem como número de quedas no último ano. Os valores variam de 0 a 10, onde 0 significa ausência de sintomas e 10 paciente extremamente sintomático. Os autores encontraram risco aumentado de redução da força muscular em testes padronizados (velocidade de marcha, força de preensão palmar, entre outros) em pacientes com escore acima de 3, recomendando que estes pacientes sejam considerados casos positivos e sejam submetidos a avaliação formal (23).

Tabela 1. Escore SARC-F para o screening de sarcopenia

Pergunta	0 pontos	1 ponto	2 pontos
Você tem dificuldade em levantar e carregar um peso de 4,5 kg? *	Nenhuma	Alguma	Muita ou incapaz
Você tem dificuldade em caminhar através de uma sala/quarto?	Nenhuma	Alguma	Muita, usa equipamento auxiliar ou incapaz
Vocês tem dificuldade em sair de uma cadeira ou cama?	Nenhuma	Alguma	Muita ou não consegue sem ajuda
Vocês tem dificuldade em subir um lance de escadas com 10 degraus?	Nenhuma	Alguma	Muita ou incapaz
Quantas quedas você teve no último ano?	Nenhuma	1-3 quedas	4 ou mais quedas

Fonte: Malmstrom et al (2016). * peso do objeto considerado como 10 libras no artigo original, valor aproximado em kg.

A avaliação da força muscular é o principal componente diagnóstico e pode ser avaliada de duas maneiras principais: força de preensão palmar e teste de sentar-levantar com cinco repetições (1). A mensuração da força de preensão palmar é simples e de baixo custo, sendo que a mesma possui limites da normalidade estabelecidos em estudos populacionais (24). Resumidamente, são consideradas

alteradas medidas de força abaixo de 27 kg em homens e 16 kg em mulheres. Quando a avaliação da força nos membros superiores não é possível ou o dinamômetro não está disponível, a avaliação dos membros inferiores pode ser necessária. Para essa finalidade, uma das opções é a utilização do teste de sentar-levantar (5 vezes) e valores acima de 15 segundos são considerados alterados (25). Por não necessitar de equipamentos e servir também como medida de resistência à atividade física, o teste de sentar-levantar é recomendado pelo Consenso para aplicação em lugares que não dispõem do material necessário para avaliação da força de preensão palmar.

Diversos métodos já foram utilizados para a mensuração da massa muscular esquelética, sendo que métodos como a avaliação da creatinina e potássio séricos e medidas antropométricas simples não são mais consideradas opções válidas (1). São clinicamente indicadas as medidas da massa muscular apendicular avaliadas por: densitometria por dupla emissão de raios-x (26); Análise de bioimpedância (27); Avaliação da área seccional de músculo esquelético ao nível da coluna lombar (28-31) por tomografia computadorizada (TC) ou ressonância magnética (RM).

A DEXA é amplamente validada e utilizada na prática clínica, sendo que os valores de referência para massa muscular apendicular já estão bem estabelecidos (32). A avaliação da massa muscular por BIA é controversa, pois diversos fatores podem influenciar as medidas, especialmente o estado nutricional e hídrico do paciente. Além disso, a validação de pontos de corte para a BIA se baseia em normalização por comparação com DEXA em estudos de populações idosas europeias, com validade externa limitada (1).

2.4. Métodos seccionais de imagem como método de mensuração

A RM e a TC já são utilizadas como métodos de mensuração da massa muscular há muitos anos (1,4). A RM é principalmente utilizada em estudos experimentais em pacientes de pesquisa, pois é um método caro e demorado, especialmente quando são obtidas imagens de corpo inteiro. A TC possui um tempo de aquisição relativamente rápido, porém os potenciais riscos ao irradiar o paciente em um exame de corpo inteiro não são justificáveis quando há alternativas menos prejudiciais. Além disso, a obtenção de volumetria de corpo inteiro é um processo demorado, sendo inviável para aplicação clínica em larga escala (31).

Em busca de uma solução para este problema, diversos pesquisadores buscaram meios de correlacionar medidas obtidas em cortes seccionais de TC ou RM com a volumetria de corpo inteiro. Estes estudos encontraram correlações excelentes de cortes tomográficos do abdome com a massa muscular corporal total, sendo que o melhor nível para a avaliação se encontra aproximadamente 5 centímetros acima do disco entre a quarta e a quinta vértebras lombares (L4-L5). Na maior parte das pessoas, esta medida corresponde ao corpo vertebral da terceira vértebra lombar (L3) ou à transição entre a terceira e a quarta vértebras lombares (L3-L4) (27-29).

Dessa forma, a avaliação seccional de músculo esquelético se tornou rapidamente um dos principais testes utilizados para o diagnóstico da sarcopenia (1). Esta utilização de cortes tomográficos do abdome é especialmente interessante pois a inclusão de diversos grupos musculares (parede abdominal, musculatura paravertebral e psoas) é capaz de refletir melhor o trofismo muscular do que a análise de músculos isoladamente (31). A disponibilidade de estudos tomográficos do abdome em grandes cortes também facilita o uso deste método, pois não é necessária obtenção de novos estudos e as imagens podem ser analisadas inclusive de forma retrospectiva.

A análise isolada da área muscular lombar é falha, pois as variações de tipo físico e altura influenciam muito na composição corporal. O conceito de Índice musculoesquelético, bastante similar ao Índice de Massa Corporal, foi inicialmente utilizado nos estudos de Prado et al. (11) e agora é amplamente difundido na literatura. Seu valor é derivado pela divisão da área musculoesquelética ao nível de L3 (cm^2) pelo quadrado da altura do paciente (m^2). Um estudo determinou os valores de limite

inferior do IM para uma população caucasiana saudável em $41,6\text{cm}^2/\text{m}^2$ para homens e $32,0\text{cm}^2/\text{m}^2$ para mulheres (33). A possível variabilidade destas referências com relação a fatores étnicos, contudo, deve ser considerada e estes dados não podem ser utilizados de forma inequívoca na população brasileira.

Além da avaliação da massa muscular, também foi introduzida recentemente nos critérios diagnósticos de sarcopenia a avaliação da qualidade do músculo esquelético. Referente ao método tomográfico, a qualidade muscular é essencialmente avaliada através da medida de sua densidade em Unidades Hounsfield (UH), como análise indireta da lipossustituição intramuscular e conseqüentemente da qualidade contrátil do músculo (1). No estudo acima relatado (33), os autores também determinaram um valor limítrofe inferior de densidade em sua amostra, sendo de 29,3 UH em homens e 22,0 UH em mulheres, abaixo dos quais o músculo pode ser considerado lipossustituído.

Valores baixos de densidade muscular estão diretamente associados a alto conteúdo lipídico intracelular (34,35), achados que foram confirmados tanto por análise direta de fibras musculares como por correlação com ressonância magnética através de técnicas avançadas de imagem que detectam a presença de água e gordura em um mesmo voxel, permitindo o cálculo direto da fração de gordura presente em um determinado tecido corporal (36).

Apesar de numerosos estudos já utilizarem essas medidas como referência, o Consenso Europeu ainda é bastante conservador na recomendação de qual exame deve ser utilizado para a avaliação da massa e da qualidade muscular, em parte devido à falta de dados sobre a repetibilidade das medidas. Citando o mesmo: “As técnicas para a mensuração da massa muscular estão disponíveis em alguns, porém não em todos os locais. Conforme os instrumentos e métodos para a avaliação da qualidade muscular sejam desenvolvidos e refinados no futuro, esse parâmetro deverá ser importante como um fator definidor de sarcopenia” (1).

Alguns estudos já avaliaram a concordância intra e interobservador para as medidas tomográficas visando a avaliação da sarcopenia. Em uma série de 12 casos, avaliados tanto por RM como por TC, identificou-se um coeficiente de correlação intraclasse (CCI) acima de 0,9 para todas as medidas tomográficas de área realizadas por dois radiologistas ($p < 0,001$) (37). Uma limitação desse estudo é o uso de radiologistas para a realização das análises, pois em teoria seriam mais habilitados a diferenciar músculo de outros tecidos corporais por sua experiência em imagiologia.

Outro estudo que incluiu 10 pacientes demonstrou CCI para análise interobservador variando entre 0,42 e 0,96 quando foram comparadas as medidas de dois observadores sem experiência prévia em radiologia (38). Nesse estudo, a primeira avaliação de alguns grupos musculares apresentou uma correlação intraclasse muito baixa, que aumentou conforme foi realizada uma segunda medida, seja por melhor definição dos parâmetros para ambos os avaliadores ou maior experiência no uso do *software*. Como limitação importante de ambos, destaca-se o pequeno número de casos avaliados.

Em um estudo com 50 pacientes, também se observou um CCI acima de 0.97 para análise interobservador de medidas de área musculoesquelética, porém o estudo não avaliou a densidade muscular dos pacientes (39).

Assim sendo, torna-se evidente a necessidade de mais estudos que avaliem a variabilidade intraobservador e interobservador das medidas de área e especialmente da densidade musculoesquelética. É importante também que se avalie o impacto que a experiência prévia em radiologia tem sobre a precisão dessas medidas.

3. Justificativa

Demonstrar que a mensuração dos parâmetros musculoesqueléticos através de cortes seccionais tomográficos é uma aplicação rápida e confiável de estimar a massa e a qualidade muscular corporal poderia levar mais pesquisadores a utilizar esta medida como subsídio diagnóstico da sarcopenia. Além disso, é importante a validação do método para uso de qualquer pesquisador com o treinamento adequado e não somente daqueles com experiência prévia em radiologia, aumentando o alcance da medida e facilitando seu uso para estudos populacionais.

4. Objetivos

- **Objetivo primário:**

Avaliar a análise intraobservador e interobservador das medidas de área e densidade musculoesqueléticas em uma amostra de pacientes do Estudo Longitudinal da Saúde do Adulto (ELSA) que realizaram cortes tomográficos abdominais.

- **Objetivo secundário:**

Determinar o impacto da experiência em radiologia abdominal dos observadores nas medidas de músculo esqueléticos obtidas por tomografia computadorizada.

Mensurar a influência de fatores tais como peso, índice de massa corporal e idade na correlação intraobservador e interobservador para as medidas de músculo esquelético.

5. Referências bibliográficas da revisão de literatura

1. Cruz-Jentoft AJ, Bahat G, Bauer J et al. Sarcopenia: revised European consensus on definition and diagnosis. *Age Ageing* 2019; 48: 16-31.
2. Cruz-Jentoft AJ, Baeyens JP, Bauer JM et al. Sarcopenia: European consensus on definition and diagnosis: report of the European working group on sarcopenia in older people. *Age Ageing* 2010; 39: 412–23.
3. Studenski SA, Peters KW, Alley DE et al. The FNIH sarcopenia project: rationale, study description, conference recommendations, and final estimates. *J Gerontol A Biol Sci Med Sci* 2014; 69: 547–58.
4. Chen LK, Liu LK, Woo J et al. Sarcopenia in Asia: consensus report of the Asian working group for sarcopenia. *J Am Med Dir Assoc* 2014; 15: 95–101.
5. Morley JE, Abbatecola AM, Argiles JM et al. Sarcopenia with limited mobility: an international consensus. *J Am Med Dir Assoc* 2011; 12: 403–9.
6. Shafiee G, Keshtar A, Soltani A et al. Prevalence of sarcopenia in the world: a systematic review and meta- analysis of general population studies. *J Diab Metab Disord* 2017; 16: 21.
7. Martin L, Birdsell L, Macdonald N, Reiman et al. Cancer cachexia in the age of obesity: skeletal muscle depletion is a powerful prognostic factor, independent of body mass index. *J Clin Oncol.* 2013; 31(12): 1539–47.
8. Fujiwara N, Nakagawa H, Kudo Y et al. Sarcopenia, intramuscular fat deposition, and visceral adiposity independently predict the outcomes of hepatocellular carcinoma. *J Hepatol* 2015; 63: 131–140.
9. Imai K, Takai K, Hanai T et al. Skeletal muscle depletion predicts the prognosis of patients with hepatocellular carcinoma treated with sorafenib. *Int J Mol Sci* 2015; 16: 9612-9624.
10. Taguchi S, Akamatsu N, Nakagawa T et al. Sarcopenia evaluated using the skeletal muscle index is a significant prognostic factor for metastatic urothelial carcinoma. *Clin Genitourin Canc* 2015; 14(3): 237-243.
11. Prado CM, Lieffers JR, McCargar LJ et al. Prevalence and clinical implications of sarcopenic obesity in patients with solid tumours of the respiratory and gastrointestinal tracts: a population-based study. *Lancet Oncol* 2008; 9(7): 629–635.

12. Montano-Loza A, Duarte-Rojo A, Meza-Junco, J et al. Inclusion of Sarcopenia Within MELD (MELD-Sarcopenia) and the Prediction of Mortality in Patients With Cirrhosis. *Clin and Translat Gastroenterol* 2015. 6(2) : e102.
13. Higashi T, Hayashi H, Taki K et al. Sarcopenia, but not visceral fat amount, is a risk factor of postoperative complications after major hepatectomy. *Int J Clin Oncol* 2015; 21(2): 310-319.
14. Chung JY, Kang HT, Lee DC, et al. Body composition and its association with cardiometabolic risk factors in the elderly: a focus on sarcopenic obesity. *Arch Gerontol Geriatr* 2013; 56:270-8.
15. Aubertin-Leheudre M, Lord C, Goulet ED, Khalil A, Dionne IJ. Effect of sarcopenia on cardiovascular disease risk factors in obese postmenopausal women. *Obesity* 2006; 14: 2277–2283.
16. Atkins JL, Whincup PH, Morris RW, Lennon LT, Papacosta O, Wannamethee SG. Sarcopenic obesity and risk of cardiovascular disease and mortality: a population-based cohort study of older men. *J Am Geriatr Soc* 2014; 62(2):253–260.
17. Chang K, Hsu TH, Wu WT, Huang KC, Han DS. Is sarcopenia associated with depression? A systematic review and meta-analysis of observational studies. *Age Ageing* 2017; 46: 1-9.
18. Beaudart C, Locquet M, Reginster JY et al. Quality of life in sarcopenia measured with the SarQoL: impact of the use of different diagnosis definitions. *Aging Clin Exp Res* 2018; 30: 307.
19. Antunes AC, Araujo DA, Verissimo MT et al. Sarcopenia and hospitalisation costs in older adults: a cross-sectional study. *Nutr Diet* 2017; 74: 46–50.
20. Sousa AS, Guerra RS, Fonseca I et al. Financial impact of sarcopenia on hospitalization costs. *Eur J Clin Nutr* 2016; 70: 1046–51.
21. <http://www.icd10data.com/ICD10CM/Codes/M00-M99/M60-M63/M62-/M62.84>. 2018 ICD-10-CM Diagnosis Code M62.84. 2018.
22. Vellas B, Fielding RA, Bens C et al. Implications of ICD-10 for sarcopenia clinical practice and clinical trials: report by the International Conference on Frailty and Sarcopenia Research Task Force. *J Frailty Aging* 2018; 7: 2–9.
23. Malmstrom TK, Miller DK, Simonsick EM et al. SARC-F: a symptom score to predict persons with sarcopenia at risk for poor functional outcomes. *J Cachexia Sarcopenia Muscle* 2016; 7: 28–36.

24. Dodds RM, Syddall HE, Cooper R et al. Grip strength across the life course: normative data from twelve British studies. *PLoS One* 2014; 9: e113637.
25. Cesari M, Kritchevsky SB, Newman AB et al. Added value of physical performance measures in predicting adverse health-related events: results from the Health, Aging And Body Composition Study. *J Am Geriatr Soc* 2009; 57: 251–9.
26. Maden-Wilkinson TM, Degens H, Jones DA et al. Comparison of MRI and DXA to measure muscle size and age-related atrophy in thigh muscles. *J Musculoskelet Neuronal Interact* 2013; 13: 320–8.
27. Yamada Y, Nishizawa M, Uchiyama T et al. Developing and validating an age-independent equation using multi-frequency bioelectrical impedance analysis for estimation of appendicular skeletal muscle mass and establishing a cutoff for sarcopenia. *Int J Environ Res Public Health* 2017; 14(7): E809.
28. Schweitzer L, Geisler C, Pourhassan M et al. What is the best reference site for a single MRI slice to assess whole-body skeletal muscle and adipose tissue volumes in healthy adults? *Am J Clin Nutr* 2015; 102: 58–65.
29. Mitsiopoulos N, Baumgartner RN, Heymsfield SB et al. Cadaver validation of skeletal muscle measurement by magnetic resonance imaging and computerized tomography. *J Appl Physiol* 1998; 85: 115–22.
30. Shen W, Punyanitya M, Wang Z et al. Total body skeletal muscle and adipose tissue volumes: estimation from a single abdominal cross-sectional image. *J Appl Physiol* 2004; 97: 2333–2338.
31. Lee SJ, Janssen I, Heymsfield SB et al. Relation between whole-body and regional measures of human skeletal muscle. *Am J Clin Nutr* 2004; 80: 1215–1221.
32. Gould H, Brennan SL, Kotowicz MA et al. Total and appendicular lean mass reference ranges for Australian men and women: the Geelong osteoporosis study. *Calcif Tissue Int* 2014; 94: 363–372.
33. van der Werf A, Langius JAE, van der Schueren MAE et al. Percentiles for skeletal muscle index, area and radiation attenuation based on computed tomography imaging in a healthy Caucasian population *Eur J Clin Nutr* 2018; 72: 288–296.
34. Goodpaster BH, Kelley DE, Thaete FL et al. Skeletal muscle attenuation determined by computed tomography is associated with skeletal muscle lipid content. *J Appl Physiol* (1985) 2000; 89: 104–10.
35. Reinders I, Murphy RA, Brouwer IA et al. Muscle quality and myosteatosis: novel associations with mortality risk: the Age, Gene/Environment Susceptibility (AGES)-Reykjavik study. *Am J Epidemiol* 2016; 183: 53–60.

36. Grimm A, Meyer H, Nickel MD et al. Evaluation of 2-point, 3-point, and 6-point Dixon magnetic resonance imaging with flexible echo timing for muscle fat quantification. *Eur J Radiol* 2018; 103: 57–64.
37. Sinelnikov A, Qu C, Fetzer DT et al. Measurements of skeletal muscle area: Comparison of CT and MR imaging. *Eur J Radiology*. 2016; 85: 1716-1721.
38. Valentin S, Yeates TD, Licka T, Elliott J. Inter-rater reliability of trunk muscle morphometric analysis. *J Back Musculoskelet* 2015; 28: 181-190.
39. Vugt JLA, Levolger S, Ghabrharan A et al. A comparative study of software programmes for cross-sectional skeletal muscle and adipose tissue measurements on abdominal computed tomography scans of rectal cancer patients. *J Cachexia Sarcopenia* 2017; 8: 285-297.

6. Artigo original

Impact of observer experience in skeletal muscle measurements in CT images.

Results from the ELSA-Brasil study

Authors: Guilherme P. Stangler^{a,b}, Eduardo N. Tavares^b, Edson K. Mizushima^a, Raúl S.J. Estépar^c, Bruce B. Duncan^d, Maria I. Schmidt^d, Cristiane Melere^{d,e}, Murilo Foppa^a, Felipe S. Torres^{a,b}

Affiliations:

^a Postgraduate Studies Program in Cardiology, School of Medicine - Universidade Federal do Rio Grande do Sul and Hospital de Clínicas de Porto Alegre. ^bDepartment of Radiology, Hospital de Clínicas de Porto Alegre, Porto Alegre, Brazil. ^cApplied Chest Imaging Laboratory, Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, USA. ^dPostgraduate Studies Program in Epidemiology, School of Medicine - Universidade Federal do Rio Grande do Sul and Hospital de Clínicas de Porto Alegre, ^e Nutrition School, Universidade do Vale do Rio dos Sinos.

Correspondence: Murilo Foppa, Hospital de Clínicas de Porto Alegre. Centro de Investigação ELSA-Brasil. Rua Ramiro Barcelos, 2350 - prédio 21 - 4º andar - Santana - CEP 90035-003 - Porto Alegre – RS, mufoppa@hcpa.edu.br, +5551-3359-6304.

Conflicts of interest: Guilherme Stangler, Eduardo Tavares, Edson Muzishima, Raúl Estépar, Bruce Duncan, Maria Schmidt, Cristiane Melere, Murilo Foppe and Felipe Torres declare that they have no conflict of interest.

Abstract

Background: Many recent publications feature the use of skeletal muscle index obtained by measuring cross-sectional area in abdominal computed tomography (CT) scans, but there is little evidence on whether previous experience in radiology or image segmentation is necessary to obtain accurate measurements. The main objective of this study is to determine the rater agreement for muscle area and density between radiology residents and a non-resident rater, in order to evaluate the role of experience in measuring skeletal muscle on CT scans.

Methods: CT images of 195 patients enrolled in the ELSA-Brasil cohort were reviewed and analyzed by three different raters. Image analysis was performed using the Chest Imaging Platform extension in the Slicer 3D software. Rater agreement was assessed using the intraclass correlation coefficient (ICC) and the coefficient of variance (CV). Measurements were then compared using Bland-Altman plots to assess bias and mean differences.

Results: Agreement measurements were excellent for skeletal muscle area and mean muscle density, with ICCs above 0.99 for all comparisons. Analysis of the coefficient of variance showed small variances, unlikely to have clinical impact. Bland-Altman analysis demonstrated a statistically significant but small mean difference between all measurements, the highest of them between both radiology residents, estimated at 2.1 cm² for skeletal muscle area and -0,42 Hounsfield Units (HU) for skeletal muscle density.

Conclusions: There is no significant impact of previous experience in CT imaging or segmentation when performing measurements of skeletal muscle area and density for the assessment of sarcopenia using this specific software. Rater agreement

was excellent for all measurements (muscle area and muscle density) and the mean differences observed in the study, however statistically significant, were considered small. These findings suggest the suitability of sarcopenia measurements in broader settings.

Keywords: intra-rater agreement, inter-rater agreement, sarcopenia, skeletal muscle index, computed tomography.

Introduction

Sarcopenia is defined by the European Consensus as a “progressive and generalized skeletal muscle disorder that is associated with increased likelihood of adverse outcomes including falls, fractures, physical disability and mortality” [1]. Other definitions were also published, but many authors agree that the sarcopenic syndrome is a disease by itself, characterized as excessive muscle atrophy and loss of muscle strength [2,3]. In recent years, measurements of skeletal muscle volume (SMV) and muscle density have been extensively used in sarcopenia research to predict disease outcomes and evaluate muscle wasting [4,5]. The diagnostic criteria of definite sarcopenia changed in the last consensus, but proof of impaired muscle function and low SMV is still required for a definitive diagnosis [1]. This relatively new research field has recently gained attention due to its implication in disease prognosis, especially in patients with chronic illnesses [6] and malignant neoplasms [2].

To estimate SMV, the alternative proposed by many authors in recent publications is to obtain the muscle area (cm²) in one CT slice at the approximate level of the third lumbar vertebrae (L3) [7,8] and normalize it by patient height [9], obtaining the Skeletal Muscle Index (SMI). This measurement has already been extensively used in research and it is accepted as a valid and faster alternative to whole-body imaging [1,3]. Muscle density, expressed in Hounsfield Units (HU), is also used as a surrogate measurement of muscle quality since some studies have demonstrated that it correlates well with intramuscular fat deposition [1,10].

The measurement of muscle area and density currently relies in manual or semi-automatic segmentation by selecting regions of interests in a CT image [11], a method which is simple and robust. However, we found no study comparing manual measurements of skeletal muscle area and density performed by operators with

previous experience in segmentation techniques in radiology and those without. Therefore, the impact of rater experience in measuring skeletal muscle area and density in CT images is unknown.

The objective of this study is to assess rater agreement for measurements of skeletal muscle area and density in abdominal CT slices at the approximate level of L3-L4 in a large sample of free-living adults, comparing the results obtained by one radiology resident with previous experience in segmentation, one radiology resident without such experience and a third rater, without experience in both fields. Through these comparisons, we aim to define whether previous experience is essential to obtain reliable measurements of both skeletal muscle area and density in the abdomen.

Materials and Methods

Population and Sample selection

The ELSA-Brasil is a prospective cohort including active and retired civil servants from six education and/or research facilities in six different Brazilian states [12]. The population eligible for this research are the subjects enrolled at the research center located in Porto Alegre, southern Brazil, which performed CT imaging for coronary calcium scoring between visits 2 and 3 of the cohort study, between the years 2015 and 2017. From the 1682 eligible patients, 200 subjects were randomly selected for this reproducibility analysis.

When performing CCS imaging, technicians also acquired one CT slice at the mid-liver level and another at the level of L3-L4, in order to study liver density and provide information about body composition. The acquisition of these additional images was pre-planned, sanctioned by the local ethics committee and authorized by the patient on written consent.

Researchers ensured that images were obtained at the correct level and included the relevant anatomic landmarks. For each patient, demographic data such as age, sex, body mass index (BMI) and comorbidities were obtained at baseline for the cohort and made available through patient code.

A sample size of at least 65 subjects was needed to estimate the intraclass correlation coefficient (ICC) with an alpha error of 0.05 and a pre-specified assurance probability of 0.9, so that our lower confidence limit was not less than 0.6 and the true value of the ICC was at least 0.8 [13]. This sample size was calculated using the ICC.sample.size package for R v3.5.0 (R Foundation for Statistical Computing, Vienna, AT).

ELSA-Brasil was approved by the ethics committee and institutional review boards in all involved centers. The supplementary study regarding coronary calcium score (CCS) and the additional abdominal CT slices was approved by the cohort's directive committee and the respective institutional review board. Patients also signed a specific informed consent for the acquisition of CCS imaging. This retrospective analysis was approved by the directive committee of the ELSA-Brasil. A waiver of additional informed consent was obtained. The study has been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments.

Imaging technique

Images were acquired in an Aquilion 64 CXL (Toshiba Medical Systems, Japan) CT Scanner, without the use of intravenous iodinated contrast media, using parameters of 120 kV and 300 mA with a large field-of-view and reconstructed into 4 mm thick transaxial slices. All images were stored in a Picture Archiving and Communication System (PACS) and later anonymized and exported to an encrypted database for analysis.

Image analysis

Measurements of skeletal muscle area were performed using the free workstation Slicer 3D®, using the Body Composition module available in the Chest Imaging Platform (CIP) extension (www.chestimagingplatform.org, Applied Chest Imaging Laboratory - Brigham and Women's Hospital). The module provides a “brush” and a stencil tool for segmentation, using the lower threshold of -29 Hounsfield Units (HU) and upper threshold of 150 HU for the selection of skeletal muscle, both validated in previous studies [1,14]. The Body Composition module generated the following quantitative outputs for the segmented results: Skeletal muscle area (mm²), minimum density (HU), maximum density (HU) and mean attenuation (HU) with corresponding standard deviation.

The muscular groups included in the skeletal muscle area measurements were: paraspinal musculature, psoas major, quadratus lumborum and the abdominal wall musculature (including the internal and external obliques, transversus abdominis and rectus abdominis), as demonstrated in **Figure 1**. Each measurement takes around 5 minutes to perform, from opening the image until the output is recorded.

Images were independently analyzed by three researchers: two radiology residents with two years of experience in interpreting body CT, one of which had previous experience in image segmentation using CIP in Slicer 3D®, and a fifth-year medical student with no previous experience in radiology or image segmentation. All raters without segmentation experience underwent the same one-hour training sessions including information about the software, abdominal anatomy and recognition of the areas of interest for this particular segmentation process. Each of the three researchers who performed the analysis was blinded to the measurements of their peers and the patient's clinical data. The radiology resident with previous experience in segmentation also performed the measurements a second time, after an interval of 3 months, to assess intra-rater reliability.



Figure 1. Example of segmentation using the Body Composition module in Slicer 3D® software. Using a fixed threshold brush tool, the rater can manually paint the regions of interest. This sample includes all relevant muscular groups: paraspinal musculature, psoas major, quadratus lumborum, internal and external obliques, transversus abdominis and rectus abdominis. **Accessible caption:** The figure shows an axial CT slice of the abdomen used for analysis in our study. The muscular regions of interest are colored in pink.

Comparisons between raters

Comparisons were made between raters in order to assess the repeatability of measurements performed in the Slicer 3D software.

Intra-rater reliability was determined by comparing both measurements performed by Resident 1, who had previous experience in image segmentation. To assess the impact of software experience, inter-rater reliability was calculated by comparing two groups: one group contained both measurements by Resident 1 and the other contained the measurements of the software-inexperienced raters, Resident 2 and Non-resident. Another analysis group was planned to define the role of previous radiology experience, comparing the results of Resident 2 and the Non-resident reader. A secondary analysis was made between the mean results for Resident 1 and the results for Resident 2 and Non-resident, to test global repeatability.

Statistical analysis

Continuous data are presented as mean with its standard deviation and data normality was assessed using the Shapiro-Wilk test. Differences in mean muscle area and density for male and female subjects were calculated through t-test for independent samples.

Rater agreement for cross-sectional skeletal muscle area, density and density standard-deviation was calculated using intraclass correlation coefficients (ICCs) with 95% confidence interval using a two-way random model for single measures with absolute agreement. The coefficient of variation was also calculated. Confidence intervals of ICCs were obtained by the bootstrap method: 1000 subsamples were taken with replacement from the original sample, the ICC was computed for each and the

95% confidence interval was derived from percentiles 2.5 and 97.5 of the distribution of ICCs. Additionally, we investigated the effect of covariates on reproducibility by meaningful stratification of continuous variables (age and body mass index). Statistical analyses were performed using IBM SPSS Statistics for Windows v.21.0 (IBM Corp. Armonk, NY, USA) and STATA v13.0 (Statacorp, College Station, TX, USA).

Bland-Altman plots were generated using MedCalc statistical software v. 18.5 (Ostend, BE) [15,16], plotting the difference between measurements of raters against the mean of the measures. Graphs present the results by mean difference with its 95% confidence interval and limits of agreement.

Results

Of the initial 200 patient images, 195 were considered adequate for analysis (Table 1). The images of four subjects had severe motion or beam-hardening artifacts and one patient had severe scoliosis limiting proper evaluation. Distribution of muscle area and density are displayed with histogram plots (**Figures 2 and 3**). Analysis of the mean between measurements obtained by the experienced reader showed that male subjects have higher skeletal muscle area ($156 \pm 22 \text{ cm}^2$ for males and $102 \pm 17 \text{ cm}^2$ for females; $p < 0.001$) and mean muscle density ($31 \pm 6 \text{ HU}$ for males and $27 \pm 6 \text{ cm}^2$ for females; $p < 0.001$) compared to female subjects. Analysis of measured variables through *t*-test for independent samples revealed a mean difference of 53.5 cm^2 between male and female subjects for skeletal muscle area ($p < 0.05$) and a mean difference of 4.5 HU for skeletal muscle mean density ($p < 0.05$), both higher in male subjects.

Table 1. Sample characteristics and descriptive analysis of skeletal muscle features of 195 random subjects from those enrolled in a single ELSA-Brasil investigation center.

	Investigated sample (n = 195)	All study center participants submitted to CT (n = 1612)
Sex		
Female	112 (57.4)	939 (57.9)
Age (years)	61.1 ± 8.3	55.7 ± 9.3
Waist-hip ratio		
Male	0.98 ± 0.06	0.98 ± 0.07
Female	0.88 ± 0.08	0.88 ± 0.07
Waist circumference (cm)		
Male	99.8 ± 9.7	99.6 ± 10.9
Female	91.8 ± 14.8	91.7 ± 12.9
Body mass index (kg/m ²)		
Male	27.2 ± 3.3	27.4 ± 3.9
Female	26.9 ± 5.6	27.4 ± 5.1
Sectional area of skeletal muscle (cm ²) *		
Male	156.2 ± 22.0	N/A
Female	102.6 ± 16.7	N/A
Mean skeletal muscle mean attenuation (HU) *		
Male	31.2 ± 5.9	N/A
Female	26.7 ± 6.3	N/A

Data is presented as n (%) for sex and mean ± standard deviation for continuous variables. * Aggregated mean of measurements obtained by the most experienced reader (resident 1). N/A: Not available.

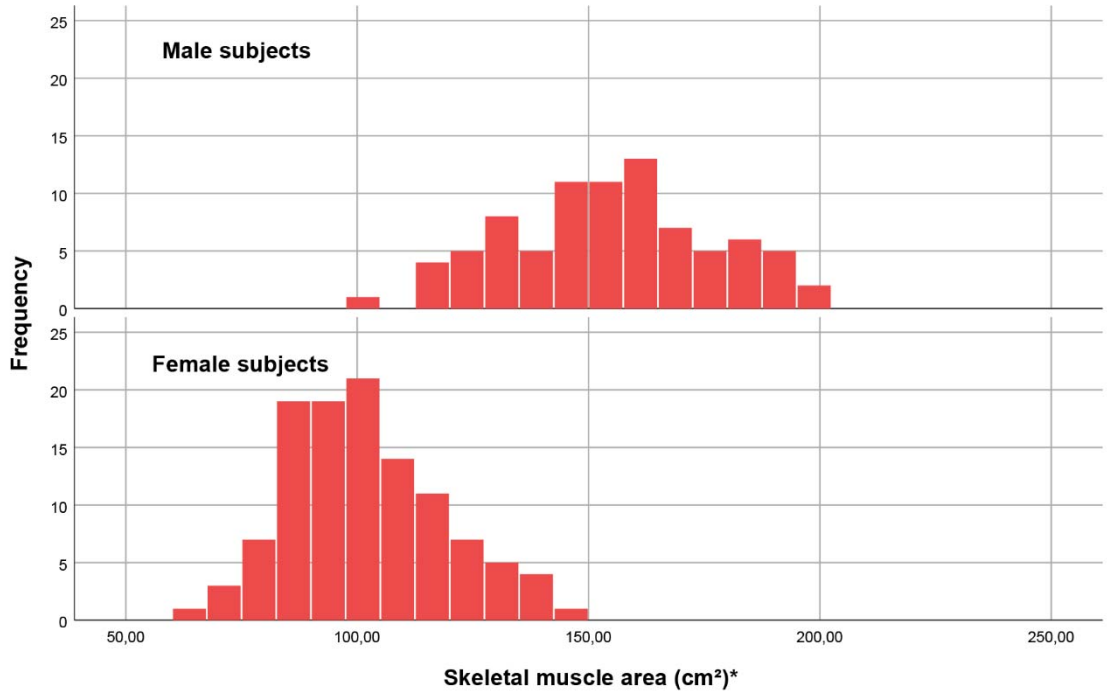


Figure 2. Histogram plots with normality curves for skeletal muscle area stratified by subject sex. * Mean of measurements by Residents 1 and 2.

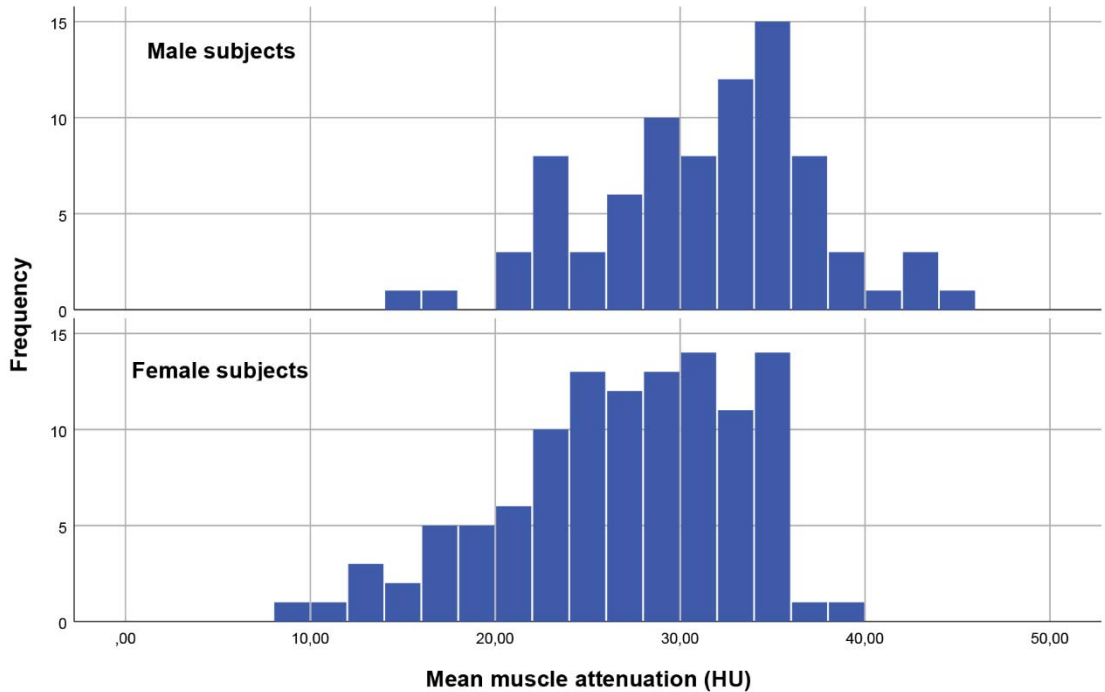


Figure 3. Histogram plots with normality curves for mean skeletal muscle attenuation filtered by subject sex. * Mean of measurements by Residents 1 and 2. **A**

Measurements of skeletal muscle area (Table 2) were similar between the radiology residents and between both residents and the non-resident rater: for Resident 1, mean SMA was $125.4 \pm 32.7 \text{ cm}^2$; for Resident 2, mean SMA was $123.4 \pm 32.5 \text{ cm}^2$ and for the Non-resident rater, mean SMA was $124.3 \pm \text{SD } 32.6 \text{ cm}^2$. Agreement was excellent in all comparisons between raters, with estimated ICCs above 0.99.

Table 2. Intra and inter-rater agreement with coefficients of variance for skeletal muscle measurements derived from CT images at the level of L3-L4. Sample of 195 subjects from the ELSA-Brasil cohort.

Rater	ICC (95% CI)	CV (%)
Skeletal muscle area (cm^2)		
Resident 1 – Two measurements	0.999 (0.999 - 1.000)	0.63
Resident 2 – Non-resident	0.998 (0.996 - 1.000)	1.13
Resident 1 (mean) – Resident 2 and Non-resident (mean)	0.998 (0.996 - 1.000)	1.14
Resident 1 (mean) - Resident 2 - Non-resident	0.998 (0.996 - 0.999)	
Skeletal muscle mean attenuation (HU)		
Resident 1 – Two measurements	1.000 (0.999 – 1.000)	0.35
Resident 2 – Non-resident	0.998 (0.996-1.000)	1.06
Resident 1 (mean) – Resident 2 and Non-resident (mean)	0.999 (0.997-1.000)	0.83
Resident 1 (mean) - Resident 2 - Non-resident	0.998 (0.997-0.999)	

CI confidence interval; CV Coefficient of variation, ICC intraclass correlation coefficient;

The coefficients of variation for skeletal muscle area ranged between 0.63% and 1.14%. Bland-Altman analysis and plots demonstrate raters differed among themselves in a systematic manner with mean difference of 1.5 cm^2 (95% Limits of

agreement [LOA] -1.1 to 4.1; $p < 0.01$) (**Figure 4**). Graphic analysis of the plots reveal measurements are roughly uniformly dispersed, and implies differences between measurements were not significantly influenced by total muscle area.

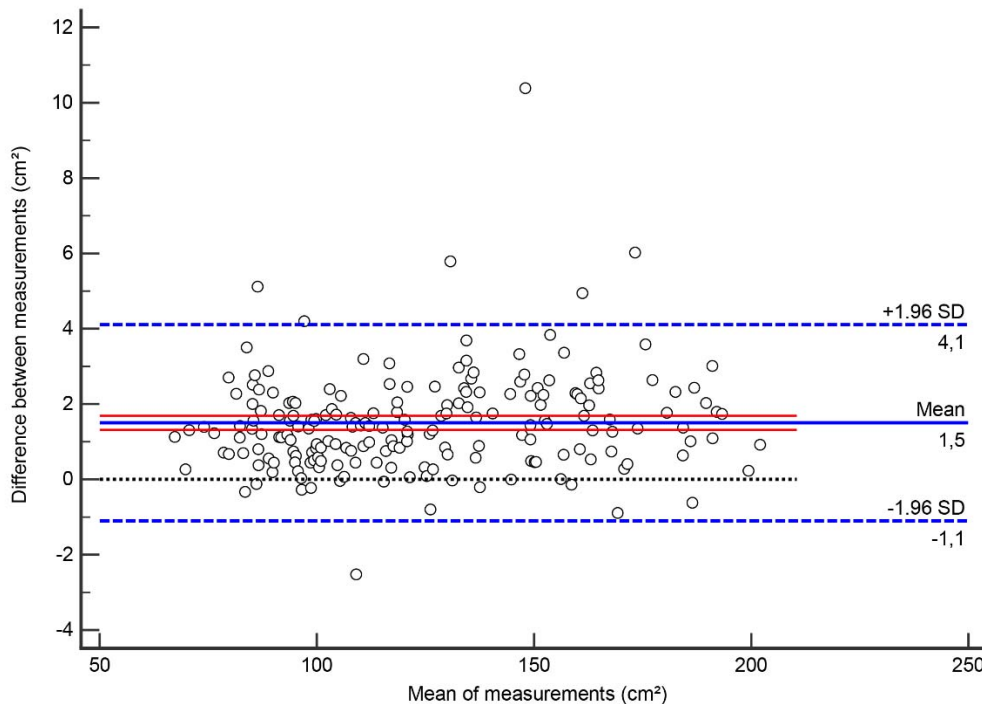


Figure 4. Bland-Altman plot for Skeletal muscle area (cm^2) comparing the measurements of Resident 1 (mean) and the mean of measurements of Resident 2 and the Non-Resident rater. Continuous line: mean difference, Dotted line: line of zero, Dashed line: upper and lower limits of agreement, Whiskers from mean difference line: confidence interval for the mean difference. **Accessible caption:** The figure shows a Bland-Altman plot for Skeletal muscle area, a scatter plot in which the differences between measurements of both readers is plotted against the mean of both measurements obtained. This graphic demonstrates a mean difference of $1,5\text{cm}^2$ that is significantly different from zero. Most of the observations in the study are within the 1,96 standard deviations limit of agreement.

Regarding skeletal muscle mean density, ICCs were also excellent, estimated to be greater than 0.99. Mean muscle density calculated was 28.6 ± 6.5 HU for Resident 1, 29.0 ± 6.5 HU and 28.7 ± 6.6 HU for the Non-resident rater. The estimated coefficient of variation for mean muscle density varied from 0.35% to 1.06%, with the highest differences observed between the measurements of both resident 2 and the Non-resident. Analysis of the Bland-Altman for mean muscle density showed raters

differed among themselves in a systematic manner with mean difference of -0.27 HU (95% Limits of agreement [LOA] -0,67 to 0,13; $p < 0.01$). Graphic analysis showed also slightly less precision for higher mean densities, especially above 25 HU (**Figure 5**).

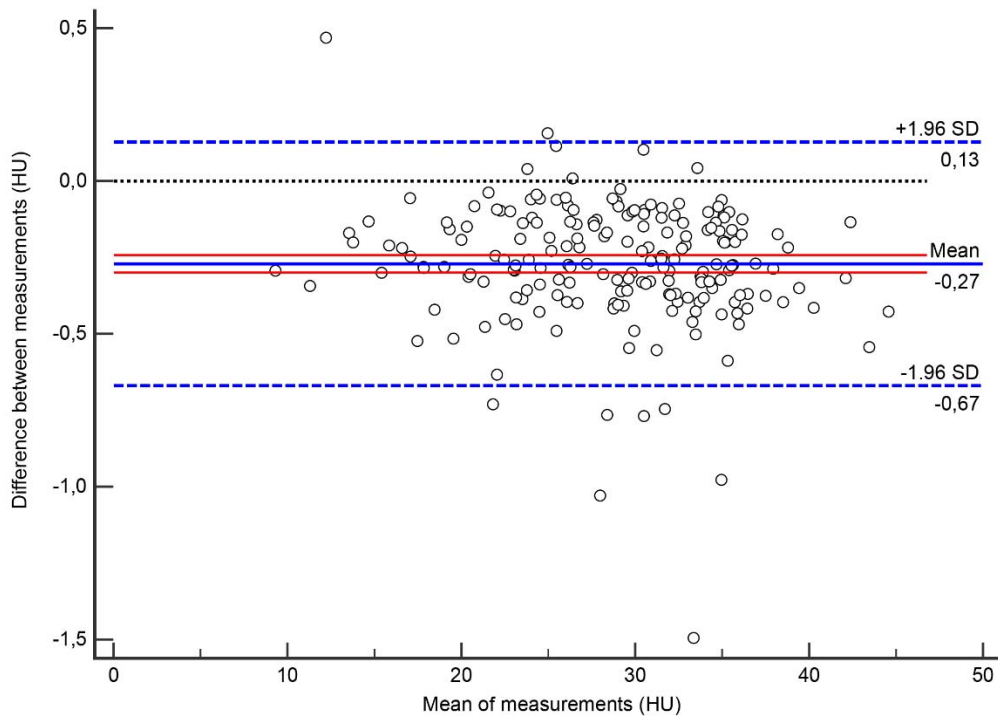


Figure 5. Bland-Altman plot for Mean muscle attenuation (cm^2) comparing the measurements of Resident 1 (mean) and the mean of measurements of Resident 2 and the Non-Resident rater. Continuous line: mean difference, Dotted line: line of zero, Dashed line: upper and lower limits of agreement, Whiskers from mean difference line: confidence interval for the mean difference. **Accessible caption:** The figure shows a Bland-Altman plot for Mean muscle attenuation, a scatter plot in which the differences between measurements of both readers is plotted against the mean of both measurements obtained. This graphic demonstrates a mean difference of -0,27 Hounsfield Units that is significantly different from zero. Most of the observations in the study are within the 1,96 standard deviations limit of agreement.

Intraclass correlation coefficients and coefficients of variation were also studied separately for the factors of sex, BMI and age. We found no impact of these measurements in ICCs for all comparisons and the coefficients of variation were also small for all comparisons (available in Supplemental material – Appendix S1).

Discussion

In our sample of 195 patients with CT studies, rater agreement for muscle area and muscle density were considered excellent for all comparisons. These findings support the high reproducibility for measurements of skeletal muscle area and density, regardless of previous experience in radiology or image segmentation. Our Bland-Altman plots demonstrated statistically significant mean differences for all comparisons and a slightly lower precision for measurements of skeletal muscle density for mean densities above 25HU. One possible explanation for this finding is that patients with higher muscle density and area tend to have more muscle tissue adjacent to bony structures, and while insertion of some pixels close to the bone may not have a large impact in total muscle area, it may influence mean muscle density since these pixels might have a density closer to the upper threshold (150HU). These mean differences, however, are considered small and are unlikely to have clinical impact, particularly regarding measurements of muscle density.

Previous studies, with smaller sample sizes, have already demonstrated excellent agreement for skeletal muscle measurements. Sinelnikov et al. [17] studied 12 subjects with chronic liver disease and found extremely high inter-rater agreement in both CT and MRI (T2 weighted images) images for skeletal muscle area when analyzed by two radiologists (0.979 and 0.909, respectively). Similarly, another author [18] demonstrated excellent intra and inter-observer agreement regarding skeletal muscle area in 50 patients. Both studies are limited by the relatively small sample size and the fact that measurements were made by people trained in abdominal imaging, radiologists in the first and doctors with extensive experience in reviewing abdominal images in the latter.

Regarding muscle density, Strandberg et al. [19] studied the inter-rater agreement for mean density measurements in mid-thigh muscles of 31 patients who had undergone anterior cruciate ligament reconstruction and found excellent agreement for measurements on all muscle groups (0.98 - 1.00). Similarly, another author [20] studied 29 patients with low back pain and measured both area and muscle density in paraspinal muscles, finding a good to excellent intraclass correlation coefficient (ICC) of about 0.82-0.85 for muscle area and 0.92-0.93 for muscle density.

Our study adds further evidence to the high ICCs for muscle area and density in a larger sample and establishes that previous training in radiology of segmentation does not affect reproducibility in a clinically significant manner. Subgroup analysis also suggest sex, weight and age do not interfere in the reliability of these measurements.

It also validates the use of Slicer 3D® as a tool for such measurements, regardless of previous experience in radiology or image segmentation. Van Vugt et al. [18] demonstrated excellent correlation between different softwares used for CT segmentation in 50 patients. The software used in our research (Slicer 3D®), combined with the CIP extension, can perform measurements in under 5 minutes and with minimal training, being a powerful tool for researchers worldwide.

At the moment, CT measurements are considered to be highly accurate for the prediction of total muscle mass [1], but the method is listed by the European consensus under the label “research studies”, and not under “clinical practice”. However, despite its lower accuracy and reproducibility, bioimpedance analysis is widely used due to its availability [22]. A recent article evaluated the performance of machine learning algorithms in measuring CT skeletal muscle features, with promising results [23]. The incorporation of these algorithms, or simpler segmentation tools in PACS softwares, as well as a better definition of reference cutoff values, might encourage radiologists

to report these data, enabling clinicians to use these measurements for clinical assessment. Fully-automated measurements are probably going to be a powerful feature in the future, but since this technology is still not widespread, and requires some degree of machine learning expertise to operate, manual or semi-automated segmentation is still a valuable tool.

Our study has some limitations, such as the relative low experience (approximately 2 years) of both radiology residents in reading CT images. However, this further adds strength to the fact that radiological experience is not a major factor when performing these measurements. Only 2.5% of images were considered inadequate for analysis and this small number of losses is unlikely to have influenced the results. Images selected for our study were obtained specifically for evaluation of body composition and further studies might be necessary to assess if images obtained in other settings (such as cancer staging) have the same repeatability. Another limitation is the fact that we are only investigating one aspect of variability, and other factors such as variability between CT scans have not been studied. Therefore, variations of muscle density and area which may be influenced by slice thickness, CT parameters and daily variations in body and water composition were not contemplated and could be assessed by future research.

Conclusion

Our study demonstrates that measures of skeletal muscle area and density have excellent intra and inter-rater agreement, even when performed by raters without previous experience in CT imaging or segmentation. Even though mean differences were statistically significant, these differences are too small to have significant clinical

impact. By analyzing a larger sample than previous repeatability studies in the field, our research demonstrated this method is suitable for measurements at a population level since measurements obtained by this method are reliable and feasible to be used in research with minimal training.

Acknowledgements

This study was funded by the Brazilian Ministry of Health (Science and Technology Department) and the Brazilian Ministry of Science, Technology and Innovation (FINEP and CNPq). Grant numbers: 01.06.0010.00 and 01.10.0643.03 (RS); 01.06.0212.00 and 01.10.0742-00 (BA); 01.06.0300.00 and 01.12.0284.00 (ES); 01.06.0278.00 and 01.10.0746.00 (MG); 01.06.0115.00 and 01.10.0773-00 (SP); and 01.06.0071.00 and 01.11.0093.01 (RJ). The authors of this manuscript certify that they comply with the ethical guidelines for authorship and publishing in the Journal of Cachexia, Sarcopenia and Muscle.

References

1. Cruz-Jentoft AJ, Bahat G, Bauer J et al. Sarcopenia: revised European consensus on definition and diagnosis. *Age Ageing* 2019; 48: 16-31.
2. McLean RR, Kiel DP. Developing consensus criteria for sarcopenia: an update. *J Bone Miner Res.* 2015; 30(4): 588-592.
3. Boutin RD, Yao L, Canter RJ, Lenchik L. Sarcopenia: Current concepts and imaging implications. *Am J Radiology.* 2015; 205: W255-W266.
4. Gibson DJ, Burden ST, Strauss BJ, Todd C, Lal S. The role of computed tomography in evaluating body composition and the influence of reduced muscle mass on clinical outcome in abdominal malignancy: a systematic review. *Eur J Clin Nutr.* 2015; 69(10): 1079-1086.
5. Prado CMM, Heymsfield SB. Lean tissue imaging: a new era for nutritional assessment and intervention. *J Parenter Enteral Nutr.* 2014; 38(8): 940-953.
6. van Vugt JL, Levolger S, de Bruin RW et al. Systematic Review and Meta-Analysis of the Impact of Computed Tomography-Assessed Skeletal Muscle Mass on Outcome in Patients Awaiting or Undergoing Liver Transplantation. *Am J Transplant* 2016; 16(8): 2277-2292.
7. Shen W, Punyanitya M, Wang ZM et al. Total body skeletal muscle and adipose tissue volumes: estimation from a single abdominal cross-sectional image. *J Appl Physiol* 2004; 97: 2333-2338.
8. Mitsiopoulos N, Baumgartner RN, Heymsfield SB, Lyons W, Gallagher D, Ross R. Cadaver validation of skeletal muscle measurement by magnetic resonance imaging and computerized tomography. *J Appl Physiol* 1998; 85(1): 115-122.
9. Baumgartner RN, Koehler KM, Gallagher D et al. Epidemiology of sarcopenia among the elderly in New Mexico. *Am J Epidemiol* 1998; 147(8): 755-763.
10. Rier HN, Jager A, Sleijfer S, Rosmalen, J, Kock MCJ, Levin MD. Low muscle attenuation is a prognostic factor for survival in metastatic breast cancer patients treated with first line palliative chemotherapy. *The Breast* 2017; 31: 9-15.
11. Gomez-Perez SL et al. Measuring abdominal circumference and skeletal muscle from a single cross-sectional CT image: a step-by-step guide for clinicians using National Institutes of Health ImageJ. *Journal of Parenteral and Enteral Nutrition* 2016; 40: 308-318.
12. Schmidt MI, Duncan BB, Mill JG et al. Cohort Profile: Longitudinal Study of Adult Health (ELSA-Brasil). *Int J Epidemiol* 2015; 44(1): 68-75.

13. Zou, G. Y. Sample size formulas for estimating intraclass correlation coefficients with precision and assurance. *Statistics in medicine* 2012; 31: 3972-3981.
14. Aubrey J et al. Measurement of skeletal muscle radiation attenuation and basis of its biological variation. *Acta Physiologica* 2014; 210: 489-497.
15. Altman DG, Bland JM. Measurements in medicine: the analysis of method comparison studies. *The Statistician* 1983; 32: 307-317.
16. Giavarina D. Understanding Bland Altman analysis. *Biochemia Medica*. 2015; 25(2): 141-151.
17. Sinelnikov A, Qu C, Fetzer DT et al. Measurements of skeletal muscle area: Comparison if CT and MR imaging. *Eur J Radiology* 2016; 85:1716-1721.
18. Van Vugt JLA, Levolger S, Gharbharan A et al. A comparative study of software programmes for cross-sectional skeletal muscle and adipose tissue measurements on abdominal computed tomography scans of rectal cancer patients. *J Cachexia, Sarcopenia Muscle* 2017; 8:285-297.
19. Strandberg S, Wretling ML, Wredmark T, Shalabi A. Reliability of computed tomography measurements in assessment of thigh muscle cross-sectional area and attenuation. *BMC Med Imaging* 2010; 10:18.
20. Hu ZJ, He J, Zhao FD, Fang XQ, Zhou LN, Fan SW. An assessment of the Intra- and inter-reliability of the lumbar paraspinal muscle parameters using CT scan and magnetic resonance imaging. *Spine* 2011; 36: E868-E874.
21. Fosbøl M, Zerahn B. Contemporary methods of body composition measurement. *Clin Physiol Funct Imaging* 2015; 35:81-97.
22. Buckinx F, Landi F, Cesari M et al. Pitfalls in the measurement of muscle mass: a need for a reference standard. *J Cachexia, Sarcopenia Muscle* 2018; 9:269-278
23. Burns JE, Yao J, Chalhoub D, Chen JJ, Summers RM. A machine learning algorithm to estimate sarcopenia on abdominal CT. *Acad Radiol* 2019.

7. Conclusões e Considerações finais

Nosso trabalho demonstrou que mesmo pesquisadores sem experiência prévia em segmentação de imagens tomográficas, quando treinados de forma adequada, podem realizar medidas de músculo esquelético de forma consistente e confiável, tanto para área como para densidade muscular. Também foi demonstrada uma excelente repetibilidade das medidas realizadas pelo mesmo observador, confirmando achados de outros pesquisadores. Apesar de terem sido encontradas diferenças médias entre as medidas significativas em nosso trabalho, as mesmas possuem uma magnitude muito pequena, como demonstrado pela diferença absoluta e pela análise dos coeficientes de variação. Estas pequenas variações podem ser consideradas pouco relevantes clinicamente e dificilmente mudariam a classificação do paciente por representar uma mudança muito pequena no cálculo do índice musculoesquelético.

Não havíamos encontrado na literatura dados para reprodutibilidade utilizando o *software* gratuito Slicer 3D. Em nosso estudo, o mesmo foi capaz de gerar medidas consistentes e reprodutíveis. Este achado é importante pois pode favorecer o estudo da sarcopenia em locais com recursos limitados que não possam arcar com os custos de *softwares* específicos para segmentação.

Muitos pacientes realizam tomografia de abdome como parte da avaliação de suas doenças de base, especialmente aqueles com doenças crônicas graves ou neoplasias. As informações sobre o trofismo e a qualidade do músculo esquelético podem ser extraídas destes estudos sem a realização de imagens adicionais e exposição do paciente a mais radiação. Para tal, basta que seja realizada uma reconstrução específica (4mm) e que sejam realizadas medidas semiautomáticas ou automáticas, das quais proviriam informações adicionais com valor diagnóstico e prognóstico para este paciente.

O fato de área e densidade muscular serem medidas nas quais a experiência do observador parece não ser um fator importante de erro aumenta a possibilidade de implementação de rotinas de extração de dados de forma rápida. Dessa forma, essas análises poderiam ser realizadas em equipes menores e sem obrigatoriedade de experiência prévia neste *software* ou em radiologia.

A maior barreira a ser quebrada continua sendo o fato de que a sarcopenia é objeto de pesquisas, porém sua abordagem na prática clínica é insuficiente. Com a melhoria dos métodos para avaliação da massa e qualidade muscular, se espera aumento no número de diagnósticos e conseqüentemente desenvolvimento de novas terapias e políticas públicas para lidar com esta patologia cada vez mais prevalente em nossa população.

**APÊNDICE A – Tabelas de coeficientes de correlação intraclass completa
para análises interobservador, incluindo análise por sexo, idade e IMC**

Table S1 – Intraclass correlation coefficients and coefficients of variation for skeletal muscle area and mean muscle attenuation for male and female subjects

Reader	ICC (95% CI)	CV (%)
Skeletal muscle area (cm²)		
Resident 1 – Two measurements		
Male subjects (n= 83)	0.998 (0.999 - 1.000)	0,38
Female subjects (n = 112)	0.997 (0.994 - 1.000)	0,88
Resident 2 – Non-radiologist		
Male subjects (n= 83)	0.995 (0.990 - 1.000)	1,01
Female subjects (n = 112)	0.994 (0.989 - 1.000)	1,25
Mean Resident 1 – Mean of Resident 2 and Non-radiologist		
Male subjects (n= 83)	0.994 (0.988 - 1.000)	1,15
Female subjects (n = 112)	0.995 (0.991 - 1.000)	1,11
Mean measurements of Resident 1 - Resident 2 - Non-radiologist		
Male subjects (n= 83)	0.993 (0.989 – 0.998)	
Female subjects (n = 112)	0.994 (0.991 - 0.998)	
Skeletal muscle mean attenuation (HU)		
Resident 1 – Two measurements		
Male subjects (n= 83)	1.000 (0.997 – 1.000)	0,24
Female subjects (n = 112)	1.000 (0.999 - 1.000)	0,44
Resident 2 – Non-radiologist		
Male subjects (n= 83)	0.997 (0.995 – 1.000)	0,96
Female subjects (n = 112)	0.998 (0.995 - 1.000)	1,14
Mean Resident 1 – Mean of Resident 2 and Non-radiologist		
Male subjects (n= 83)	0.998 (0.997 – 1.000)	0,76
Female subjects (n = 112)	0.999 (0.997 - 1.000)	0,90
Mean measurements of Resident 1 - Resident 2 - Non-radiologist		
Male subjects (n= 83)	0.998 (0.996 – 1.000)	
Female subjects (n = 112)	0.998 (0.996 – 0.999)	

CI confidence interval; CV Coefficient of variation, ICC intraclass correlation coefficient;

Table S2 – Intraclass correlation coefficients and coefficients of variation for skeletal muscle area and mean muscle attenuation for different Body Mass Index categories.

	Eutrophic (BMI < 25kg/m²)		Overweight (BMI 25 - 29,9 kg/m²)		Obese (BMI ≥ 30kg/m²)	
	ICC (95% CI)	CV (%)	ICC (95% CI)	CV (%)	ICC (95% CI)	CV (%)
Skeletal muscle area						
Resident 1 (Intra)	1.000 (0.999 – 1.000)	0,53	0.999 (0.998 – 1.000)	0,69	0.999 (0.999 – 1.000)	0,59
Resident 2 - Student	0.998 (0.996 – 1.000)	1,17	0.998 (0.996 – 1.000)	1,08	0.998 (0.995 – 1.000)	1,14
Mean R1 – Mean R2/NR	0.998 (0.995 – 1.000)	1,24	0.998 (0.995 – 1.000)	1,12	0.998 (0.996 – 1.000)	1,05
Mean R1 – R2 - NR	0.997 (0.996 – 1.000)		0.997 (0.996 – 0.999)		0.998 (0.996 – 0.999)	
Skeletal Muscle Density						
Resident 1 (Intra)	1.000 (1.000 – 1.000)	0,31	1.000 (0.999 – 1.000)	0,38	1.000 (1.000 – 1.000)	0,35
Resident 2 - Student	0.998 (0.996 – 1.000)	0,86	0.997 (0.994 – 1.000)	1,27	0.999 (0.997 – 1.000)	0,90
Mean R1 – Mean R2/NR	0.998 (0.996 – 1.000)	0,94	0.999 (0.998 – 1.000)	0,78	0.999 (0.998 – 1.000)	0,72
Mean R1 – R2 - NR	0.998 (0.996 – 0.999)		0.998 (0.996 – 0.999)		0.998 (0.997 – 0.999)	

CI confidence interval; CV Coefficient of variation, ICC intraclass correlation coefficient;

Table S3 – Intraclass correlation coefficients and coefficients of variation for skeletal muscle area and mean muscle attenuation for different age categories

	< 60 years old		≥ 60 years old	
	ICC (95% CI)	CV (%)	ICC (95% CI)	CV (%)
Skeletal muscle area				
Resident 1 (Intra)	1.000 (1.000 – 1.000)	0,58	0.999 (0.999 – 1.000)	0,66
Resident 2 - Student	0.998 (0.997 – 1.000)	1,11	0.998 (0.996 – 1.000)	1,14
Mean R1 – Mean R2/NR	0.999 (0.998 – 1.000)	0,96	0.997 (0.995 – 1.000)	1,25
Mean R1 – R2 - NR	0.998 (0.998 – 0.999)		0.997 (0.995 – 1.000)	
Skeletal Muscle Density				
Resident 1 (Intra)	1.000 (0.999 – 1.000)	0,28	1.000 (0.999 – 1.000)	0,41
Resident 2 - Student	0.996 (0.991 – 1.000)	1,03	0.998 (0.996 – 1.000)	1,07
Mean R1 – Mean R2/NR	0.998 (0.995 – 1.000)	0,80	0.999 (0.998 – 1.000)	0,86
Mean R1 – R2 - NR	0.996 (0.994 – 0.999)		0.998 (0.997 – 0.999)	

CI confidence interval; CV Coefficient of variation, ICC intraclass correlation coefficient;