Assessment of knee extension muscle strength and thickness to indirectly measure function-related parameters

JOÃO ALBERTO PINTO RAMOS

TESE DE DOUTORAMENTO APRESENTADA
À FACULDADE DE MEDICINA DA UNIVERSIDADE DO PORTO EM
INVESTIGAÇÃO CLÍNICA E EM SERVIÇOS DE SAÚDE

Assessment of knee extension muscle strength and thickness to indirectly measure function-related parameters

JOÃO ALBERTO PINTO RAMOS

DOCTORAL THESIS IN
CLINICAL AND HEALTH SERVICES RESEARCH PRESENTED TO THE
FACULTY OF MEDICINE OF THE UNIVERSITY OF PORTO

SUPERVISOR – Bernardo Manuel Sousa Pinto, MD PhD
CO-SUPERVISOR – Joana de Brito Barroso Monteiro, MD PhD

PORTO | 2023

Jury of the Doctoral Defense | *Constituição do Júri*

President | *Presidente*:

Professora Doutora Maria Dulce Cordeiro Madeira

Voting members | *Vogais*:

Professor Doutor João Paulo Neves Branco

Professor Doutor Luís Gabriel Gago Horta

Professor Doutor António Pedro Pinto Cantista

Professora Doutora Teresa Sarmento Henriques

Professor Doutor Bernardo Manuel de Sousa Pinto

Corpo Docente Catedrático da Faculdade de Medicina da Universidade do Porto

Professor Doutor Patrício Manuel Vieira Araújo Soares Silva

Professor Doutor Alberto Manuel Barros da Silva

Professor Doutor José Henrique Dias Pinto de Barros

Professora Doutora Maria Fátima Machado Henriques Carneiro

Professora Doutora Maria Dulce Cordeiro Madeira

Professor Doutor Altamiro Manuel Rodrigues Costa Pereira

Professor Doutor Manuel Jesus Falcão Pestana Vasconcelos

Professor Doutor João Francisco Montenegro Andrade Lima Bernardes

Professora Doutora Maria Leonor Martins Soares David

Professor Doutor Rui Manuel Lopes Nunes

Professor Doutor José Manuel Pereira Dias de Castro Lopes

Professor Doutor Joaquim Adelino Correia Ferreira Leite Moreira

Professora Doutora Raquel Ângela Silva Soares Lino

Professor Doutor Fernando Manuel Mendes Falcão dos Reis

Professor Doutor Francisco José Miranda Rodrigues Cruz

Professor Doutor José Paulo Alves Vieira de Andrade

Professor Doutor Jorge Manuel Silva Junqueira Polónia

Professor Doutor José Luís Dias Delgado

Professora Doutora Isaura Ferreira Tavares

Professor Doutor Fernando Carlos de Landér Schmitt

Professor Doutor Acácio Agostinho Gonçalves Rodrigues

Professora Doutora Maria de Fátima Moreira Martel

Professor Doutor João Tiago de Sousa Pinto Guimarães

Professor Doutor José Carlos Lemos Machado

Professor Doutor José Carlos de Magalhães Silva Cardoso

List of Publications

In accordance with the Doctoral studies regulation of Portuguese Universities, this dissertation comprises the following publications:

Pinto-Ramos J, Moreira T, Costa F, Tavares H, Cabral J, Costa-Santos C, Barroso J, Sousa-Pinto B. Handheld dynamometer reliability to measure knee extension strength in rehabilitation patients - A cross-sectional study. PLoS One. 2022;17(5):e0268254. doi:10.1371/journal.pone.0268254. Impact Factor: 3.752 (Quartile 1, Scimago/ Quartile 2, Web of Science)

Pinto-Ramos J, Costa-Santos C, Costa F, Tavares H, Cabral J, Moreira T, Brito R, Barroso J, Sousa-Pinto B. Reliability of Point-of-Care Ultrasound for Measuring Quadriceps Femoris Muscle Thickness. European Journal of Physical and Rehabilitation Medicine. 2022; 58(5):767-773. doi:10.23736/S1973-9087.22.07432-9. Impact Factor: 5.313 (Quartile 1, Scimago and Web of Science)

Pinto-Ramos J, Moreira T, Costa L, Costa F, Barroso J, Sousa-Pinto B. Association between knee extension strength and functional capacity after intensive care unit discharge: A six-month prospective cohort study. American Journal of Physical Medicine & Rehabilitation:10.1097/PHM.0000000000002137. 2022. doi: 10.1097/PHM. 0000000000002137. Impact Factor: 3.412 (Quartile 1, Scimago and Web of Science)

Abreviations

1MSTS - 1-Minute Sit-to-Stand

6MWT - 6 Minutes Walking Test

95%CI – 95% Confidence Interval

Apache II - Acute Physiology Assessment and Chronic Health Evaluation II

BMI – Body Mass Index

HHD - Handheld Dynamometer

ICC - Intraclass Correlation Coefficients

ICU – Intensive Care Unit

IKD - Isokinetic Dynamometer

Kg – Kilogram

m – Meter

MCID - minimal clinically important differences

mRankin - Modified Rankin Scale

MRCSS - Medical Research Council Sum Score

N – Newtons

NIHSS - Initial National Institute of Health Stroke Scale

P25;P75 – 25th and 75th Percentile

PMR - Physical Medicine and Rehabilitation

POCUS – Point-of-Care Ultrasound

PROMs - Patient-Reported Outcome Measures

s – Seconds

SAPS II - Simplified Acute Physiology Score II

SD – Standard Deviation

SF-36 - Short Form 36

SPPB - Short Physical Performance Battery

TUG - Timed Up and Go

US – Ultrasound

Acknowledgments

To Professor **Bernardo Sousa Pinto**, supervisor of this thesis. Words are not enough to thank you for the supervision of my thesis – your work ethic, knowledge and also patience were of capital importance for its conclusion. In addition, my interest in investigation and constant need to find good evidence-based support on every topic is greatly influenced by the last 12 years together, since we were young medical students. I also have to thank you for your availability to discuss all topics at all times, for the prompt feedback when needed and all the kindness without which it would have been impossible to finish this thesis in three years. It’s been an honour to work with one of the most brilliant minds I have ever known.

To Professor **Joana Barroso**, co-supervisor of this thesis, for all her teachings, support and availability. You are one of the most gifted people I know, which together with your courage makes you one of the most promising investigators in the world. You’ve inspired me to have the courage to proceed with this thesis, and your insights, corrections and inputs were paramount for this thesis to be completed. Thank you for furthering the name of Physical Medicine and Rehabilitation.

To all the staff from the PhD program, namely professors and assistants, I have to thank for the organization, carefully chosen content and assistance. A special mention to the director of the PhD program, Professor **Luís Azevedo** for his leadership and for all the knowledge transmitted during his lessons.

To all co-authors of the studies, specially to Dr. **Tiago Moreira**, Dr. **Frederico Costa** and Dr. **Liliana Costa** for their help in the study design and data collection. This thesis would have not been possible without your contribution.

To my colleagues from physical medicine and rehabilitation department, specially Dr. **João Barroso** and Dr. **Fernando Parada**, current and previous directors, and Dr. **Maria**

José Festas, my residency tutor, for their support and contribution towards my development and for the opportunity to work with them in the department.

To my friend, **Mariana Marinho** for her help with the linguistic revision.

To my parents, **César** and **Aurora**. All I have reached so far I owe it to your teachings, incentive to do more and better. Thank you for always believing in me.

To my wife, **Rita** for being my everyday companion and for sharing this journey with me, and to my daughter, **Luísa**, for giving me a new perspective of life.

To all my friends and family.

Table of contents

Abstract

Background:

Disability is a generic term for activity impairment and limitation, being an important cause of morbidity. Although there are functional related tools that estimate disability and its impact in physical rehabilitation patients, there is not a clear single gold-standard for functioning. This prompts the need to use a combination of tools, in a process which is usually time-consuming. This limits our capability to estimate patients' functioning in the outpatient setting, as well as to estimate functional improvements (and, consequently, treatment effectiveness).

The identification of accurate and reliable methods able to assess functioning in a fast and non-expensive way is, therefore, much needed. Muscle strength and thickness have been hypothesized as potential estimators of functioning, capacity and performance. The isokinetic dynamometer is the gold-standard to measure muscle strength, but its use is not always feasible in the clinical setting, given its costs and slowness. Such limitations may be overcome by the use of hand-held dynamometers (HHD) or, for the assessment of muscle thickness, point-of-care ultrasound (POCUS). While interest on these two inexpensive and portable tools to estimate functioning has been growing, their reliability has not been fully evaluated, especially in the rehabilitation setting. In addition, few studies have prospectively measured whether changes in muscle strength or thickness are correlated to those observed with function-related scales and patient-reported outcome measures (PROMs).

Aims:

This thesis main aims are: (i) to assess the intra-rater and inter-rater reliability of HHD and POCUS for the estimation of knee extension muscle strength and quadriceps

and rectus femoris thickness in patients from a rehabilitation setting, and (ii) to measure the association between changes in knee extension muscle strength or quadriceps femoris and rectus femoris muscle thickness, and changes in function-related measurements in post-Intensive Care Unit (ICU) patients. The assessment of post-ICU patients is grounded on the fact that ICU admissions tend to be associated with a rapid loss of muscle mass and strength, often resulting in physical impairment and disability. Being able to predict functional outcomes and assess functional improvement after ICU discharge is particularly challenging but would be highly beneficial.

Methods:

Three different studies were conducted with different methodological approaches:

Handheld Dynamometer Reliability to Measure Knee Extension Strength: Patients admitted into an inpatient Physical Medicine and Rehabilitation unit were consecutively included in a cross-sectional study. Each of the four observers (two experienced and two inexperienced) used HHD to perform a set of two measurements (assessment AB) followed by a second set of two measurements (assessment CD) three hours later. Two-way mixed-effects model ICCs were used to calculate intra-rater and inter-rater reliability for average, maximum and first value of the assessments. Furthermore, experienced and inexperienced observers were separately analysed for ICC comparison. Absolute and relative differences between average, maximum and first measurements of each assessment were also calculated. Lastly, we estimated Spearman correlation coefficients to assess the correlation between average knee extension peak force of the different participants and other function-related parameters.

Point-of-Care Ultrasound for Measuring Quadriceps Femoris Muscle Thickness: Inpatients were consecutively selected after admission into a Physical Medicine and Rehabilitation Department of a tertiary care hospital for a cross-sectional study. Four observers, two experienced and two inexperienced, used POCUS to measure quadriceps femoris and rectus femoris thickness. Furthermore, two experienced and two inexperienced physicians performed two sets of two assessments (assessment AB and assessment CD) of knee extension strength with HHD. The assessments were performed three hours apart and registered the average, maximum, minimum and first value. Two-way model intra and inter-rater intraclass correlation coefficients were then calculated in order to estimate the reliability between those assessments. A sub-analysis was performed with ICC values being independently estimated for experienced and inexperienced observers. We also calculated absolute and relative percent differences between the average, maximum, minimum and first measurements of AB assessments versus those of CD assessments. Finally, Spearman correlation coefficients were estimated between average muscle thickness of rectus femoris and quadriceps femoris from all observers and function-related parameters measured by a fifth independent researcher.

Knee Extension Strength and Functional Capacity after Intensive Care Unit Discharge - A six-month prospective cohort study: Adult patients without previous disability were consecutively selected after ICU discharge in this prospective cohort study. We measured knee extension strength using HHD and Rectus and Quadriceps Femoris thickness using POCUS, as well as SF-36 physical functioning, Barthel Index, International Physical Activity Questionnaire (IPAQ), 6 Minutes Walking Test (6MWT), 1-Minute Sit-to-Stand (1MSTS), Short Physical Performance Battery (SPPB), Timed Up and Go (TUG) test, handgrip strength and Medical Research Council Sum Score

(MRCSS) at baseline and three and six months after discharge. We assessed the correlation and built regression models to assess the association between evolution in knee extension strength or muscle thickness as assessed by HHD and POCUS, and evolution in function-related tests.

Results:

Handheld Dynamometer Reliability to Measure Knee Extension Strength: Twenty-nine patients were assessed. Intra and inter-rater ICC of HHD measurements were overall high (≥ 0.950 and 0.927, respectively). Higher values were found when the average of two measurements were made for estimating intra-rater ICC (ICC=0.978; 95%CI=0.969-0.985) but not for inter-rater ICC. The ICC were not statistically significantly different when calculated based on measurements performed by both inexperienced and experienced physicians. There was a moderate correlation between strength and function-related variables (including current modified Rankin, 1MSTS and Handgrip Strength).

Point-of-Care Ultrasound for Measuring Quadriceps Femoris Muscle Thickness: Twenty-nine patients were assessed. Both intra-rater and inter-rater ICC were higher than 0.888 for both quadriceps and rectus femoris muscle thickness measurements with POCUS. Reliability was highest when ICC were calculated based on the average of two measurements, with the intra-rater ICC being of 0.956 (95%CI=0.937-0.970) for rectus femoris and of 0.966 (95%CI=0.951-0.976) for quadriceps femoris, and with the inter-rater ICC being of 0.919 (95%CI=0.863-0.957) for rectus femoris and 0.945 (95%CI=0.907- 0.971) for quadriceps femoris. Experienced and inexperienced observers did not have significantly different ICC values.

Knee Extension Strength and Functional Capacity after Intensive Care Unit Discharge - A six-month prospective cohort study: Thirty patients completed the follow-up. Moderate correlation was found between knee extension strength change and changes in the SF-36 Physical Function (correlation coefficient [ρ]=0.53), 6MWT (ρ=0.38), 1MSTS (ρ=0.52) and SPPB (ρ=0.38). Baseline values and changes in knee extension strength moderately predicted evolution in SF-36 Physical function (r^2=0.32; p=0.006). Changes in muscle thickness were overall not associated with changes in function-related variables.

Conclusions:

HDD and POCUS seem to be reliable tools to measure knee extension strength and quadriceps femoris muscle thickness. Our studies suggest that for both of these tools high reliability can be achieved by either experienced or inexperienced clinicians after a short training course. There seems to be a moderate association between knee extension strength change measured with HHD and function-related parameters' change (SF-36 Physical Functioning, 6MWT, and 1MSTS) within the first 6 months of patients' discharge from ICU. Such association was not observed for quadriceps and rectus femoris muscle thickness. While further confirmatory studies using larger samples sizes, more assessments and longer follow-up periods are needed in order to confirm these results, our thesis suggests that HHD is a reliable tool that can potentially be used in the clinical practice to track functional recovery in a portable and inexpensive way.

Resumo

Introdução:

O conceito "incapacidade" é usado de forma genérica para traduzir limitações nas atividades, sendo uma importante causa de morbilidade. Embora existam instrumentos relacionados com a funcionalidade que estimam a incapacidade e o seu impacto em pacientes em Medicina Física e de Reabilitação, não existe um *gold-standard* único e claro para a avaliação de funcionalidade. Isto resulta na necessidade de usar uma combinação de intrumentos, num processo que é geralmente moroso, e que limita a capacidade de estimar o estado funcional dos pacientes em ambulatório, bem como de estimar as suas melhorias funcionais (e, consequentemente, a eficácia do tratamento instituído).

A identificação de métodos reprodutíveis e com boa acuidade, capazes de avaliar a funcionalidade de forma rápida e pouco dispendiosa é, portanto, necessária. A força e a espessura muscular foram já elencadas como potenciais intrumentos para estimar funcionalidade, capacidade e desempenho. O dinamómetro isocinético é o *gold-standard* para medir a força muscular, mas o seu uso nem sempre é viável em contexto clínico, devido ao seu custo e ao facto de a sua aplicação ser morosa. Estas limitações podem eventualmente ser ultrapassadas com o uso de dinamómetros de mão (HHD) ou, no caso da avaliação da espessura muscular, com a ecografia portátil (POCUS). Embora o interesse por estes dois instrumentos acessíveis e portáteis para estimar a capacidade funcional esteja em crescimento, a sua acuidade ainda não foi completamente avaliada, especialmente em contexto de reabilitação. Além disso, poucos estudos mediram prospectivamente se as alterações na força ou espessura muscular estão associadas com as alterações observadas em escalas relacionadas com a função e em *patient-reported outcome measures* (PROMs).

Objetivos:

Esta tese teve dois objetivos principais: (i) avaliar a acuidade intra e inter-observador do HHD e POCUS para estimar a força muscular de extensão do joelho e a espessura do quadríceps e reto femoral em pacientes no contexto de reabilitação, e (ii) medir a associação entre alterações na força muscular de extensão do joelho ou espessura dos músculos quadríceps e reto femoral e alterações nas medidas relacionadas com a função em pacientes pós-Unidade de Cuidados Intensivos (UCI). A avaliação dos pacientes pós-UCI baseia-se no facto de que os internamentos em UCI tendem a associar-se a uma rápida perda de massa e força muscular, muitas das vezes resultando em limitação funcional e incapacidade. Ter uma forma de prever os resultados funcionais e avaliar a melhoria funcional após a alta da UCI é particularmente importante para a avaliação destes doentes.

Métodos:

Três estudos diferentes foram realizados com diferentes abordagens metodológicas:

Acuidade do Dinamómetro de Mão na Medição da Força de Extensão do Joelho: Foi realizado um estudo transversal em que foram incluídos consecutivamente pacientes internados num Serviço de Medicina Física e Reabilitação. Quatro observadores (dois experientes e dois inexperientes) usaram o HHD para realizar um conjunto de duas medições (avaliação AB) seguido, três horas depois, por um segundo conjunto de duas medições (avaliação CD). Um coeficiente de correlação intra-classe (ICC) de modelo de efeitos mistos foi usado para calcular a acuidade intra- e inter-observador para a primeira medição e para a média e máximo de cada avaliação. Além disso, foram analisados separadamente os observadores experientes e inexperientes para se proceder à comparação dos respetivos ICC. Foram ainda calculadas as diferenças absolutas e relativas entre a média, a máxima e a primeira medição de cada avaliação. Por último, estimamos os coeficientes de correlação de Spearman para avaliar a correlação entre o valor médio da força de extensão do joelho dos diferentes participantes e um conjunto de diferentes variáveis relacionadas com a funcionalidade.

Ecografia Portátil para a Medição da Espessura do Músculo Quadríceps Femoral: Foram selecionados consecutivamente pacientes internados no Serviço de Medicina Física e Reabilitação de um hospital terciário para a realização de um estudo transversal. Quatro observadores, dois experientes e dois inexperientes, usaram o POCUS para medir a espessura do reto e quadríceps femoral. Cada um dos observadores fez duas séries de duas medições (avaliação AB e avaliação CD) da espessura muscular do reto e quadríceps femoral, separadas por três horas, e registaram o valor médio, máximo, mínimo e a primeira medição. Em seguida, foram calculados os ICC intra- e inter-observador para estimar a acuidade entre essas avaliações. Foi ainda realizada uma subanálise, tendo sido estimados os valores de ICC para observadores experientes e inexperientes separadamente. Também foram calculadas as diferenças absolutas e relativas entre a média, máxima, mínima e primeiras medições entre as avaliações AB e CD. Por fim, os coeficientes de correlação de Spearman foram estimados entre a espessura média dos músculos reto e quadríceps femoral e um conjunto de variáveis relacionadas com a funcionalidade medidas por um quinto observador independente.

Força de Extensão do Joelho e Capacidade Funcional após Alta da Unidade de Cuidados Intensivos – Um Estudo de Coorte Prospetivo de 6 meses: Foram selecionados consecutivamente pacientes adultos, sem limitações funcionais prévias, após a alta da UCI. Foi medida a força de extensão do joelho com o HHD e a espessura do reto e quadríceps femoral com o POCUS. Foi ainda aplicado o SF-36 *Physical Functioning*, o Índice de Barthel, o Questionário Internacional de Atividade Física (IPAQ), Teste de 6 Minutos de Marcha (6MWT), o Teste de Sentar e Levantar em 1 minuto (1MSTS), o *Short Physical Performance Battery* (SPPB), o *Timed Up and Go* (TUG), a Força de Preensão Manual e o *Medical Research Council Sum Score* (MRCSS). Todos estes testes foram realizados imediatamente após alta da UCI e três e seis meses após a alta. Foi avaliada a correlação e foram construídos modelos de regressão linear para avaliar a associação entre evolução da força de extensão do joelho ou da espessura muscular e a evolução nos parâmetros relacionados com a função.

Resultados:

Acuidade do Dinamómetro de Mão na Medição da Força de Extensão do Joelho: Vinte e nove pacientes foram avaliados. Os valores de ICC intra e inter-observador das medições com o HHD foram globalmente altos (≥ 0,950 e 0,927, respetivamente). Foram obtidos valores superiores de ICC quando se utilizou a média de duas medições para estimar o ICC intra-observador (ICC=0,978; IC95%=0,969-0,985), mas não para o ICC inter-observador. O ICC não mostrou diferenças estatisticamente significativas quando calculado com base em medições realizadas por médicos experientes ou inexperientes. Houve ainda uma correlação moderada entre a força e as variáveis relacionadas com a função (incluindo o Rankin modificado atual, o 1MSTS e a força de preensão manual).

Ecografia Portátil para a Medição da Espessura do Músculo Quadríceps Femoral: Vinte e nove pacientes foram avaliados. Relativamente às medições de espessura do músculo quadríceps e reto femoral com POCUS, tanto o ICC intra- como o inter-observador foram superiores a 0,888. A acuidade foi ainda maior quando o ICC foi calculado com base na média de duas medições, com um ICC intra-observador de 0,956 (IC95%=0,937-0,970) para o reto femoral e de 0,966 (IC95%=0,951-0,976) para o quadríceps femoral, e um ICC inter-observador de 0,919 (IC95%=0,863-0,957) para o reto femoral e 0,945 (IC95%=0,907-0,971) para o quadríceps femoral. Os observadores experientes e inexperientes não apresentaram valores de ICC significativamente diferentes.

Força de Extensão do Joelho e Capacidade Funcional após Alta da Unidade de Cuidados Intensivos – Um Estudo de Coorte Prospetivo de 6 meses: Trinta pacientes completaram o estudo. Foi encontrada uma correlação moderada entre a alteração da força de extensão do joelho e as respetivas alterações no SF-36 *Physical Functioning* (coeficiente de correlação [ρ]=0,53), 6MWT (ρ=0,38), 1MSTS (ρ=0,52) e SPPB (ρ=0,38). Os valores basais e variações na força de extensão do joelho previram moderadamente a evolução da função física do SF-36 (r^2=0,32; p=0,006). Por outro lado, mudanças na espessura do músculo não estiveram, em geral, associadas a alterações importantes nas variáveis relacionadas com a função.

Conclusões:

O HDD e o POCUS parecem ser instrumentos reprodutíveis para medir a força de extensão do joelho e a espessura do músculo quadríceps femoral. Os nossos estudos sugerem uma alta reprodutibilidade para ambos os instrumentos após um curto período de treino, quer por médicos experientes, quer por médicos inexperientes. Parece ainda haver uma associação moderada entre a variação da força de extensão do joelho medida com o HHD e a alteração dos parâmetros relacionados com a função (SF-36 *Physical Functioning*, 6MWT e 1MSTS) durante os primeiros 6 meses após a alta da UCI. Esta mesma associação não foi observada para a espessura dos músculos reto e quadríceps femoral. Embora estudos adicionais com maior tamanho amostral, maior número de avaliações e períodos de acompanhamento mais longos sejam necessários para confirmar estes resultados, esta tese sugere que o HHD é uma ferramenta reprodutível e que pode ser potencialmente usada na prática clínica para avaliar a recuperação funcional de forma acessível.

1. Introduction

1.1 Outcomes in Physical Medicine and Rehabilitation

Disability is an important cause of morbidity, being defined by the World Health Organization as an "umbrella term for impairments, activity limitations and participation restrictions" and may lead to limitations in activities of daily living, work and even personal relationships.[1, 2] Disability is not only impactful, affecting mobility and independent living, but also widely prevalent. According to the Centers for Disease Control and Prevention, one-quarter of all adults in the United States have some kind of disability; fourteen percent have severe difficulty walking and climbing stairs, and seven percent have difficulty performing daily tasks without assistance.[3] In Portugal, forty percent of the population reported some kind of long-term disability.[4] As a consequence, disability is increasingly being taken into account in the decision process, as attested by the concept of disability-adjusted life years.[5] Trying to address and attempt to eliminate or minimize disability is the main purpose of Physical Medicine and Rehabilitation (PMR), which acts, among others, to promote a normal interaction between the musculoskeletal and neurological systems, seeking to optimize functioning in all domains.[2, 6] In this context, PMR aims to improve domains, such as capacity and performance, which have a different meaning but are closely inter-related.[2] In this thesis, they will be considered altogether using the umbrella expression of "function-related parameters".[1]

The need for PMR is particularly frequent following prolonged hospitalization in the context of an acute disease or event, leading to a marked muscle mass loss resulting

[1] In the literature, similar expressions such as "function-related outcomes" can be found. However, we did not use this expression, in order not to create confusion, as not all parameters were defined as outcomes in all performed studies.[7]

in a significant functional decline.[8] While this muscle mass loss is frequently transient, there are many patients who never recover to their previous functioning level and others who have progressive and irreversible loss of muscle mass and capacity.[9]

Patients in Intensive Care Units (ICU) are particularly prone to lose muscle mass and strength, and concomitantly, functioning. During the first days of an ICU admission, an average of 3-4% of muscle mass is lost per day.[10, 11] Anti-gravity muscles are the most affected with immobilization – for example, the rectus femoris loses an average of 9% of its thickness and cross-sectional area after three ICU admission days, and 30% after ten days. [12] Nevertheless, some of these problems are temporary, with most patients subsequently recovering muscle mass, strength and functioning lost during the ICU admission.[13] However, understanding the post-ICU evolution of the patient's muscle mass, strength and function-related parameters requires the ability to accurately measure such properties.

Being able to measure functioning and its evolution over time is crucial, not only to estimate it but also because accurate measurements can help in the selection of more effective treatments or in the interruption of the interventions that are already unnecessary, leading to health gains and/or cost savings.[14] Some function-related scales, tests and questionnaires are already used in the clinical practice (including in the specific post-ICU context) to measure different dimensions of functioning and its progression. The short form 36 (SF-36) is the most frequently used function-related questionnaire in the ICU setting,[15] where it has been shown to have an acceptable validity and reliability.[16-18] Other frequently used questionnaires and tests which have been validated to measure functioning and performance include the Barthel index, handgrip strength, 6 Minutes Walking Test (6MWT), 1-Minute Sit-to-Stand (1MSTS), Timed Up and Go (TUG) test and Short Physical Performance Battery (SPPB).[19-23] Some of these parameters (such as SF-36 Physical function, 6MWT, 1MSTS and SPPB) are predictors of mortality and health-related quality of life.[24-27] Table 1 summarises the function-related parameters usually evaluated in ICU context as well as the time to perform them and their properties.

Table 1 – Function-related parameters evaluated in ICU patients

Functional parameter	**Items**	**Time to perform**	**Properties**
SF-36[28, 29]	36: Physical functioning (10 items) Physical role limitations (4 items) Bodily pain (two items) General health perceptions (5 items) Energy/vitality (4 items) Social functioning (2 items) Emotional role limitations (3 items) Mental health (5 items)	10 minutes	ICC physical functioning: 0.75 Cronbach α physical functioning: 0.94
Barthel Index[30, 31]	10: Feeding Bathing Grooming Dressing Bowel control Bladder control Toileting Chair transfer Ambulation Stair climbing	Self-report: 2-5 minutes Direct Observation: 20 minutes	Cronbach α: 0.81 Interrater ICC: 0.98 (95%CI, 0.97-0.98) κ statistic for the individual items: 0.54-0.94 Standard error of measurement: 7.22 Minimal detectable change: 20.01
6 minutes walking test[32, 33]	1: Gait Aerobic Capacity	<15 minutes	Minimal detectable change: 20 meters Minimal clinically important difference: 50 meters Test-retest reliability: 0.95
1-Minute Sit-to-Stand[34]	1: Functional Mobility Strength	<3 minutes	Minimal clinically important difference: 3 repetitions ICC: 0.99 (95%CI 0.97-1.00)
Timed Up and Go test[32, 35, 36]	1: Gait and balance	<3 minutes	Minimal clinically important difference: 3.4 seconds Intra-rater ICC: 0.92 Inter-rater ICC: 0.91
Short Physical Performance Battery[33, 37]	3: Balance Gait Functional Mobility Strength	10 minutes	Minimal detectable change: 0.5 points Minimal clinically important difference: 1.0 points ICC: 0.92

Despite being so important for patients' quality of life, and in spite of the existence of the aforementioned tests and questionnaires, functioning measurement is particularly challenging, namely in the rehabilitation context, due to difficulties in capturing all its dimensions plus the time needed to evaluate different patient-reported outcome measures (PROMs) or to complete physical tests to evaluate it (Table 1).[38] In fact, all these different PROMs or tests measure specific dimensions of functioning, and it isn't possible to obtain a complete evaluation with a single test (that is, there is no clear gold-standard to fully estimate patient's functioning).[39] Also, since some of those outcomes are dependent on patient's experiences and expectations, there may be some degree of subjectivity, which can lead to some variation.[40] For example, the physical environment, social environment, social support and patients' own personality all influence reported outcomes, since patients with better house structures (less architectural barriers), home support (living with other people willing to help) or state or other organizations' support may report different physical outcomes (performance) even if they have the same levels of capacity.[38] All those shortcomings have not only an impact in the clinical practice, but also in obtaining scientific evidence of functioning recovery, since different studies use different outcomes function-related parameters, which leads to challenges in making comparisons.[41, 42]

This prompts the need to identify methods both (i) capable of effectively assessing functional outcomes and (ii) easily implementable in all rehabilitation areas, and specifically in post-ICU patients. Some of the possibilities that have been gaining some attention include the measurement of knee extension muscle strength and quadriceps femoris muscle mass with HHD and POCUS respectively.[43, 44] The rationale for their potential use in this setting lies on the fact that patients not only recover functioning and capacity after discharge, but also muscle mass and strength, being hypothesised that there is a correlation in the evolution of these properties. However, such association between the evolution of these variables has not yet been assessed, despite that being key to understand if HHD and POCUS can help assessing evolution of functioning and capacity.[45-47] There are additional aspects which have not been assessed on the performance of HHD and POCUS and which are particularly relevant for their eventual implementation in routine clinical practice – for example, their reliability has not yet been fully evaluated.

Addressing these problems may help overcome a major challenge in everyday clinical practice, as it is challenging to assess functioning and performance status improvement during rehabilitation programs, assessing when the patients should adjust rehabilitation treatment or even if they are still benefiting from them.

1.2 Muscle strength measurement with handheld dynamometer

Muscle strength is important for functional tasks and there seems to be a negative correlation between muscle strength and (i) disability (ii) and capacity.[6, 43, 48] One of the joint movements that has been most commonly assessed in its association with capacity is knee extension, as it is a pivotal movement for functional activities such as walking, sitting, dressing or having a shower.[49-51] The gold-standard device for evaluating muscle strength in knee extension is the isokinetic dynamometer (IKD), but it is particularly difficult to apply in the clinical practice as it is expensive, non-portable and requires previous specialized training.[52] A possible alternative to measure muscle strength is the handheld dynamometer (HDD) which is a less expensive and portable device that requires less training, rendering it potentially more applicable in clinical practice.[52]

In fact, some studies already showed the HHD to display high validity and reliability in measuring muscle strength and its measurements have been shown to have strong correlation with those obtained using an IKD. [52-60] (table 2) This demonstrates that HHD can be possibly seen as an alternative to IKD and also that it can help to overcome the limitations of the methods currently available to effectively estimate the functioning in an objective way. Table 2 summarises some of the previous studies evaluating the use of HHD to measure muscle strength.

Although there are some studies assessing the accuracy and reliability of HHD some questions still remain unanswered as these studies have some limitations. In fact, only a few of the studies calculated sample size, used more than two observers or compared experienced with unexperienced observers. In addition, to the best of our knowledge, no previous studies have been conducted in a rehabilitation setting.[53-59]

Table 2. Validity and Reliability of instruments used to measure muscle strength

Instrument	Comparison	Population	Outcome
- Belt-stabilized HHD[53]	- IKD	26 healthy adults	Strength: ICC: KE 0.91(R), 0.93(L); KF: 0.66(R), 0.62 (L) r: KE 0.78(R), 0.87(L); KF: 0.74(R), 0.69(L)
- 2 different HHD[54]	- IKD	30 healthy young adults	Peak force: ICC: HHD KE 0.89-0.92, IKD KE 0.98; HHD KF 0.89-0.96, IKD KF 0.94 r: KE 0.82-0.90; KF 0.64-0.73 Rate of force development: ICC: HHD KE 0.71-0.84, IKD KE 0.98; HHD KF 0.78-0.91, IKD KF 0.93 r: KE 0.36-072; KF 0.18-0.58
- Portable dynamometer anchoring system[55]	- IKD	39 healthy adults (20-40 y)	Maximal isometric knee extensor strength: ICC: 0.98 maximal voluntary knee extensor torque: r: 0.927
-HHD[61]	- IKD	11 active and healthy participants	ICC HHD for KF and KE: >0.938 Correlation between HHD e IKD: KF: r= 1.000 KE: r= 0.996
- HHD[57]	- IKD	- 216 professional football players	ICC for HHD isometric KE: 0.96 ICC for HHD isometric KF: 0.91 Total correlations: r= 0.322-0.617 KF and KE measured by HHD and IKD 60°/s r= 0.519-0.617 KF and KE measured by HHD and IKD 300°/s r= 0.322-0.472
- HHD with belt resistance[62]	- HHD with examiner resistance	- 30 patients with total knee arthroplasty	ICC HHD with belt resistance: 0.82 ICC HHD with therapist resistance: 0.80
- HHD - Portable fixed dynamometer[58]	- IKD	- 16 young adults	ICC: HHD KE 0.76

			Portable fixed dynamometer KE 0.92 IKD KE 0.93 HHD KF 0.49 Portable fixed dynamometer KF 0.96 IKD KF 0.89
- HHD KE and KF[63]	- 5-Repetition Sit to stand test time scores - Berg Balance Scale - limits of stability test using dynamic posturography	- 12 subjects with chronic stroke	- ICC: 0.970–0.999 - Spearman Correlation Coefficient Between 5-Repetition Sit to Stand Test and HHD KF: (p= –0.753 to –0.830) - No significant associations were found between 5-repetition Sit to stand test and HHD KE, Berg Balance Scale and limits of stability tests in subjects with stroke
- HHD stabilized by a belt [59]	- HHD stabilized by an examiner	- 24 young and healthy participants	ICC for both HHD: 0.93-0.97
- belt-stabilized HHD[64]	- dynamometer chair	52 healthy adult subjects (36.5y) and 21 haemodialysis patients (72.4y)	Spearman coefficients Healthy adult subjects: 0.63 Haemodialysis patients: 0.75
- HHD[65]	- IKD	- 25 young male	Associations between HHD and IKD: KE male examiner: r= 0.406 KE female examiner: r= -0.086 KF male examiner: r= 0.664 KF female examiner: r= 0.214
- HHD KE [66]	- minimum sit-to-stand height test	152 patients (72.5y)	low correlation between HHD KE and lowest chair height: r= 0.30-0.42
- Fixed HHD KE and KF[67]	Functional performance: - 30-second lateral hop for endurance - triple hop for distance - crossover hop for distance - single hop for distance - single-legged vertical jump	62 recreationally athletic subjects (21y)	Correlation with p<0.05 (r=0.26-0.49) - KE for Triple hop for work - KE and KF for Single-legged vertical jump for work - KF for Triple hop for distance and for work - KE and KF for Single-legged vertical jump for distance and for work
- HHD[68]	- Isometric IKD - Static balance - Maximal step length	- 18 adults (65-92y)	- Isometric IKD: r= 0.58-0.75 - Static balance: r= 0.09-0.14

			- Maximal step length: r 0.37-0.62
- HHD without belt-stabilization[69]	- HHD with belt-stabilization	- 20 testers	KE forces measured using the HHD without belt-stabilization (470.6N) were significantly lower than those measured with belt-stabilization (866.9N). Pearson correlations between tester characteristics (weight, gender, BMI, push force, grip force, etc.) and knee extension forces measured with no belt-stabilization were all statistically significant ($p \leq 0.01$); however, the correlations were not statistically significant under the belt-stabilization condition.
- HHD KE[70]	- physical performance tests - total leg extension muscle strength	- 47 older adults (>70y)	- HHD KE muscle strength correlation to 10 m gait speed and timed up-and-go test: r=-0.41 to -0.45 - balance tests were poorly correlated with muscle strength: r= 0.17-0.40.
- HHD maximal quadriceps isometric strength[71]	- estimated metabolic equivalents (eMETs) exercise capacity levels	- 621 patients (60.6y) with recent coronary artery bypass grafting or myocardial infarction	- HHD maximal quadriceps isometric strength: r= 0.42 - HHD maximal quadriceps isometric strength % body weight: r= 0.43 - positive predictive values of 0.72, 0.66, and 0.67 for 5, 7 and 10 eMETs
- HHD[72]	- Manual Muscle Testing	- 20 children with spina bifida(10y)	- Manual Muscle Testing ICC KE: 0.40 - HHD ICC KE: 0.83
- HHD KE[73]	- Twelve-step Ascend and Descend Test	- 35 subjects with chronic stroke - 29 healthy elderly subjects.	- HHD KE: r= - 0.24 to - 0.26
- HHD KE % body weight[74]	- Sit-to-stand independence	- 31 adults	r= 0.711

HHD – Handheld dynamometer, ICC – Intraclass correlation, IKD, Isokinetic dynamometer, KE – Knee extension, KF – Knee flexion, L – Left, R – Right, y - years

1.3 *Muscle thickness measurement with point-of-care ultrasound*

Muscle thickness is another biomarker which has been hypothesised to be potentially associated with functioning. Particularly in the case of knee extension – which, as previously discussed, is particularly important for functional tasks - muscle thickness appears to have a correlation with muscle strength. In the clinical practice and for investigational purposes, muscle mass and thickness are usually measured using Dual-energy X-ray absorptiometry and magnetic resonance imaging.(75-77) Nevertheless, these techniques are also not easily applicable, which makes them difficult to be applied in the clinical practice. In addition, magnetic resonance imaging is an expensive procedure. The POCUS is a possible alternative, being more affordable and easier to apply in an outpatient setting, rendering it a potentially useful tool to measure muscle thickness.(78) Some studies have already shown that POCUS has a high acuity and reliability to measure muscle thickness of different muscles, namely the quadriceps femoris.(12, 75, 79, 80) (Table 3) Therefore, measuring knee extensors thickness is conceivable as a simple, easy to obtain and reliable metric, which may aid to adequately modulate treatment choices in patients with disability.

Table 3: Validity and Reliability of instruments used to measure muscle thickness

Instrument	Observers	Population	Outcome
US(81)	1 observer	13 healthy Caucasian volunteers	Mid-tight intra-rater ICC = 1.00 (95%CI =0.99-1.00)

US[82]	2 trained physicians	29 critically hill patients	Intra-observer reliability's ICC: QF: 0.74 [95% CI 0.63; 0.84] and 0.83 [95% CI 0.75; 0.9] in different locations Inter-observer reliability's ICC: 0.76 [95% CI, 0.66; 0.86] and 0.81 [95% CI, 0.7; 0.9] in different locations
US[83]	2 observers	26 patients in hemodialysis	Intra-observer ICC QF: ranging between 0.98-1.00 Inter-observer ICC QF: ranging between 0.98-0.99
Portable US[84]	2 observers (1 expert and 1 novice)	10 patients from Emergency Department	Intra-observer ICC QF: Expert - 0.952 (95%CI 0.821 - 0.988) Novice - 0.999 (95%CI 0.998 - 1.000) Inter-observer ICC QF: 0.990 (95%CI 0.970 - 0.997)
Portable US[85]	4 novice observers 1 expert observer	12 patients (ICU, recovery group and healthy group)	Inter-observer ICC: QF - 0.940 (95%CI 0.903 – 0.966) RF – 0.915 (95%CI 0.863 – 0.952)
US[86]	2 critical care fellows	20 patients with sepsis	Intra-observer ICC QF: observer 1 - 0.925 (95%CI: 0.851–0.967 observer 2 - 0.835 (95%CI: 0.689–0.925) Inter-observer ICC QF: 0.992 (0.979–0.997)

95%CI – 95% confidence interval, ICC – Intraclass correlation, ICU – Intensive Care Unit, QF - Quadriceps Femoris, US – Ultrasound

However, some questions are still unanswered, namely regarding the reliability of this method, since the majority of the studies done with either ultrasound (US) or POCUS included only two observers, did not calculate sample size, did not assess users with different levels of experience, did not compare the performance of having two POCUS measurements *versus* one, and only assessed patients outside of the rehabilitation setting.[82-86]

2. Aims and Hypotheses

The main aim of this project consists in assessing the association between changes over time in knee extension muscle strength (measured using HHD) and thickness (measured using POCUS) and function-related changes in post-ICU patients, while at the same time contributing towards a better understanding of the properties of muscle strength and thickness measurements using HHD and POCUS.

In accordance with this general aim, three specific aims were defined, based on which three studies were conducted:

i. Measuring the intra-observer and inter-observer reliability of HHD for the assessment of knee extension strength and analysing (i) if the reliability of HHD measurements differ according to the experience of observers or the number of measurements, and (ii) if there is an association between knee extension strength measured by HHD and functioning;

ii. Assessing the intra-rater and inter-rater reliability of POCUS for assessing muscle thickness in a rehabilitation patient population and analysing (i) if the reliability of POCUS differs according to the observers' experience or number of measurements, and (ii) if there is an association between muscle thickness measured by POCUS and functioning.

iii. Assessing – in an exploratory way - the association between changes in knee extensor muscle thickness and strength (measured using POCUS and HHD) and the corresponding changes in functioning changes in ICU patients, as measured using the SF-36, the Barthel Index, handgrip strength, 6MWT, 1MSTS, TUG test and SPPB.

With regard to these specific aims, we postulated the following hypothesis:

i. Intra-observer and inter-observer reliability of both HHD and POCUS measurements is expected to be very high according to preliminary studies with lower number of observers;

ii. A higher reliability of HHD and POCUS measurements is expected to be achieved (i) if these tools are used by previously experienced observers (rather than inexperienced observers), and (ii) if obtained based on two measurements rather than based on one measurement alone.

iii. Both knee extensor muscle thickness and strength are expected to be moderately correlated with changes in functioning, functional capacity and performance in ICU patients as measured with SF-36, the Barthel Index, handgrip strength, 6MWT, 1MSTS, TUG test and SPPB.

iv. Knee extensor muscle thickness and strength, as well as function-related parameters are expected to improve faster in the first months of recovery after ICU discharge.

3. PhD Studies: Methods and Results

This thesis consists in three different studies planned to collectively answer the main aim of this project: (i) better understand the properties of muscle strength and thickness measurements using HHD and POCUS and (ii) use these methods to assess the association between changes over time in knee extension muscle strength and quadriceps femoris thickness and function-related changes in post-ICU patients.

To reach this aim, we started by assessing intra-observer and inter-observer reliability of the HHD to assess knee extension muscle strength and of the POCUS to assess quadriceps femoris muscle thickness. Subsequently, we performed a study to evaluate if changes in muscle strength and thickness as assessed by those methods would be associated with functioning progression over time. (Figure 1)

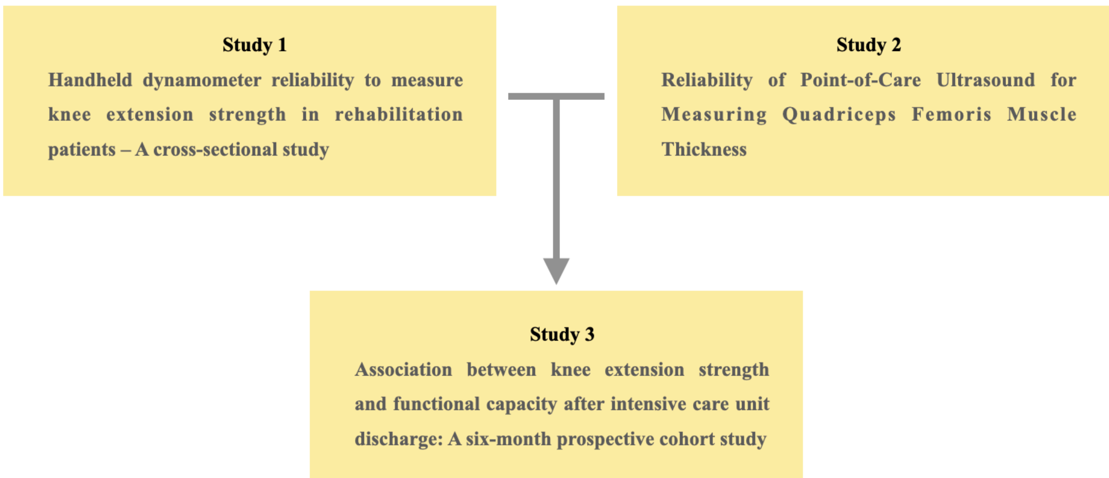


Figure 1 – Diagram of the rationale of the studies performed

3.1 Handheld dynamometer reliability to measure knee extension strength in rehabilitation patients - A cross-sectional study

Reference of the published version of this study: Pinto-Ramos J, Moreira T, Costa F, Tavares H, Cabral J, Costa-Santos C, Barroso J, Sousa-Pinto B. Handheld dynamometer reliability to measure knee extension strength in rehabilitation patients - A cross-sectional study. PLoS One. 2022;17(5):e0268254.

3.1.1 Methods

Study Design

Twenty-nine patients were consecutively selected for this cross-sectional study at the Inpatient unit, PMR Department, Centro Hospitalar e Universitário de São João, which is a tertiary care hospital in Northern Portugal. For each patient, four different observers, two previously experienced and two inexperienced in HHD use, made four measurements of knee extension peak force with an HHD. Do to schedule logistics and in order to systematize the measurements, two measurements were taken in early afternoon (measurement A and measurement B – assessment AB), and two three hours after (measurement C and measurement D – assessment CD) (Fig 2). All observers were physicians with at least two years of clinical experience. Intra-rater and inter-rater Intraclass Correlation Coefficients (ICC), absolute differences between measurements within each assessment, and correlation with functional variables were calculated. The study was approved by Ethical Committee of the respective hospital (Research Project 289/20). Patients were asked for written consent. This article was written according to STROBE statement guidelines.[87]

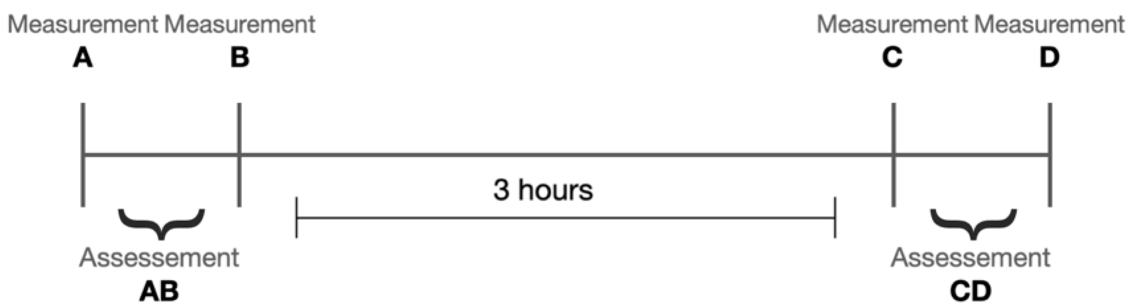


Figure 2 – Schematic representation of the measurements performed by each observer in each participant

Setting and Participants

We assessed participants which had already been included in a previous study from our research group, with sample size of thirty patients being calculated. In brief we consecutively included all patients admitted at our Service from March to June 2021 which satisfied defined inclusion and exclusion criteria. Patients aged ≥ 18 years-old performing rehabilitation treatment for any cause fulfilled the inclusion criteria. Exclusion criteria were: patients unable to ambulate without assistance before hospitalization, morbid obesity, actual lower limb bone fractures, serious pressure or venous ulcers, cardiorespiratory instability, major psychiatric disorders, severe cognitive impairment or neurologic conditions such as multiple sclerosis, traumatic brain injury or severe stroke. Patients admitted to our acute rehabilitation service were hospitalized due to acute events which have impacted their functional capacity, having been transferred for personalized and intense rehabilitation programs. Each day, patients had a 24-hour rehabilitation nursery care with activities of daily living training sessions, two physical therapy sessions of one and half hours each, one occupational therapy session and, in case of need, psychology and speech therapy.

Variables and Measurements

From each patient, knee extension strength was measured four times by each of the four observers with the HDD. That is, the four observers performed two knee extension peak force measurements with HHD (Micro FET®2 HHD; Hoggan Health Industries, Draper, UT, USA) at each patient (measurements A and B, comprising the assessment AB) followed by two other measurements three hours later (measurements C and D, comprising the assessment CD) (Fig 2). The three-hour interval was defined to allow some interval between assessments in order to reduce the risk of memory bias while allowing all assessments to be done in the same day (ensuring clinical stability of the patients during the interval between assessments). Patients were seated in the edge of the stretcher with the feet and hands suspended and knee flexed at 60º. The observer was in squatting position with the back against the wall for support and both arms extended to the patient dominant leg for stabilization, so that the results dependence on patient or observer strength would be minimized. The HHD was placed in the anterior leg of each patient, five centimetres above the distal part of the medial malleolus (Fig 3). Patients were asked to make and maintain maximum knee extension strength for five seconds. One-minute rest was used between measurements to decrease fatigue impact. Results displayed on HHD were registered by an independent observer, so that the patients and observers applying the HHD were blinded. Strength was measured in Newtons (N).

Figure 3 – Standardized evaluation method of the patients. Observer was squatting with back against the wall and stretch arms in order to increase stability.

In order to standardize the procedures, every observer was given a thirty-minute training before the beginning of the study on how to perform HHD measurements correctly, even though two of the observers had more than two years of experience using the HHD.

From each patient, information was also collected (by a different researcher) on his/her sex and age as well as height, weight, body mass index, diagnosis, pain on visual analogue scale and current and previous modified Rankin Scale. Functional variables were also collected, namely handgrip strength, TUG test, 1MSTS test and Medical Research Council Sum Score.(21, 22, 88, 89)

Study size

The sample size for this study was calculated simultaneously with that for the previous study. Our primary effect size measure was the intra and inter-rater ICC. For us to estimate the sample size, we conducted a comprehensive literature search in order to identify previous ICC estimates. We identified seven studies providing such estimates,(53-59) which we pooled by random-effects meta-analysis, resulting in a pooled ICC of 0.944 (95% confidence interval (95%CI) = [0.902;0.986]) for intra-rater reliability and 0.977 (95%CI=[0.959;0.995]) for inter-rater reliability. If more than one ICC were calculated in the same study (i.e., in left and right limbs), the lower ICC value was used in the meta-analysis. The minimum sample size estimated based on those meta-analytic values using a 95%CI and a semi-width of 5% was of 13 for intra-rater and 4 for inter-rater reliability. Since the required sample size calculated for the previous study was higher (30 patients), we ended up enrolling a larger number of patients.

Statistical analyses

For numerical (continuous) variables, we used means and standard-deviations (SD) for describing variables with normal distribution and medians and 25^{th} and 75^{th} percentiles (P25;P75) for variables with non-normal distribution. Categorical variables were described using absolute and relative frequencies (in percentage). Reliability of HHD measurements was estimated using ICC. For each set of two measurements (i.e., assessment AB, and assessment CD), we registered the average, maximum and first value. Two-way mixed-effects model ICCs were used to calculate intra-rater and inter-rater reliability. ICCs were calculated using single measurements per assessment set for maximum and first values. For intra-rater ICC, we calculated it taking in to account the measurements of all observers, comparing average, maximum and first values of the AB assessments with those of CD assessments. Individual intra-rater ICC were also presented for average, maximum and first values. For inter-rater ICC, we compared average, maximum and first values of AB assessments between observers (in an ancillary analysis, such comparisons were also performed for CD assessments). Experienced and inexperienced observers were also separately analysed for ICC comparison between different levels of training. Reliability was considered poor if ICC<0.50, moderate if 0.50≤ICC<0.75, good if 0.75≤ICC≤0.90 and excellent if ICC>0.90.[90]

Since ICC values are dependent on the heterogeneity of the population,[90] absolute differences between average, maximum and first measurements of each assessment were calculated, with differences being tested using the Friedman test. We also calculated relative differences, dividing the aforementioned absolute differences by average knee extension peak force of each participant. The correlation between absolute differences and average knee extension peak force was also calculated to assess if absolute differences between assessments depend on average knee extension peak force. If correlations were found to be sufficiently strong (r>0.4 or r<-0.4), univariable linear regression models were applied.

We also estimated the correlation between average knee extension peak force of the different participants with other functional variables. Such functional tests/variables include the handgrip strength, TUG test, 1MSTS test, Medical Research Council Sum Score, actual and previous Modified Rankin, and POCUS measured rectus and quadriceps femoris muscle thickness. If correlations were found to be sufficiently strong (r>0.4 or r<-0.4), univariable linear regression models were applied, with knee extension peak force being the independent variable and the dependent variables corresponding to the results of each functional variable.

Correlations were estimated using Spearman correlation coefficients. A p-value inferior to 0.05 was considered statistically significant. The Bonferroni correction was applied in order to control for multiple comparisons. ICC values were calculated using R software and the remaining statistical analyses were done in IBM SPSS Statistics 28 (IBM, Armonk, NY).

3.1.2 Results

Twenty-nine patients were included in this study. Participants' age ranged between 19 and 82 years old with mean age of 58.8 (SD=14.1) years old; 75.9% of patients were male (Table 4). Participants displayed a median body mass index of 23.1 kg/m^2 (P25;P75=21.5;27.4), and the current modified Rankin median score was of 3 (P25;P75=2;4). Nineteen patients (65.5%) were not able to stand up from a chair whereas thirteen patients could not ambulate (44.8%). Patients able to ambulate had a mean TUG test of 18.9 seconds (SD=11.6) and patients able to stand from a sitting position had a mean 1MSTS test of 17.2 (SD=7.6). Median muscle strength was of 189.8 N (P25;P75=130.7;274.4).

Table 4: Demographic characteristics of the sample and descriptions of functional tests

Variables	Patients (N=29)
Age (years) – mean (SD)	58.8 (14.1)
Males – *n* (%)	22 (75.9)
BMI (Kg/m^2) – median (P25;P75)	23.1 (21.5;27.4)
Previous mRankin – median (P25;P75)	0 (0;1)
Current mRankin – median (P25;P75)	3 (2;4)
Medical Research Council Sum Score - median (P25;P75)	46 (43;47.5)
Handgrip Strength (kg) – median (P25;P75)	18.3 (11.7;23.3)
Timed Up and Go (s) – mean (SD)[1]	18.9 (11.6)
Sit to Stand test – mean (SD)[2]	17.2 (7.6)

BMI – Body Mass Index, P25;P75 – 25th and 75th Percentile, Kg – Kilogram, m – Meter, s – Seconds, SD – Standard Deviation; 1 – Only 16 participants could complete the test, 2 - Only 10 participants could complete the test

The most common admission diagnoses were Intensive Care Unit Acquired Weakness (ten patients) and non-severe (Initial National Institute of Health Stroke Scale (NIHSS) <15) Stroke (six patients).

Observers rapidly adapted to the use of HHD and measurements took less than one minute in the majority of times, although positioning the most disabled patients in the stretcher was often a challenge. Also, patients with higher knee extension strength were difficult to resist in order to measure their full strength.

Intra-Rater Reliability

The intra-rater ICC for knee extension strength of the 464 measurements (corresponding to 4 measurements done by each of the 4 observers to the 29 patients) was excellent with results higher than 0.950 when considering either average, maximum and first measurements. The intra-rater ICC was significantly higher when calculated based on the average of two measurements (average of AB versus average of CD) (ICC=0.978, 95%CI=0.969-0.985) than when based on the first measurement alone of each set (A and C) (ICC=0.950, 95%CI=0.928-0.965) (Table 5).

Individual intra-rater ICC were similar or higher than 0.948 for the four observers, either using mean, maximum or first values. (Table 6)

Table 5: Intra-rater and Inter-rater intra-class correlation coefficients (ICC) of knee extension strength measured with hand-held dynamometers using Average, Maximum and First values within AB and CD assessments.

	ICC calculated based on average of measurements within each assessment (95%CI)	ICC calculated based on maximum of measurements within each assessment (95%CI)	ICC calculated based on first measurements within each assessment (95%CI)
Knee Extension Strength intra-rater ICC	0.978 (0.969-0.985)	0.961 (0.945-0.973)	0.950 (0.928-0.965)
Knee Extension Strength inter-rater ICC			
Assessment AB	0.932 (0.864-0.967)	0.936 (0.874-0.969)	0.927 (0.859-0.964)
Assessment CD	0.952 (0.908-0.976)	0.952 (0.911-0.976)	0.943 (0.899-0.971)

95%CI – 95% Confidence Interval

Table 6: Individual intra-rater intra-class correlation coefficients (ICC) of knee extension strength measured with hand-held dynamometers using Average, Maximum and First values within AB and CD assessments for all the observers

	ICC calculated based on average of measurements within each assessment (95%CI)	ICC calculated based on maximum of measurements within each assessment (95%CI)	ICC calculated based on first measurements within each assessment (95%CI)
Knee Extension Strength intra-rater ICC			
Experienced Observer 1	0.980 (0.953-0.991)	0.952 (0.895-0.978)	0.949 (0.887-0.977)
Experienced Observer 2	0.982 (0.963-0.992)	0.972 (0.943-0.987)	0.950 (0.897-0.976)
Inexperienced Observer 1	0.973 (0.943-0.987)	0.957 (0.911-0.980)	0.955 (0.906-0.978)
Inexperienced Observer 2	0.980 (0.957-0.991)	0.965 (0.927-0.983)	0.948 (0.892-0.975)

95%CI – 95% Confidence Interval

Inter-Rater Reliability

Considering measurements of the assessment AB, we observed an inter-rater ICC between 0.927 and 0.936 of knee extension strength measured with HHD. There were no substantial differences between ICC calculated based on average, maximum and first values of the assessment AB. (Table 5). Similar results were found from assessment CD with ICC between 0.943 and 0.952. (Table 5)

Absolute and Relative Differences

The median of the absolute differences between knee extension strength assessments (AB vs CD) ranged between 15.0 and 15.4 N (depending on whether average, maximum or first measurements in each assessment were being considered). No significant differences were found between median differences calculated based on average, maximum or first measurements within each assessment (p=0.23). Relative differences between assessments ranged between 8.7% and 10.5%. (Table 7)

Table 7: Absolute difference (in Newton) of knee extension strength and percentual difference of absolute difference over average muscle strength of all assessments with Hand Held Dynamometer (HHD) using mean, maximum and first measurements between assessments AB and CD for the same observer.

	Average of measurements within each assessment - Median (P25;P75)	**Maximum of measurements within each assessment - Median (P25;P75)**	**First measurement within each assessment - Median (P25;P75)**
Absolute Difference (N) Between Knee Extension Strength	15.2 (7.9;34.7)	15.0 (6.2;33.0)	15.4 (8.9;38.0)
Percentual Difference Between Absolute Difference and Average Knee Extension Strength	8.7% (5.1;16.8)	8.8% (4.5;17.4)	10.5% (4.9;18.3)

HHD – Hand Held Dynamometer, P25;P75 – 25th and 75th Percentiles, N – Newton;

Average HHD strength were moderately correlated with absolute differences for average (*r*=0.477 (95%CI=0.318;0.610) [p<0.001]), maximum (*r*=0.415 (95%CI=0.246;0.559) [p<0.001]) and first measurements (*r*=0.414 (95%CI=0.245;0.558) [p<0.001]). Univariable linear regression coefficients were subsequently applied, with coefficients being of 0.112 (95%CI=0.080;0.144) [p<0.001] for average measurements, 0.103 (CI95%=0.070;0.136) [p<0.001] for maximum measurements, and 0.117 (95%CI=0.080;0.155) [p<0.001] for first measurements. On the other hand, correlations between relative differences and average HHD strength were not significant, with Spearman correlation coefficients ranging between -0.063 and -0.109.

Experienced and Inexperienced Observers

Similar results were observed when comparing the reliability of measurements of experienced versus inexperienced observers. For experienced observers intra-rater ICCs were of 0.981, 0.963, 0.949, depending on whether such ICCs were estimated based on average, maximum or first measurements, respectively. For inexperienced observers, such values were of 0.976, 0.961, 0.950, respectively.

Muscle Strength and Functional Outcomes

Except for the Previous Modified Rankin, muscle strength measured by knee extension peak force in the HHD was moderately correlated (correlation coefficient ≥0.4) with all functional variables, including the current modified Rankin, the TUG test, the 1MSTS test, the Medical Research Council Sum Score, the Handgrip Strength and the Quadriceps and Rectus Femoris Muscle Thickness. (Table 8) Correlations after Bonferroni correction were statistically significant for all functional tests, except the Previous Modified Rankin, the Medical Research Council Sum Score and the TUG. Linear regression coefficients ranged from -0.062 (association between knee extension peak force and TUG) to 0.045 (association between knee extension peak force and 1MSTS). (Table 8)

Table 8: Correlation and Linear Regression Coefficient between Handheld Dynamometer and different Functional Variables

Functional Variables	Correlation coefficient (95%CI) [p-value]	Linear regression coefficient (95%CI) [p value]
Previous Modified Rankin	-0.378 (-0.660; -0.002) [0.043]*	**
Current Modified Rankin	-0.565 (-0.776; -0.239) [0.001]	-0,005 (-0,008; -0.002) [0.004]
TUG test	-0.612 (-0.854; -0.151) [0.012]*	-0,062 (-0.106; -0.019) [0.009]
1MSTS test	0.499 (0.151; 0.737) [0.006]	0.045 (0.018; 0.073) [0.002]
Medical Research Council Sum Score	0.484 (0.132; 0.728) [0.008]*	0.019 (-0.005; 0.043) [0.109]
Handgrip Strength	0.545 (0.213; 0.765) [0.002]	0.030 (0.004; 0.055) [0.023]
Quadriceps Femoris Muscle Thickness	0.511 (0.167; 0.744) [0.005]	0.002 (0.000; 0.003) [0.012]
Rectus Femoris Muscle Thickness	0.536 (0.201; 0.759) [0.003]	0.002 (0.000; 0.005) [0.027]

95%CI – 95% Confidence Interval, TUG – Timed Up and Go; 1MSTS – 1-minute sit-to-stand; *non-significant after Bonferroni correction; ** Not applicable since r<0.40

3.2 Reliability of Point-of-Care Ultrasound for Measuring Quadriceps Femoris Muscle Thickness

Reference of the published version of this study: Pinto-Ramos J, Costa-Santos C, Costa F, Tavares H, Cabral J, Moreira T, Brito R, Barroso J, Sousa-Pinto B. Reliability of Point-of-Care Ultrasound for Measuring Quadriceps Femoris Muscle Thickness. European Journal of Physical and Rehabilitation Medicine. 2022.

3.2.1 Methods

Study Design

This single-centre cross-sectional study was conducted at the PMR Department of *Centro Hospitalar e Universitário de São João*, a tertiary care hospital in Northern Portugal. Twenty-nine patients were assessed by four different observers using a portable ultrasound to evaluate muscle thickness of rectus femoris and quadriceps femoris. Each observer performed four measurements on each patient, two in the early afternoon (assessments A and B) and two in the late afternoon (3 hours after the first measurements; assessments C and D) (Figure 4). Intra and Inter ICC were calculated, as well as absolute differences between assessment AB and CD values. The study was approved by the institution's Ethical committee and written consent was obtained from each patient. This article was written according to the STROBE statement for cross-sectional studies.[91]

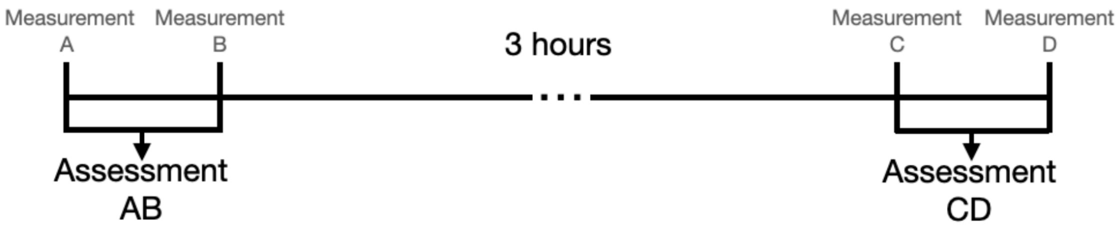


Figure 4 – Schematic representation of the measurements performed by each observer in each participant

Setting and Participants

Patients admitted to the PMR Department of Centro *Hospitalar e Universitário de São João* from March to June 2021 were consecutively included in the study if they fulfilled all the eligibility criteria. To fulfil the inclusion criteria, patients ought to be aged ≥ 18 years old and to have performed rehabilitation treatment for any cause. We excluded patients who were unable to walk independently before starting rehabilitation treatment, as well as patients with severe cognitive impairment, neurological/neuromuscular conditions (such as multiple sclerosis, traumatic brain injury or amyotrophic lateral sclerosis), major psychiatric disorders, traumatic bone fractures of the lower limb, morbid obesity (body mass index ≥ 35), serious pressure or venous ulcers, or cardiorespiratory instability.

Variables and measurements

For all participants we assessed the current and previous to admission modified Rankin (mRankin) scale, the current handgrip strength, and current pain intensity (by means of visual analogue scale). Handgrip strength was measured using Camry® Electronic Hand Dynamometer (Camry Scale, USA).

Patients were asked to sit with their elbows flexed at 90º and to produce and maintain maximum handgrip strength for five seconds with their dominant hand. The maximum values of each pair of attempts were recorded. In addition, we performed TUG and 1MSTS tests, Medical Research Council Sum Score and Knee Extension Peak force measured with a Hand Held Dynamometer. Participants' demographic characteristics (sex and age), height, weight, body mass index and diagnosis at admission were also retrieved.

For each patient four different observers performed two measurements of rectus femoris and quadriceps femoris thickness in the early afternoon (assessments A and B) and two measurements of the same muscles' thickness 3 hours after (assessments C and D) (Figure 4). Measurements were performed using a handheld portable POCUS (Butterfly iQ+, Butterfly Network, Inc., Guilford, CT) configured for musculoskeletal image acquisition. (Figure 5) The probe was applied 15 cm above the superior pole of the patella of the dominant leg while the patients were in supine position with both legs stretched. Minimal hand force was applied at the ultrasound probe in order to avoid the effect of pressure on muscle thickness.[92] The probe was positioned perpendicular to muscle fibres in order to avoid increases in muscle thickness due to angulation.[93] Maximal distances between the cortex of femur and the most superficial muscular fascia were measured.

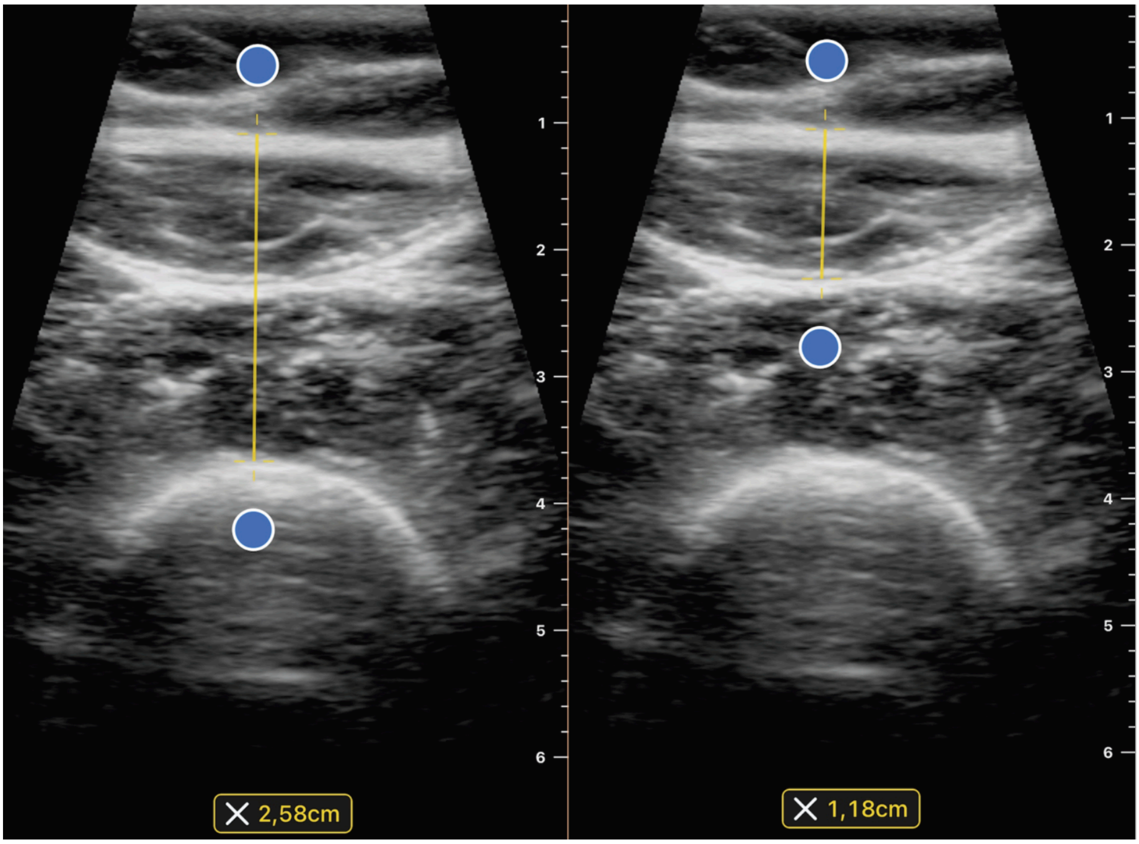


Figure 5 – Ultrasound measurements of quadriceps femoris (left) and rectus femoris (right); Yellow line represent the measurement of muscle thickness.

Two of the four observers (observer 1 and 2) had previously received formal training and had already been familiar with musculoskeletal ultrasound imaging for at least two years; nonetheless one hour of training was provided for all observers. The four observers were blinded for each other's results, and a fifth researcher statistically analysed the data.

Study size

This study primary effect size measure consisted of the ICC for intra-rater and inter-rater reliability. In order to estimate the adequate sample size, we performed a comprehensive literature search, followed by random-effects meta-analysis of ICC values published in identified studies.[82-86] The estimated meta-analytical intra-rater ICC value for quadriceps femoris muscle thickness POCUS measurements was 0.91 (95% confidence interval [CI]=0.82;1.00), while the meta-analytical inter-rater ICC was of 0.95 (95%CI=0.89;1.00). Using the meta-analytical value, and in order to obtain 95%CI with a semi-width of 5%, with a power of 80%, and considering four observers we estimated a minimum sample size of thirty patients for intra-rater reliability ICC and ten for inter-rater reliability ICC. Thirty-five patients were selected although we were only able to assess twenty-nine patients, due to loss of follow-up.

Statistical analyses

Means and SD were used to describe variables with normal distribution, and medians and 25th and 75th percentile (P25-P75) were used to describe asymmetrical variables.

For each set of two assessments (i.e., assessments AB, and assessments CD), we registered the average, maximum, minimum and first value. Two-way model intraclass correlation coefficients were then calculated in order to estimate the reliability between average, maximum, minimum and first AB versus CD assessments from the same observer (intra-rater ICC). In addition, ICC were calculated to estimate the reliability among average, maximum, minimum and first AB measurements of the four different observers (inter-rater ICC).[90] A subanalysis was performed with ICC values being separately estimated for experienced and inexperienced observers. The ICC was considered poor for values < 0.50, moderate for values between 0.50–0.75, good for values between 0.75–0.90, and excellent for values >0.90.[90]

We calculated absolute and relative percent differences (absolute differences divided by the average value of the four assessments from the four observers) between the average, maximum, minimum and first measurements of AB assessments versus those of CD assessments. Spearman correlation coefficients were calculated to assess the correlation between computed differences and average muscle thickness, so as to assess if increased thickness associated with higher absolute differences.

Finally, as a post hoc analysis, we computed the Spearman correlation coefficient between average muscle thickness of rectus femoris and quadriceps femoris from all observers and functional variables measured by a fifth independent researcher (namely current and previous mRankin scale, handgrip strength, Medical Research Council Sum Score, Timed Up and Go and Sit to Stand tests) and Knee Extension Peak force measured with an Hand Held Dynamometer.[94]

Statistical significance was considered at p-value <0.05. All analyses were performed using IBM SPSS Statistics 27 software and the R software.

3.2.2 Results

A total of 29 patients were included in the study, of whom 22 (75.9%) were male. (Table 9) Participants' mean age was of 58.8 years old (SD=14.1). The median mRankin scale was 0 (P25-P75 0-1) before hospitalization, and of 3 (P25-P75 2-4) at time of observation. The most common admission diagnosis was intensive care unit acquired weakness (ICUAW) (10 patients; 34.5%), followed by ischemic stroke (6 patients; 20.7%). At the time of assessment, thirteen (44.8%) patients were unable to walk, and 19 (65.5%) patients were not capable of standing up from a sitting position.

Table 9: Demographic characteristics of included patients

	Included participants (N=29)
Age (years) – mean (SD)	58.8 (14.1)
Males – *n* (%)	22 (75.9)
Weight (Kg) – mean (SD)	72.7 (13.4)
Hight (m) – mean (SD)	1.72 (0.10)
BMI (Kg/m^2) – median (P25-P75)	23.1 (21.5-27.4)
Previous mRankin – median (P25-P75)	0 (0 - 1)
Current mRankin – median (P25-P75)	3 (2 - 4)
MRCSS - median (P25-P75)	46 (43 – 47.5)
Handgrip Strength (kg) – median (P25-P75)	18.3 (11.7 – 23.3)
Timed Up and Go test (s) – mean (SD)[a]	18.9 (11.6)
1 minute Sit to Stand test – mean (SD)[b]	17.2 (7.6)

BMI – Body Mass Index, P25-P75 – 25th and 75th percentile, Kg – Kilogram, m – Meter, MRCSS - Medical Research Council Sum Score s – Seconds, SD – Standard Deviation; a – Only 16 participants could complete the test, b - Only 10 participants could complete the test

Intra-rater ICC

In all assessments, intra-rater ICCs for rectus femoris displayed values over 0.90 (Table 10). Intraclass correlation coefficients were significantly higher when calculated based on the average measurements of AB and CD (ICC=0.956; 95%CI=0.937-0.970) than when calculated based on maximum, minimum or first measurements. (Table 10)

Table 10: Intra-rater and Inter-rater ICC of rectus femoris and quadriceps femoris thickness measurements using Average, Maximum, Minimum and First values between AB and CD assessment and between AB measurements, respectively

	ICC calculated based on			
	Average of measurements (95%CI)	**Maximum of measurements (95%CI)**	**Minimum of measurements (95%CI)**	**First measurements (95%CI)**
Intra-rater:				
Rectus Femoris	0.956 (0.937-0.970)	0.908 (0.870-0.935)	0.911 (0.874-0.937)	0.901 (0.861-0.9319)
Quadriceps Femoris	0.966 (0.951-0.976)	0.932 (0.904-0.953)	0.927 (0.896-0.949)	0.922 (0.889-0.945)
Inter-rater:				
Rectus femoris	0.919 (0.863-0.957)	0.912 (0.853-0.953)	0.908 (0.847-0.951)	0.888 (0.815-0.940)
Quadriceps femoris	0.945 (0.907-0.971)	0.941 (0.900-0.969)	0.941 (0.900-0.969)	0.937 (0.892-0.966)

95%CI – 95% Confidence Interval, ICC - Intraclass Correlation Coefficient.

Similar results were observed for quadriceps femoris, with all ICC values being superior to 0.90 (Table 10). Intraclass correlation coefficients calculated based on the average values (ICC=0.966; 95%CI=0.951–0.976) were significantly higher than those estimated based on singular measurements (except for maximum values). All ICC values for quadriceps femoris were higher than those for rectus femoris measurements, but differences were not statistically significant. (Table 10)

Inter-rater ICC

Inter-rater ICC for recuts femoris assessments was higher when estimated based on average measurements (ICC=0.919; 95%CI=0.963–0.957) than single measurements, although differences were not statistically significant. (Table 10) Similar results were observed for ICC for quadriceps femoris, with estimates based on average measurements (ICC=0.945; 95%CI=0.907–0.971) having a non-significantly higher ICC value than when based on single measurements. All ICC values were found to be over 0.90. All quadriceps femoris ICC values were higher than those for rectus femoris, although such differences were not significant (Table 10). Similar results were observed when calculating ICC based on CD values. (Table 11)

Table 11: Inter-rater ICC of Assessment CD of rectus femoris and quadriceps femoris thickness using Mean, Maximum, Minimum and First values between CD measurements

	ICC calculated based on average of measurements (95%CI)	**ICC calculated based on maximum of measurements (95%CI)**	**ICC calculated based on minimum of measurements (95%CI)**	**ICC calculated based on first measurement (95%CI)**
Rectus femoris	0.888 (0.813-0.940)	0.880 (0.801-0.935)	0.876 (0.797-0.933)	0.863 (0.776-0.925)
Quadriceps femoris	0.951 (0.911-0.975)	0.949 (0.908-0.974)	0.939 (0.893-0.968)	0.935 (0.887-0.966)

95%CI – 95% Confidence Interval; ICC - Intraclass Correlation Coefficient

Muscle Thickness Absolute Difference

Median absolute differences in measurements of rectus femoris muscle thickness from the same observer ranged between 0.9 and 1.2 mm (corresponding to an average percent difference of 5.1-9.2%). (Tables 12-13) Differences in muscle thickness measurements were not correlated with average muscle thickness with all correlations being inferior to 0.037.

Median values of the absolute differences between the same observer muscle thickness measurements in different assessments were higher for the quadriceps femoris than for rectus femoris, with results ranging from 1.4 to 1.5 mm (corresponding to an average percent difference of 5.9-6.8%). (Tables 12-13) Differences in muscle thickness were weakly correlated with average muscle thickness with all correlations showing values inferior to 0.210.

Table 12: Absolute difference [in millimetres] of rectus femoris and quadriceps femoris muscle thickness mean, maximum, minimum and first measurement between assessment AB and CD for same observer

	Absolute difference [mm] between average of measurements (P25-P75)	**Absolute difference [mm] between maximum of measurements (P25-P75)**	**Absolute difference [mm] between minimum of measurements (P25-P75)**	**Absolute difference [mm] between first measurement (P25-P75)**
Rectus Femoris	1.0 (0.4 – 1.7)	0.9 (0.4 – 1.7)	1.1 (0.4 – 1.9)	1.2 (0.5 – 2.0)
Quadriceps Femoris	1.5 (0.6 – 2.3)	1.4 (0.5 – 2.7)	1.5 (0.7 – 2.6)	1.5 (0.7 – 2.7)

P25-P75 – 25th and 75th percentile, mm - millimetres.

Table 13: Percentual difference of absolute difference in millimetres of rectus femoris and quadriceps femoris muscle thickness average, maximum, minimum and first measurement over average rectus femoris and quadriceps femoris (respectively) muscle thickness between assessment AB and CD for same observer

	Percentual difference between average measurement and total thickness (P25-P75)	**Percentual difference between maximum measurement and total thickness (P25-P75)**	**Percentual difference between minimum measurement and total thickness (P25-P75)**	**Percentual difference between first measurement and total thickness (P25-P75)**
Rectus Femoris	7.1% (2.2 – 11.8)	5.1% (2.2 – 11.6)	6.9% (2.6 – 13.2)	9.2% (3.2 – 16.0)
Quadriceps Femoris	6.7% (3.7 – 9.9)	6.1% (3.4 – 10.3)	6.8% (3.5 – 8.6)	5.9% (3.8 – 10.3)

P25-P75 – 25th and 75th percentile.

Experienced vs Inexperienced Observers

Differences were not statistically significant between ICC values of observers already experienced for the use of POCUS and observers with no previous experience, either for rectus or quadriceps femoris thickness assessments.

Correlation between Muscle Thickness and Other Variables

We observed that rectus femoris thickness was significantly correlated with Knee Extension Peak Force measured with an HHD (r=0.536; p=0.003) and handgrip strength (r=0.574; p<0.001), but not for current and previous mRankin (r=-0.325; p= 0.086 and r=0.063; p= 0.746, respectively), Medical Research Council Sum Score (r=0.070; p=0.718), Timed up Go test on patients able to walk (r=-0.297; p=0.264), or Sit to Stand test on patients able to stand up from a sitting position (r=0.433; p=0.211).

We observed a moderate correlation between quadriceps femoris muscle thickness and current mRankin score (r=-0.376; p=0.044), Knee Extension Peak Force measured with an HHD (r=0.511; p=0.005), and handgrip strength (r=0.588; p<0.001). Correlations were weaker and not significant for other functional variables such as previous mRankin (r=0.141; p=0.466) and Medical Research Council Sum Score (r=0.064; p=0.742). Timed Up Go (n=16; r=-0.247; p=0.356) and Sit to Stand test (n=10; r=0.482; p=0.159) had a poor correlation for patients that could perform them.

3.3 Association between knee extension strength and functional capacity after intensive care unit discharge: A six-month prospective cohort study

Reference of the published version of this study: Pinto-Ramos J, Moreira T, Costa L, Costa F, Barroso J, Sousa-Pinto B. Association between knee extension strength and functional capacity after intensive care unit discharge: A six-month prospective cohort study. American Journal of Physical Medicine and Rehabilitation. 2022

3.3.1 Methods

Study Design

This study was written according to STROBE statement guidelines.[95] In this prospective study, we assessed a consecutive sample of participants after ICU discharge. Patients were initially assessed during hospitalization within three days after walking capability was recovered - Assessment 1 (A1). Patients were subsequently assessed three months (Assessment 2 - A2) and six months after A1 (Assessment 3 - A3). (Figure 6). For each patient assessment, we measured knee extension strength and quadriceps femoris thickness respectively with HHD and POCUS, and we applied a set of questionnaires (SF-36 physical functioning, Barthel Index, and International Physical Activity Questionnaire (IPAQ)) and physical tests (6MWT, 1MSTS, TUG test, handgrip strength, Medical Research Council Sum Score (MRCSS) and SPPB).

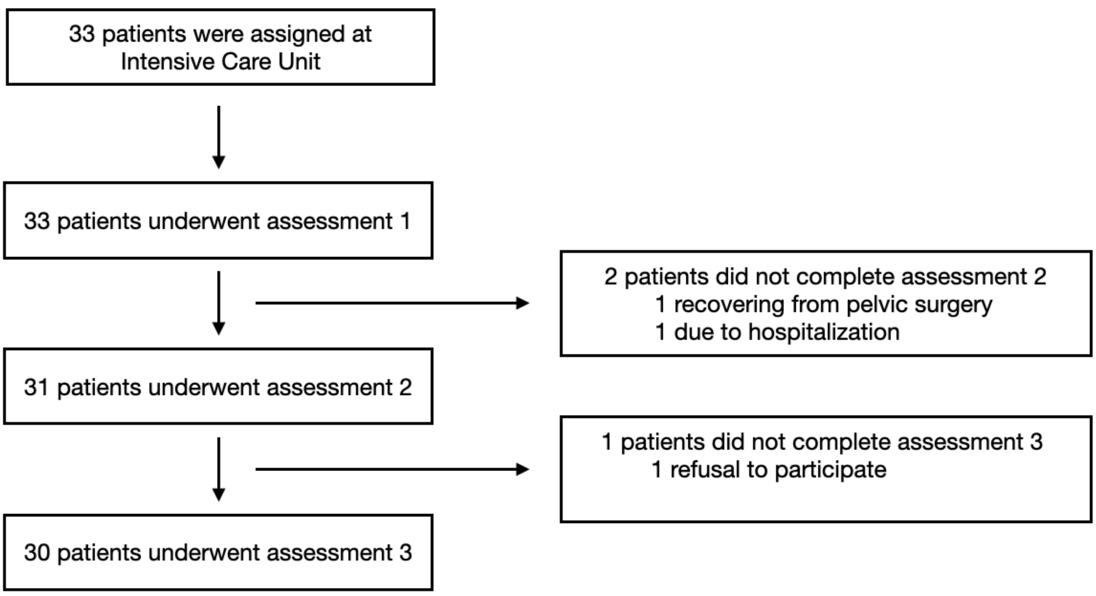


Figure 6 – Flow diagram of patients included in the study and withdrawals at each stage

This study was approved by the Ethical Committee of the hospital where this study was conducted. Patients provided written informed consent before participation to be included in the study.

Setting and Participants

Patients from a single centre (*Centro Hospitalar Universitário de São João*, a tertiary care hospital in Northern Portugal) meeting eligibility criteria were consecutively included after ICU discharge from end March to mid of August 2021. Patients aged ≥ 18 years-old who have undergone more than 5 days of ICU hospitalization and were previously able to ambulate without assistance before hospitalization were included. Exclusion criteria were severe cognitive impairment, major neurological/neuromuscular conditions such as traumatic brain injury or stroke, major psychiatric conditions, unrecoverable disease, traumatic lesions with long bone fractures, severe obesity (body mass index ≥35), full thickness pressure ulcers, cardiorespiratory instability, and current or planned home ventilation.

Variables and Measurements

Each participant was assessed three times, namely (i) during admission (within three days after walking capability was recovered – A1), (ii) three months after A1 (A2), and (iii) six months after A1 (A3). Each assessment included two HHD (Micro FET®2 HHD; Hoggan Health Industries, Draper, UT, USA) measurements of knee extension strength. The HHD was placed five centimetres above the distal portion of the medial malleolus of the dominant leg, with the patient seated in the edge of the stretcher. In addition, each assessment encompassed two measurements of both rectus femoris and quadriceps femoris muscle thickness using a POCUS (Butterfly iQ+, Butterfly Network, Inc., Guilford, CT). The POCUS was placed 15 centimetres above the superior pole of the patella of the same leg of the HHD measurements, with the patient in supine position and minimal hand force applied perpendicular to muscle fibres.[96, 97]

In each assessment, we applied a set of questionnaires (namely the SF-36 physical functioning, the Barthel Index and IPAQ), and physical function tests (namely the 6MWT, 1MSTS, TUG test, handgrip strength, MRCSS and SPPB) to all participants. Additionally, at A1, we assessed patient-reported IPAQ and Barthel index previous to hospital admission (figure 7). Questionnaires and physical function tests were applied by one of three PMR residents (according to schedule availability) with at least one year of experience of physical tests performance and both HHD and POCUS experience. Residents were blinded for previous results from each patient or from other patients.

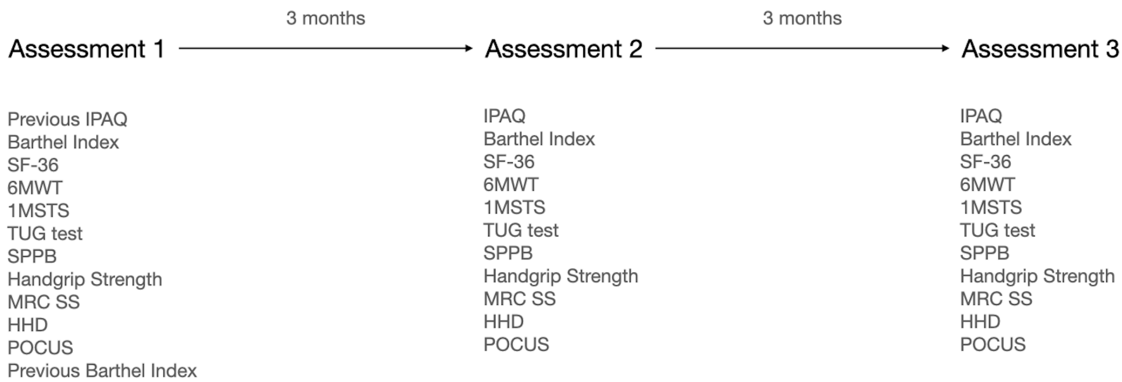


Figure 7 – Variables collected at each assessment

1MSTS - 1-Minute Sit-to-Stand; 6MWT - 6 Minutes Walking Test; HHD - Handheld Dynamometer; IPAQ - International Physical Activity Questionnaire; MRC SS - Medical Research Council Sum Score; POCUS -point-of-care ultrasound; SF-36 – 36-item Short Form Survey; SPPB - Short Physical Performance Battery; TUG – Timed Up and Go;

Information on the participants' demographic characteristics (sex and age), height, weight, body mass index, diagnosis at admission, Acute Physiology Assessment and Chronic Health Evaluation II (Apache II), Simplified Acute Physiology Score II (SAPS II) and hospitalization in ICU were retrieved from their electronic health records.

The assessments were performed in different settings, with A1 having been performed during hospitalization in the ward where the patient had been admitted, and A2 and A3 having been performed during PMR outpatient visits.

Study Size

Sample size calculation was made to detect as significant a correlation coefficient between HHD or POCUS variation and SF-36 physical functioning variation of at least 0.5 (defined by the COSMIN guidelines as indicator of good correlation).[98] A sample size of 33 was found based on a power of 80%, a level of significance of 0.05, and considering a 15% patient withdrawal.

Statistical Analysis

Means and SD were used for describing continuous variables with normal distributions whereas medians and 25th and 75th percentiles (P25;P75) were used for variables with non-normal distributions. Absolute and relative frequencies were used to describe categorical variables.

We calculated changes (arithmetic difference between results at different assessment periods) between results of HHD, POCUS, and applied questionnaires and physical tests (with such differences henceforth indicated as "Δ"). We estimated the correlations – using Spearman correlation coefficients – between Δ HHD or Δ POCUS at each time interval and (i) Δ SF-36 physical functioning, (ii) Δ Barthel Index, (iii) Δ IPAQ, (iv) Δ 6MWT, (v) Δ MRCSS, (vi) Δ 1MSTS; (vii) Δ TUG, (viii) Δ SPPB Chair Stand Test, (ix) Δ SPPB Balance Tests, (x) Δ SPPB Gait Speed Test; (xi) Δ SPPB and (xii) Δ Handgrip Strength. Whenever Spearman correlation coefficients were higher than 0.30, we built multiple linear regression models using either baseline and Δ muscle strength (as measured with HHD) or muscle thickness (POCUS) as independent variables, and Δ of each of the questionnaires and physical function tests, as dependent variable. This allowed to obtain models predicting change in functionality based on two variables: (i) baseline functionality, (ii) and change in muscle strength or thickness for each of these models, we retrieved the r^2 and the omnibus *p*-value. Alternative models were built adjusting for the baseline dependent variable (instead of the baseline muscle strength and thickness), as suggested by Husted et al.[(99)] Further covariates were not included in the models due to sample size restrictions. Additionally, we calculated, for each patient and for each of the used main function-related parameters (namely, the SF-36 physical functioning, the 1MSTS, the 6MWT, the SPPB and the TUG tests), the number of minimal clinically important differences (MCID) occurred in between the last and the first assessments ("e.g., for 6MWT, the MCID is 50 meters. A patient whose 6MWT improves by 100 meters, displays a recovery corresponding to 2 MCID"). We then calculated the average of the number of MCIDs registered for the different parameters, and we correlated it with HHD weight adjusted strength variation in the same period.

Statistical significance was defined at *p*-value <0.05. All analyses were performed using software IBM SPSS Statistics 28 (IBM, Armonk, NY).

3.3.2 Results

We included thirty-three participants who agreed to participate in this study and completed A1. Six months' follow-up was accomplished for 30 participants, with assessments being finished in February 2022. Two patients did not complete A2 and A3 due to hospitalization and pelvic surgery respectively, and one patient did not complete A3 due to refusal to participate. (Figure 6)

Among patients who completed their follow-up, 23 were male (76.7%), mean participants' age was of 60.4 years old (SD=14.3), and the average body mass index was 25.5 kg/m^2 (SD=4.2) (Table 14). Participants were hospitalized in the ICU for a median of 10 days (P25;P75=5.75;17.25), with 16 (53.3%) patients having received invasive ventilation, 14 (46.7%) non-invasive ventilation, and 2 (6.7%) Extra Corporeal Membrane Oxygenation. The ICU gravity score SAPSS II mean value was of 41.8 (SD=18.2), and the Apache II score median value was of 19 (P25;P75=12.25;26.25). Prior to admission, patients' median Barthel Index was of 100 (P25;P75=100;100), and the median IPAQ was of 627 (P25;P75=99;2973).

Table 14: Demographic and clinical characteristics of the assessed sample

Variables	**Patients who completed follow-up (N=30)**
Age (years) – mean (SD)	60.4 (14.3)
Males – *n* (%)	23 (76.7)
Weight (kg) - mean (SD)	73.3 (14.9)
Height (m) – mean (SD)	1.69 (0.07)
BMI (Kg/m^2) – mean (SD)	25.5 (4.2)
ICU Hospitalization (days) – median (P25;P75)	10 (5.75;17.25)
Previous Barthel Index – median (P25;P75)	100 (100;100)
Previous IPAQ – median (P25;P75)	627 (99;2973)
Apache II – median (P25;P75)	19 (12.25;26.25)
SAPSS II – mean (SD)	41.8 (18.2)

Non Invasive Ventilation - *n* (%)	14 (46.7)
Invasive Ventilation – *n* (%)	16 (53.3)
ECMO – *n* (%)	2 (6.7)
Tracheostomy – *n* (%)	16 (53.3)

BMI – Body Mass Index, ECMO - Extra Corporeal Membrane Oxygenation, IPAQ - International Physical Activity Questionnaire, Kg – Kilogram, m – Meter, P25;P75 – 25th and 75th Percentile, s – Seconds, SD – Standard Deviation;

Variables Progression During Follow-Up

Tests assessing physical function, such as gait speed test, TUG, 6MWT, 1MSTS, Hand grip strength and SPPB had a greater improvement from A1 to A2 than from A2 to A3 (Figure 8-9). The same pattern was seen in SF-36 Physical Function, whereas improvements in the SF-36 General Health remained stable over the follow-up period. Patients' median level of activity reported by IPAQ previous to admission (594 MET-min/week; P25;P75=165;2976) was not achieved during the follow-up period, with an A3 median value of 396 MET-min/week (P25;P75=99;1640). On the other hand, patients' disability measured by the median Barthel index matched pre-admission levels after 3 months at visit A2, with patients reaching previous levels on basic functions.

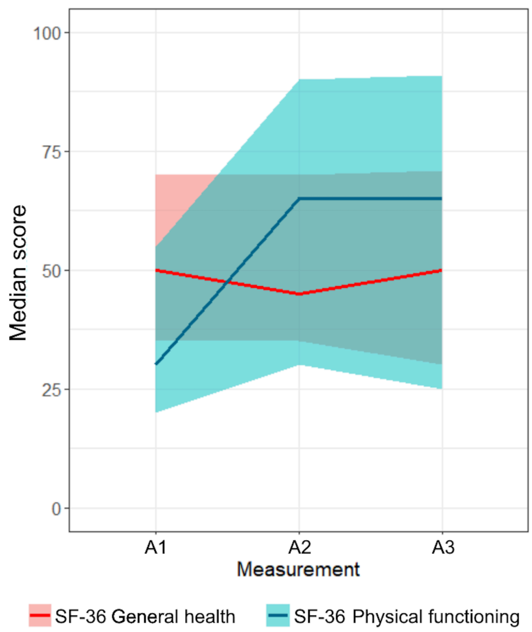


Figure 8 – Median and 25th and 75th Percentile values of 36-item Short Form Survey Physical Functioning and General Health at Baseline, 3 months and 6 months.

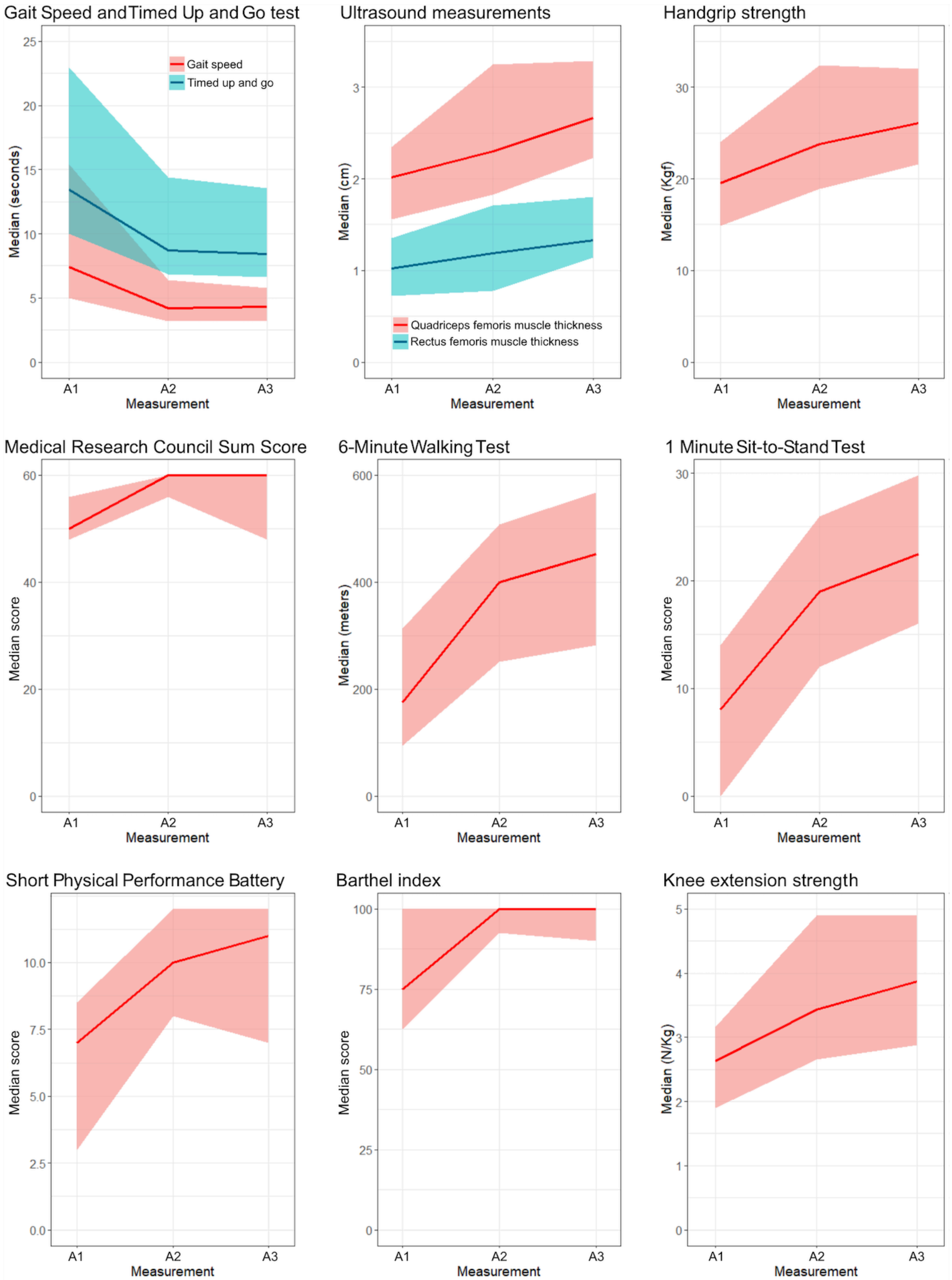


Figure 9 - Median and 25th and 75th Percentile values of different functional variables at Baseline, 3 months and 6 months.

Both quadriceps and rectus femoris muscle thickness increased steadily during the follow-up period, while knee extension muscle strength had a larger increase between A1 and A2, than between A2 and A3. (Figure 8-9)

Association between variations of functional variables and muscle strength or thickness

Moderate correlation was found between Δ knee extension muscle strength and Δ SF-36 Physical Function at Δ A1-A3 and Δ A1-A2 but not at Δ A2-A3 (Spearman correlation coefficients=0.53 (p=0.003), 0.45 (p=0.017), and -0.01 (p=0.967), respectively) (Table 15). In multiple linear regression models, Δ knee extension muscle strength was shown to be significantly associated with Δ in SF-36 physical functioning in the A1-A3 period (r^2=0.32; *p*=0.006). (Table 16) The Δ knee extension muscle strength was also correlated with other variables, with the strongest correlations being those with Δ 6MWT, Δ 1MSTS, and Δ handgrip strength (this latter one only for Δ A1-A2). These variables were also those associated with the best-performing multiple linear regression models, particularly at Δ A1-A2. Overall, correlations tended to be stronger and regression models tended to display a better performance when considering the Δ A1-A3 or Δ A1-A2 periods than the Δ A2-A3 periods. For the association between muscle strength change in the A1-A3 period and that composite measure, we obtained a correlation coefficient of 0.56 (p=0.002) and a linear regression coefficient was 1.79 (95%CI=0.63-2.94; p=0.004; r^2=0.29).

Table 15: Spearman correlation coefficients for the associations between changes of knee extension muscle strength or muscle thickness and change of functional variables for the same period

Functional variable	Change between A1 and A3 (p-value)	Change between A1 and A2 (p-value)	Change between A2 and A3 (p-value)
A. Correlations with knee extension muscle strength change			
SF-36 Physical Function	0.53 (0.003)	0.45 (0.017)	-0.01 (0.967)
Barthel Index	0.30 (0.113)	-0.05 (0.816)	0.08 (0.686)
6MWT	0.38 (0.045)	0.48 (0.009)	0.43 (0.024)
MRCSS	0.09 (0.623)	0.23 (0.224)	0.10 (0.597)
1MSTS	0.52 (0.003)	0.81 (<0.001)	0.60 (<0.001)
TUG	-0.28 (0.133)	-0.39 (0.035)	-0.41 (0.029)
SPPB Chair Stand Test	-0.04 (0.880)	-0.15 (0.573)	-0.13 (0.528)
SPPB Balance Tests	-0.02 (0.920)	0.19 (0.307)	-0.03 (0.873)
SPPB Gait Speed Test	0.13 (0.483)	0.05 (0.790)	-0.21 (0.280)
SPPB	0.38 (0.041)	0.40 (0.027)	0.26 (0.158)
Hand Grip Strength	0.19 (0.312)	0.80 (<0.001)	0.29 (0.118)
POCUS Rectus Femoris	0.34 (0.062)	0.46 (0.009)	0.35 (0.056)
POCUS Quadriceps Femoris	0.32 (0.090)	0.47 (0.008)	0.33 (0.071)
B. Correlations with rectus femoris muscle thickness change			
SF-36 Physical Function	0.03 (0.886)	0.15 (0.461)	0.34 (0.077)
Barthel Index	0.17 (0.370)	-0.15 (0.442)	-0.01 (0.948)
6MWT	0.28 (0.155)	0.10 (0.628)	0.37 (0.059)
MRCSS	0.16 (0.396)	-0.08 (0.672)	-0.01 (0.956)
1MSTS	0.24 (0.193)	0.46 (0.009)	0.33 (0.077)
TUG	-0.29 (0.124)	-0.09 (0.639)	0.12 (0.528)
SPPB Chair Stand Test	0.33 (0.217)	0.14 (0.599)	-0.34 (0.086)
SPPB Balance Tests	-0.03 (0.877)	-0.18 (0.330)	-0.15 (0.438)
SPPB Gait Speed Test	-0.25 (0.189)	0.09 (0.622)	-0.17 (0.374)
SPPB	0.15 (0.445)	-0.07 (0.716)	0.23 (0.215)
Hand Grip Strength	0.02 (0.931)	0.37 (0.043)	0.17 (0.381)
C. Correlations with quadriceps femoris muscle thickness change			
SF-36 Physical Function	0.11 (0.564)	0.19 (0.344)	0.36 (0.064)
Barthel Index	-0.01 (0.971)	-0.21 (0.285)	-0.05 (0.818)
6MWT	0.14 (0.482)	-0.05 (0.806)	0.43 (0.024)
MRCSS	0.14 (0.478)	0.06 (0.737)	-0.05 (0.787)
1MSTS	0.23 (0.213)	0.41 (0.024)	0.31 (0.091)
TUG	-0.26 (0.104)	-0.09 (0.656)	0.10 (0.602)
SPPB Chair Stand Test	0.23 (0.399)	0.10 (0.701)	-0.34 (0.083)
SPPB Balance Tests	-0.01 (0.950)	0.10 (0.611)	-0.16 (0.397)
SPPB Gait Speed Test	-0.20 (0.296)	0.20 (0.296)	-0.24 (0.201)
SPPB	0.10 (0.609)	0.01 (0.958)	0.26 (0.172)
Hand Grip Strength	0.22 (0.234)	0.41 (0.022)	0.09 (0.632)

1MSTS – 1-minute sit-to-stand; 6MWT – 6 Minutes Walking Test; MRCSS – Medical Research Council Sum Score; POCUS – Point-of-Care Ultrasound; SF-36 – 36-item Short Form Survey; SPPB – Short Physical Performance Battery; TUG – Timed Up and Go; A1=Assessment at baseline; A2=Assessment at 3 months; A3=Assessment at 6 months.

Table 16: Results of the linear regression models predicting change in functional tests based on change and baseline knee extension muscle strength

Functional variable	Change between A1 and A3		Change between A1 and A2		Change between A2 and A3	
	r^2	Predicting formula [p-value]	r^2	Predicting formula [p-value]	r^2	Predicting formula [p-value]
SF-36 Physical Function	0.32	y = 17.7x - 3.8h + 8.3 [0.006]	0.21	y = 11.7x + 0.5h + 0.5 [0.053]	-	-
Barthel Index	0.20	y = 2.5x - 7.8h + 32.0 [0.053]	-	-	-	-
6MWT	0.17	y = 27.9x - 36.9h + 253.1 [0.103]	0.33	y = 54.4x - 26.2h + 164.5 [0.006]	0.15	y = 20.9x - 4.1h + 57.0 [0.142]
1MSTS	0.22	y = 4.6x - 0.1h + 7.9 [0.038]	0.62	y = 7.1x + 0.2h + 2.2 [<0.001]	0.40	y = 3.6x + 0.6h + 0.2 [0.001]
TUG	-	-	0.16	y = -2.4x + 11.4h – 43.6 [0.100]	0.08	y = 0.2x + 1.1h – 5.2 [0.345]
SPPB	0.14	y = 0.5x - 0.5h + 3.8 [0.133]	0.14	y = 0.7x - 0.3h + 2.9 [0.117]	-	-
Hand Grip Strength	-	-	0.54	y = 5.5x + 2.5h - 8.2 [<0.001]	-	-
POCUS Rectus Femoris	0.23	y = 0.12x - 0.10h + 0.48 [0.028]	0.20	y = 0.14x - 0.10h - 0.19 [0.047]	0.26	y = 0.07x - 0.10h + 0.48 [0.019]
POCUS Quadriceps Femoris	0.19	y = 0.18x - 0.14h + 0.87 [0.056]	0.23	y = 0.28x - 0.18h - 0.25 [0.027]	0.27	y = 0.09x - 0.15h - 0.76 [0.014]

y – functional variable change between assessments; x – rectus femoris muscle thickness change between assessments, in centimeters, h – Knee extension muscle strength at first of the two assessments, in Newton per kilogram.

1MSTS – 1-minute sit-to-stand; 6MWT – 6 Minutes Walking Test; POCUS – Point-of-Care Ultrasound; SF-36 – 36-item Short Form Survey; SPPB – Short Physical Performance Battery; TUG – Timed Up and Go; A1=Assessment at baseline; A2=Assessment at 3 months; A3=Assessment at 6 months

Δ Rectus Femoris muscle thickness and Δ Quadriceps Femoris muscle thickness displayed weak correlations (correlation coefficient <0.30) with most Δ functional variables, with the exception of SF-36 Physical Function, 1MSTS, 6MWT, SPPB Chair Stand Test and handgrip strength. (Table 15) Multiple linear regression showed statistically significant results for 6MWT at Δ A2-A3 for both rectus and quadriceps femoris and 1MSTS at Δ A1-A2 for POCUS rectus femoris only.

We performed a sensitivity analysis, building multiple linear regression models adjusting for each of the baseline dependent variables. For knee extension muscle strength, those models tended to display a better performance than those based on baseline and Δ strength (Table 17). For muscle thickness, this trend was also observed, although differences in model performance tended to be much smaller (Table 18).

Table 17. Results of the linear regression models predicting change in functional tests based on change in knee extension muscle strength and baseline functional tests

Functional variable	Change between A1 and A3		Change between A1 and A2		Change between A2 and A3	
	r^2	Predicting formula [p-value]	r^2	Predicting formula [p-value]	r^2	Predicting formula [p-value]
SF-36 Physical Function	0.39	y = 14.6x - 0.35h + 15.8 [0.002]	0.38	y = 10.9x - 0.5h + 22.4 [0.003]	-	-
Barthes Index	0.86	y = 2.0x - 0.83h + 78.3 [<0.001]	-	-	-	-
6MWT	0.34	y = 45.2x - 0.38h + 220.4 [0.006]	0.55	y = 62.8x - 0.41h + 173.8 [<0.001]	0.55	y = 62.8x - 0.41h + 173.8 [<0.001]
1MSTS	0.32	y = 4.8x - 0.41h + 11.0 [0.005]	0.71	y = 7.0x - 0.39h + 6.2 [<0.001]	0.71	y = 7.0x - 0.39h + 6.2 [<0.001]
TUG	-	-	0.94	y = -1.7x - 0.90h + 11.0 [<0.001]	0.94	y = -1.7x - 0.90h + 11.0 [<0.001]
SPPB	0.35	y = 0.8x - 0.33h + 4.4 [0.003]	0.33	y = 0.90x - 0.37h + 4.2 [0.004]	-	-
Hand Grip Strength	-	-	0.50	y = 4.9x - 0.12h + 1.3 [<0.001]	-	-
POCUS Rectus Femoris	0.25	y = 0.13x - 0.21h + 0.44 [0.020]	0.15	y = 0.12x + 0.05h + 0.05 [0.100]	0.15	y = 0.12x + 0.05h + 0.05 [0.100]
POCUS Quadriceps Femoris	0.16	y = 0.21x - 0.06h + 0.61 [0.100]	0.21	y = 0.24x + 0.16h - 0.09 [0.037]	0.21	y = 0.24x + 0.16h - 0.09 [0.037]

y – functional variable change between assessments; x – rectus femoris muscle thickness change between assessments, in centimeters, h – functional variable value at first of the two assessments.
1MSTS – 1-minute sit-to-stand; 6MWT – 6 Minutes Walking Test; POCUS – Point-of-Care Ultrasound; SF-36 – 36-item Short Form Survey; SPPB – Short Physical Performance Battery; TUG – Timed Up and Go; A1=Assessment at baseline; A2=Assessment at 3 months; A3=Assessment at 6 months

Table 18. Results of the linear regression models predicting change in functional tests based on change in baseline muscle thickness and baseline functional tests

Functional variable	Change between A1 and A3		Change between A1 and A2		Change between A2 and A3	
	r^2	Predicting formula [p-value]	r^2	Predicting formula [p-value]	r^2	Predicting formula [p-value]
A. Models involving the rectus femora's muscle thickness						
SF-36 Physical Function	-	-	-	-	0.20	y = 13.9x - 0.07r + 8.5 [0.061]
6MWT	-	-	-	-	0.24	y = 68.8x - 0.10r + 72.1 [0.040]
1MSTS	-	-	0.35	y = 13.4x - 0.47r + 11.8 [0.002]	0.09	y = 3.2x – 0.09r + 4.5 [0.296]
SPPB Chair Stand Test	0.53	y = 0.08x – 0.68r + 5.6 [0.007]	-	-	0.24	y = -2.8x – 0.16r + 1.4 [0.040]
Hand Grip Strength	-	-	0.21	y = 8.7x - 0.24r + 7.5 [0.035]	-	-

B. Models involving the quadriceps femoris muscle thickness						
SF-36 Physical Function	-	-	-	-	0.23	y = 11.3z – 0.07q + 8.0 [0.042]
6MWT	-	-	-	-	0.26	y = 56.2z - 0.09q - 65.9 [0.026]
1MSTS	-	-	0.32	y = 7.1z – 0.54q + 11.9 [0.004]	0.10	y = 4.2z + 0.06q + 0.69 [0.239]
SPPB Chair Stand Test	-	-	-	-	0.26	y = -2.3z - 0.17q + 1.7 [0.026]
Hand Grip Strength	-	-	0.22	y = 5.0z - 0.24q + 7.1 [0.030]	-	-

y – functional variable change between assessments; x – rectus femoris muscle thickness change between assessments, in centimeters, r – rectus femora's muscle thickness at first of the two assessments, in centimeters; z – quadriceps femora's muscle thickness change between assessments, in centimeters, q – quadriceps femora's muscle thickness at first of the two assessments, in centimeters

1MSTS – 1-minute sit-to-stand; 6MWT – 6 Minutes Walking Test; SF-36 – 36-item Short Form Survey; SPPB – Short Physical Performance Battery; A1=Assessment at baseline; A2=Assessment at 3 months; A3=Assessment at 6 months

4. Discussion

4.1 Major Findings

These studies suggest that the HHD may be a reliable tool for estimating knee extensors muscle strength in rehabilitation patients, with both intra and inter-rater ICC being higher than 0.9 across all observers. We also observed that POCUS may be a reliable tool to measure muscle thickness in rehabilitation patients, with its intra-class and inter-class ICC displaying overall high values across all observers. For both HHD and POCUS, performing two measurements for each assessment (with subsequent computation of the average value) appears to result in increased reliability, potentially overcoming random errors in the measurements. The learning curve for both methods may be apparently short, with experienced and inexperienced observers showing similar results (i.e., similarly high ICC) after a brief course of training.

When both methods were used in a prospective way to assess patients in the post-ICU context, we found that weight-adjusted knee extension strength change over 6 months (as measured with HHD) seemed to be correlated with changes in function-related tests and questionnaires such as SF-36 Physical Function, 6MWT, 1MSTS and SPPB. Correlation was higher when information on strength change was combined with that on baseline functionality than with that on baseline strength. By contrast, changes in quadriceps and rectus femoris muscle thickness were not correlated with function-related tests and questionnaires.

In our prospective assessment of patients after ICU discharge, we also found that function-related tests and questionnaires had a larger improvement in the first three months after discharge in comparison to the subsequent three months. In fact, correlations between changes in knee extension strength and functioning tended to be better when considering the first three months after discharge than the subsequent months.

4.2 Limitations and Strengths

This thesis has some limitations which are common to all studies, including the fact that all studies were done in a single centre, and caution is required when generalising their results to other populations. Also, as patients were seen in the context of rehabilitation programmes (either in the PMR outpatient setting or following admission in an acute tertiary care ward), results may not necessarily apply to patients from other settings.

On the other hand, we ensured that all studies were adequately powered to obtain precise estimates, namely by performing sample size calculations. Also, to the best of our knowledge, there are no previous studies evaluating (i) reliability of HHD or POCUS in rehabilitation settings or (ii) the correlation between function-related parameters evolution and changes in muscle strength and thickness measurements using HHD and POCUS.

In addition, there are some limitations and strengths which are specific to cross-sectional or the prospective studies, and which will be discussed in the following subsections.

4.2.1 Reliability of Handheld Dynamometer and Point-of-Care Ultrasound studies

Some limitations were common to both cross-sectional studies. One potential limitation concerns the possibility of memory biases from the observers, given the three-hour gap between the two sets of measurements. However, we believe that this limitation may not have had such a relevant impact, as in each day there were seven to nine patients being assessed by each independent observer (resulting in at least 56 measurements to remember before the second set of assessments). Another limitation stems from the fact that patients' muscular conditions were heterogeneous (as patients displayed different pathologies and stages of disease) and muscle thickness and strength variability was substantial in the assessed sample, possibly resulting in higher ICC values. In fact, lower values could have been obtained if a more homogenous population had been assessed.[90] To overcome this limitation, and in order to estimate the actual differences, absolute and percentual differences between assessments were calculated.

Regarding specific limitations of each performed study, in the HHD reliability study, patients were observed in two assessments by four observers, so that at each set of measurements the patient was subject to eight maximum knee strength evaluations performed within a short period (being subject to sixteen of such evaluations within a period of three hours). As a result, patient fatigue may have modified the results between measurements and lead to underestimation of results. Authors tried to manage this limitation by setting a minimum of one-minute rest between each measurement, so that maximum peak force could be reached every time.

On the other hand, in POCUS reliability study, the probe was applied 15 centimetres above the superior pole of the patella, for simplicity purposes, which may not correspond to the thickest part of the muscle. Nevertheless, other previous studies took the same approach as we did.[100, 101] Finally, we performed each pair of measurements separated by a three-hour interval, so that there may be hourly changes in patients' muscle thickness that may have resulted in a reduced reliability of measurements. However, previous studies did not find significant differences in rectus femoris thickness when measurements were done with 1, 6 and 24 hours intervals.[102]

These cross-sectional studies have also important strengths. Firstly, they were performed with more than two observers, differentiating them from most other studies assessing the reliability of both HHD or POCUS. Also, these studies compared different ways of calculating muscle strength and thickness, namely comparing the performance of two measurements *versus* a single measurement, in order to understand what could be the most reliable approach. Having four observers also allowed us to include observers with different experience levels, making us able to assess whether past experience in HHD or POCUS use was associated with different results.

4.2.2 Knee Extension Strength and Functional Capacity after Intensive Care Unit Discharge – A six-month prospective cohort study

The first important limitation of the prospective study concerns the possibility of selection biases. In fact, this study was conducted in a tertiary care hospital, where patients with higher Apache II and SAPSS II scores are expected to be overrepresented, since more severely diseased patients from other hospitals are transferred to this centre. Although all patients were from the ICU setting, there was heterogeneity in their severity scores and admission periods, limiting the results' generalizability. Patients were evaluated in the ward, three days after starting ambulation, and some may have overestimated their real capacity when answering SF-36 Physical Function, since they had a short time to perceive their disability. This may justify why some patients reported deterioration of capacity/performance in SF-36 from A1 to A3 while improving in function-related tests at the same time. This potential information bias may have led to an underestimation of the association between knee extension muscle strength change and SF-36 change. Patients were assessed in different locations, with A1 assessments having been performed in the ward during admission, and A2 and A3 in the outpatient setting. While this may have influenced the performed measurements, the effect of this phenomenon is probably non-differential (i.e, probably not associated with patients' specific baseline characteristics). Another limitation concerns the absence of a single gold-standard comparator assessing functioning – while we were able to compare changes in knee extension muscle strength or thickness with changes in several function-related tests, none of these tests is a perfect gold-standard for functioning and associations might have been stronger if there were a better way to measure their change over time.[15, 103, 104]

Finally, the assessed sample size precluded us from including further covariates in linear regression models, including the simultaneous testing of knee extension muscle strength and muscle thickness changes, although this was just an exploratory study.

This study has also some important strengths. Firstly, this is, to our knowledge, the first prospective study to evaluate the correlation of the change in function-related tests, questionnaires, muscle strength and muscle thickness after ICU hospitalization. These function-related tests and questionnaires have been validated in different settings. According to this thesis' previous studies, HDD and POCUS were found to display strong intra- and inter-rater reliability.[96, 97] In addition, these methods are portable, fast to use, and inexpensive, rendering them easily used in the clinical practice. The consistency of the obtained results also appears to be an important strength of our study.

4.3 Discussion of the obtained results

4.3.1 Handheld dynamometer reliability to measure knee extension strength in rehabilitation patients - A cross-sectional study

The HHD was found to be a reliable tool of measuring muscle strength, with results showing excellent reliability, in accordance with what has been previously reported in previous studies assessing other types of patients.[53-59]

Absolute differences between assessments were relatively low, despite being higher in patients with higher muscle strength. This might occur because knee extension of stronger patients is possibly more difficult to resist, leading to higher errors, specially between observers with less capability to resist knee extension strength. In fact, other studies already suggested that stronger observers tend to report higher values on HHD and also that external stabilization of the HHD, namely with a belt, can be more reliable than human-resisted evaluations.[59, 69, 105, 106] Nevertheless, belt stabilization is more difficult and time-consuming, possibly rendering it impractical to apply in clinical practice. By contrast, the percentual differences did not vary with increase of muscle strength.

Results were similar when considering measurements performed by both experienced and previously inexperienced observers, which suggest that the use of HHD to measure muscle strength has a brief learning curve and could be generalized across physicians if a short course of training is previously done. Since this study was conducted throughout a period of four months, we cannot estimate if reliability could be maintained this high for longer periods if observers were not trained again for standardized evaluations. This was not the first study to show that inexperienced subjects with a short course training could achieve good reliabilities with HDD, although in different muscular groups.[107, 108]

Knee extension strength measured with HHD was moderately correlated with the majority of analysed function-related variables, confirming that muscle strength, including for knee extension, may be important for these tasks, as it was also suggested by other studies.[109-112] In these participants, correlations between function-related tests and muscle strength were higher than those observed between function-related tests and muscle thickness (except for hand grip strength). This suggests that, compared with muscle thickness, muscle strength correlates better with function-related variables for this specific population, which can be justified by many factors, such as muscular or polyneuropathic changes due to hospitalization, which can affect thickness and strength in a different manner.[113, 114]

The HHD overcomes many of the limitations associated with the use of IKD (considered the gold-standard for the assessment of strength). In fact, while the latter requires patients to move to a different location to be tested, well-trained users and expensive initial investment, HHD is portable, fast to use (with measurements done in less than two minutes during a consultation/medical evaluation) and displays lower costs. This may propel the measurement of muscle strength in clinical practice which has been very limited so far.

Therefore, this study allowed us to advance knowledge in muscle strength measurements using HHD, namely that their reliability is very high, even in patients from a rehabilitation setting. Moreover we have shown that HHD can be used with physicians with low experience (if a small set of training is given) and that more reliable information may be obtained with the average of two measurements rather than with the use of just one measurement.

4.3.2 Reliability of Point-of-Care Ultrasound for Measuring Quadriceps Femoris Muscle Thickness

The reliability of POCUS to measure rectus femoris and quadriceps femoris muscle thickness was found to be very high, which is in line with previous studies.[82-86] Our results also suggest that measurements are reliable even when done by different observers and at different time points.

The learning curve for this procedure appears to be faster than with other US procedures, since experienced and inexperienced observers displayed similar reliability results. Absence of significant differences between experienced and inexperienced observers with the use of POCUS were also observed in other fields, such as rheumatology, vascular and emergency medicine as well as for quadriceps muscle thickness measured in intensive care units.[115-118] This suggests that this technique can be used across a wider set of physicians with a single and short training course in order to standardize data collection.

The absolute difference between measurements of the same and different observers had a poor correlation with muscle thickness, which suggests that between-measurement differences are usually low, independently of the muscle thickness differences. This is the opposite to what might have been intuitively thought (i.e., that assessments in thicker muscles would have higher errors).

Muscle thickness was moderately correlated with knee extension peak force, sit to stand test and handgrip strength, and weakly correlated with other function-related variables. Previous studies found similar results; this may be due to the fact that thickness does not perfectly correlate with the muscle cross-sectional area and, more importantly, with muscle quality, such as fat and fibrosis levels.(119-123)

The high reliability of POCUS and the possibility of its wider use across physicians is especially important for the diagnosis of conditions such as sarcopenia which, according to some algorithms, requires the measurement of quadriceps femoris muscle thickness.(124-126) Sarcopenia is a condition related with loss of muscle mass and strength, which is defined as “age-related loss of muscle mass and muscle function” and primarily targets the anterior tight muscles, being one of the most common conditions leading to disability in the elderly.(124, 125) The existence of a method to reliably assess muscle thickness and, therefore, contribute to the diagnosis of sarcopenia is all the more relevant as sarcopenia is an underdiagnosed condition.(127)

Therefore, this study allowed us to confirm the high reliability of POCUS, also in the PMR area, while also suggesting that POCUS can be used for a wider set of physicians, which can be useful in the detection of some conditions such as sarcopenia.

4.3.3 Association between knee extension strength and functional capacity after intensive care unit discharge: A six-month prospective cohort study

Knee extensor muscle strength measured with HHD may be helpful for estimating functioning gains over same period. This was particularly suggested by the results observed in its association with SF-36 Physical Function, 6MWT, 1MSTS and the composite outcome reflecting different function-related tests and questionnaires. In the case of SF-36 Physical Function, such association was not found for the variation between 3 and 6 months' assessments, which might be explained for a difficulty in this variable to measure the minor capacity/performance changes which are seen in this period. For other function-related tests, weaker associations between their change and that of knee extension muscle strength were observed. This difference may occur because knee extension strength is more critical for some functional tasks than others, namely 1MSTS, while some depend more on endurance, balance or coordination. This diversity of results and the absence of a single gold-standard to assess functioning prompts the need for further studies assessing how changes in strength predict functioning gains. Our results suggest that changes in strength can at least inform on the evolution of some dimensions of functioning in an easy and less time-consuming way. In fact, preparing the patient, assessing and scoring the SF-36 Physical Function, the 6MWT and the 1MSTS could take more than 20 minutes – (i) the SF-36 may be difficult to use for some patients and time-consuming to score by the observers, and (ii) the 6MWT and 1MSTS usually require more than theoretically-expected 7 minutes to be performed, given the time needed to provide patients with instructions on how to execute the tests and the time needed to relocate them. On the other hand, the HHD assessment could take less than 2 minutes since a direct value is obtained after a couple of knee extension tests. This is particularly relevant for patients' follow-up in the outpatient setting (in fact, the possibility of collecting function-related variables at baseline can improve the estimation of those variables improvement, complementing information on strength variation).

Additionally, when prospectively assessing post-ICU patients, patient progression was found to be faster in the first than in the second three months after patients started ambulation. In fact, previous studies had already pointed that improvement is faster in the first months after hospitalization, probably because function-related parameters are furthest away from patients' previous state which gives them more functioning potential.[45, 128] Quadriceps and Rectus Femoris thickness were the only exception, with steady muscle growth not necessarily meaning similar functional improvement. This might happen because muscle thickness does not entirely correlate with muscle volume and because strength also depends on muscle quality (such as levels of fibrosis and fat) which was not evaluated.[120, 121] Also, we were able to observe that the Barthel Index reached levels previous to hospitalization at 3 months post-discharge. In fact, at that time point, the maximum score was achieved, probably due to the ceiling effect of this test, which limits differentiation of patients with higher levels of capacity and performance.[129] On the other hand, patients' previous IPAQ was never reached, even after 6 months after hospitalization, probably due to loss of capacity or fear of overexertion, although memory bias can also be a factor.

4.4 Implications for practice and future research

This thesis aimed to study the correlation of changes in muscle strength and thickness with functioning variation. Having a practical method to systematically estimate function-related parameters (some of which are predictors of mortality and quality of life) can give physicians crucial information on patients' improvement. Having such information can be used to manage patients' expectations and inform clinicians on their prognosis, since patients who are not improving or who are improving slower may have their treatment adjusted and intensified in order to potentially have more gains.[38] Assessing patients' prognosis and monitoring their evolution is critical for understanding which patients are still in need of additional rehabilitation sessions and which patients have already reached their full performance. This information can prevent us from treating patients who no longer need additional treatments, giving the opportunity for other patients with more potential to recover to be treated. This may have relevant implications not only from a clinical, but also from a health services point-of-view.

This thesis also contributes to promote POCUS and HHD as reliable tools to measure quadriceps thickness and muscle strength respectively, that can be used by a wide range of physicians with short course of training without significant loss of reliability. It also contributes to encourage physicians to use an average of two measurements in order to increase reliability. Reliability of POCUS and HHD seems now to be better studied, in a wider range of patients, namely in the rehabilitation setting as studied during this thesis.

Although HHD seems to be a potentially useful tool in clinical practice to estimate functioning progression in patients after intensive care unit, more studies should take place to confirm this hypothesis, namely to understand if larger improvements in knee extension strength lead to better functional and vital prognosis. This relationship should also be studied in other populations in the rehabilitation setting such as musculoskeletal, neurologic, respiratory, geriatric or cardiac rehabilitation in order to understand if these results could be applied in other contexts. Also, studies with higher sample size may be useful to develop multivariable models of prognosis with a potential to more accurately predict patients' evolution. It would also be important to have studies assessing if introducing HHD in PMR consultation would lead to a better treatment capacity, with more gains in health and function-related parameters. That is, it would be important to assess whether a better assessment of patients' functioning using a HHD may result in improved outcomes for the patient. Future studies should also assess the feasibility of using an HHD in the PMR consultation, with quantification of the extra time (or less time, if it could replace other evaluations) it would require, if clinicians could naturally incorporate it in their daily routine and if they would feel satisfied in using it systematically. Other questions that should be the target of future studies are if we should apply these tools to all patients or only in some specific cases and how economically viable or cost-effective it would be for hospital departments to provide clinicians with these tools. In fact, from a health economic point of view, it would be important to understand if using HHD could lead to cost savings in unnecessary treatments or if expenses would be compensated for health or quality-of-life gains due to treatment optimization.

On the other hand, Quadriceps and Rectus Femoris muscle thickness measured with POCUS does not seem to be a consistently valid method of estimating changes in function-related variables and its use may not be indispensable in clinical practice for assessment of those changes. In fact, muscle mass gains are consistent over time and not necessarily correlated with functioning improvements although different results could have possibly been found in other populations or sample size. This could challenge the assumption usually applied in the clinical practice that the amyotrophic status, namely as a function of thigh circumference, is a clear indicator of patients' functional condition. More studies should be done to assess the relationship between amyotrophy, thigh circumference and functional status. Irrespectively of that, POCUS may still play an important role in the PMR practice which may be explored in future studies – for example, it may be a useful tool to help diagnosing some conditions such as sarcopenia. Further studies (including cost-effectiveness and feasibility assessments) should be done to understand if a wider use of POCUS could lead to a better diagnosis of sarcopenia and if such could lead to health gains and/or cost savings.

In the future, our group purposes to study the applicability and economic viability in clinical practise of HHD usage in post-ICU patients and study these associations in a higher number of patients. Also, we aim to extend this study to other populations.

5. Conclusion

Both HHD and POCUS seem to be reliable resources to objectively measure knee extension strength and quadriceps femoris muscle thickness on patients in rehabilitation programs, respectively. These methods seem feasible to apply with good reproducibility, even for physicians with no previous experience with these tools after a short training course. This is all the more relevant as we observed that knee extension strength change measured with HHD (but not quadriceps and rectus femoris muscle thickness measured with POCUS) was found to have a moderate association with function-related parameters' change (notably SF-36 Physical Functioning, 6MWT, and 1MSTS), rendering HHD as an easily-applicable tool to estimate changes in functioning in patients after ICU discharge. Nevertheless, further confirmatory studies – with larger samples, more frequent assessments, and/or longer follow-up periods - are needed to confirm the findings obtained in this exploratory study.

6. References

1. Brach JS, VanSwearingen JM. Physical impairment and disability: relationship to performance of activities of daily living in community-dwelling older men. Phys Ther. 2002;82(8):752-61.

2. European P, Rehabilitation Medicine Bodies A. White Book on Physical and Rehabilitation Medicine (PRM) in Europe. Chapter 1. Definitions and concepts of PRM. Eur J Phys Rehabil Med. 2018;54(2):156-65.

3. [Available from: https://www.cdc.gov/ncbddd/disabilityandhealth/infographic-disability-impacts-all.html.

4. Raposo Dos Santos JR, Cascão Torres AR, Neto M, Namorado S. Prevalence of long-term disability in Portugal, 2014: Evidence of variation by personal and contextual factors. Disabil Health J. 2021;14(2):101043.

5. Murray CJ, Acharya AK. Understanding DALYs (disability-adjusted life years). J Health Econ. 1997;16(6):703-30.

6. Rantanen T. Muscle strength, disability and mortality. Scand J Med Sci Sports. 2003;13(1):3-8.

7. Hilfiker R, Bachmann LM, Heitz CA, Lorenz T, Joronen H, Klipstein A. Value of predictive instruments to determine persisting restriction of function in patients with subacute non-specific low back pain. Systematic review. Eur Spine J. 2007;16(11):1755-75.

8. Rommersbach N, Wirth R, Lueg G, Klimek C, Schnatmann M, Liermann D, et al. The impact of disease-related immobilization on thigh muscle mass and strength in older hospitalized patients. BMC Geriatr. 2020;20(1):500.

9. Zorowitz RD. ICU-Acquired Weakness: A Rehabilitation Perspective of Diagnosis, Treatment, and Functional Management. Chest. 2016;150(4):966-71.

10. Puthucheary ZA, Phadke R, Rawal J, McPhail MJ, Sidhu PS, Rowlerson A, et al. Qualitative Ultrasound in Acute Critical Illness Muscle Wasting. Crit Care Med. 2015;43(8):1603-11.

11. Dirks ML, Wall BT, van de Valk B, Holloway TM, Holloway GP, Chabowski A, et al. One Week of Bed Rest Leads to Substantial Muscle Atrophy and Induces Whole-Body Insulin Resistance in the Absence of Skeletal Muscle Lipid Accumulation. Diabetes. 2016;65(10):2862-75.

12. Parry SM, El-Ansary D, Cartwright MS, Sarwal A, Berney S, Koopman R, et al. Ultrasonography in the intensive care setting can be used to detect changes in the quality and quantity of muscle and is related to muscle strength and function. J Crit Care. 2015;30(5):1151.e9-14.

13. Fan E, Dowdy DW, Colantuoni E, Mendez-Tellez PA, Sevransky JE, Shanholtz C, et al. Physical complications in acute lung injury survivors: a two-year longitudinal prospective study. Crit Care Med. 2014;42(4):849-59.

14. Anekwe DE, Biswas S, Bussières A, Spahija J. Early rehabilitation reduces the likelihood of developing intensive care unit-acquired weakness: a systematic review and meta-analysis. Physiotherapy. 2020;107:1-10.

15. Oeyen SG, Vandijck DM, Benoit DD, Annemans L, Decruyenaere JM. Quality of life after intensive care: a systematic review of the literature. Crit Care Med. 2010;38(12):2386-400.

16. Black NA, Jenkinson C, Hayes JA, Young D, Vella K, Rowan KM, et al. Review of outcome measures used in adult critical care. Crit Care Med. 2001;29(11):2119-24.

17. Chrispin PS, Scotton H, Rogers J, Lloyd D, Ridley SA. Short Form 36 in the intensive care unit: assessment of acceptability, reliability and validity of the questionnaire. Anaesthesia. 1997;52(1):15-23.

18. Heyland DK, Hopman W, Coo H, Tranmer J, McColl MA. Long-term health-related quality of life in survivors of sepsis. Short Form 36: a valid and reliable measure of health-related quality of life. Crit Care Med. 2000;28(11):3599-605.

19. Silveira LTYD, Silva JMD, Soler JMP, Sun CYL, Tanaka C, Fu C. Assessing functional status after intensive care unit stay: the Barthel Index and the Katz Index. Int J Qual Health Care. 2018;30(4):265-70.

20. Parry SM, Nalamalapu SR, Nunna K, Rabiee A, Friedman LA, Colantuoni E, et al. Six-Minute Walk Distance After Critical Illness: A Systematic Review and Meta-Analysis. J Intensive Care Med. 2021;36(3):343-51.

21. Bohannon RW, Crouch R. 1-Minute Sit-to-Stand Test: SYSTEMATIC REVIEW OF PROCEDURES, PERFORMANCE, AND CLINIMETRIC PROPERTIES. J Cardiopulm Rehabil Prev. 2019;39(1):2-8.

22. Kear BM, Guck TP, McGaha AL. Timed Up and Go (TUG) Test: Normative Reference Values for Ages 20 to 59 Years and Relationships With Physical and Mental Health Risk Factors. J Prim Care Community Health. 2017;8(1):9-13.

23. Fisher S, Ottenbacher KJ, Goodwin JS, Graham JE, Ostir GV. Short Physical Performance Battery in hospitalized older adults. Aging Clin Exp Res. 2009;21(6):445-52.

24. Grundtvig M, Eriksen-Volnes T, Ørn S, Slind EK, Gullestad L. 6 min walk test is a strong independent predictor of death in outpatients with heart failure. ESC Heart Fail. 2020;7(5):2904-11.

25. Crook S, Frei A, Ter Riet G, Puhan MA. Prediction of long-term clinical outcomes using simple functional exercise performance tests in patients with COPD: a 5-year prospective cohort study. Respir Res. 2017;18(1):112.

26. Pavasini R, Guralnik J, Brown JC, di Bari M, Cesari M, Landi F, et al. Short Physical Performance Battery and all-cause mortality: systematic review and meta-analysis. BMC Med. 2016;14(1):215.

27. Lowrie EG, Curtin RB, LePain N, Schatell D. Medical outcomes study short form-36: a consistent and powerful predictor of morbidity and mortality in dialysis patients. Am J Kidney Dis. 2003;41(6):1286-92.

28. Ferreira PL. [Development of the Portuguese version of MOS SF-36. Part I. Cultural and linguistic adaptation]. Acta Med Port. 2000;13(1-2):55-66.

29. Ferreira PL. [Development of the Portuguese version of MOS SF-36. Part II -- Validation tests]. Acta Med Port. 2000;13(3):119-27.

30. Dos Reis NF, Figueiredo FCXS, Biscaro RRM, Lunardelli EB, Maurici R. Psychometric Properties of the Barthel Index Used at Intensive Care Unit Discharge. Am J Crit Care. 2022;31(1):65-72.

31. MAHONEY FI, BARTHEL DW. FUNCTIONAL EVALUATION: THE BARTHEL INDEX. Md State Med J. 1965;14:61-5.

32. Steffen TM, Hacker TA, Mollinger L. Age- and gender-related test performance in community-dwelling elderly people: Six-Minute Walk Test, Berg Balance Scale, Timed Up & Go Test, and gait speeds. Phys Ther. 2002;82(2):128-37.

33. Perera S, Mody SH, Woodman RC, Studenski SA. Meaningful change and responsiveness in common physical performance measures in older adults. J Am Geriatr Soc. 2006;54(5):743-9.

34. Crook S, Büsching G, Schultz K, Lehbert N, Jelusic D, Keusch S, et al. A multicentre validation of the 1-min sit-to-stand test in patients with COPD. Eur Respir J. 2017;49(3).

35. Gautschi OP, Stienen MN, Corniola MV, Joswig H, Schaller K, Hildebrandt G, et al. Assessment of the Minimum Clinically Important Difference in the Timed Up and Go Test After Surgery for Lumbar Degenerative Disc Disease. Neurosurgery. 2017;80(3):380-5.

36. Nordin E, Rosendahl E, Lundin-Olsson L. Timed "Up & Go" test: reliability in older people dependent in activities of daily living--focus on cognitive state. Phys Ther. 2006;86(5):646-55.

37. Olsen CF, Bergland A. "Reliability of the Norwegian version of the short physical performance battery in older people with and without dementia". BMC Geriatr. 2017;17(1):124.

38. Wade DT. Outcome measures for clinical rehabilitation trials: impairment, function, quality of life, or value? Am J Phys Med Rehabil. 2003;82(10 Suppl):S26-31.

39. Chaiwat O, Sathitkarnmanee B, Dajpratham P, Thanakiattiwibun C, Jarungjitaree S, Rattanamung S. The impact of physical medicine and rehabilitation consultation on clinical outcomes in the surgical intensive care unit: A prospective observational cohort study. Medicine (Baltimore). 2022;101(9):e28990.

40. Kluzek S, Dean B, Wartolowska KA. Patient-reported outcome measures (PROMs) as proof of treatment efficacy. BMJ Evid Based Med. 2022;27(3):153-5.

41. Duncan Millar J, van Wijck F, Pollock A, Ali M. Outcome measures in post-stroke arm rehabilitation trials: do existing measures capture outcomes that are important to stroke survivors, carers, and clinicians? Clin Rehabil. 2019;33(4):737-49.

42. Ali M, English C, Bernhardt J, Sunnerhagen KS, Brady M, Collaboration V-R. More outcomes than trials: a call for consistent data collection across stroke rehabilitation trials. Int J Stroke. 2013;8(1):18-24.

43. Wang DXM, Yao J, Zirek Y, Reijnierse EM, Maier AB. Muscle mass, strength, and physical performance predicting activities of daily living: a meta-analysis. J Cachexia Sarcopenia Muscle. 2020;11(1):3-25.

44. Burtin C, Clerckx B, Robbeets C, Ferdinande P, Langer D, Troosters T, et al. Early exercise in critically ill patients enhances short-term functional recovery. Crit Care Med. 2009;37(9):2499-505.

45. Denehy L, Skinner EH, Edbrooke L, Haines K, Warrillow S, Hawthorne G, et al. Exercise rehabilitation for patients with critical illness: a randomized controlled trial with 12 months of follow-up. Crit Care. 2013;17(4):R156.

46. Parry SM, Nalamalapu SR, Nunna K, Rabiee A, Friedman LA, Colantuoni E, et al. Six-Minute Walk Distance After Critical Illness: A Systematic Review and Meta-Analysis. J Intensive Care Med. 2019:885066619885838.

47. Herridge MS, Tansey CM, Matté A, Tomlinson G, Diaz-Granados N, Cooper A, et al. Functional disability 5 years after acute respiratory distress syndrome. N Engl J Med. 2011;364(14):1293-304.

48. Jung H, Tanaka S, Iwamoto Y, Kawano T, Yamasaki M, Tanaka R. Reductions in Muscle Strength and Range of Motion Cause Locomotion Disability via Locomotion-Related Functional Limitation in Japanese Older Adults: A Cross-Sectional Study. J Aging Res. 2021;2021:6627767.

49. Arnold AS, Anderson FC, Pandy MG, Delp SL. Muscular contributions to hip and knee extension during the single limb stance phase of normal gait: a framework for investigating the causes of crouch gait. J Biomech. 2005;38(11):2181-9.

50. Corrigan D, Bohannon RW. Relationship between knee extension force and stand-up performance in community-dwelling elderly women. Arch Phys Med Rehabil. 2001;82(12):1666-72.

51. Kojima N, Kim H, Saito K, Yoshida H, Yoshida Y, Hirano H, et al. Association of knee-extension strength with instrumental activities of daily living in community-dwelling older adults. Geriatr Gerontol Int. 2014;14(3):674-80.

52. Stark T, Walker B, Phillips JK, Fejer R, Beck R. Hand-held dynamometry correlation with the gold standard isokinetic dynamometry: a systematic review. PM R. 2011;3(5):472-9.

53. Martins J, da Silva JR, da Silva MRB, Bevilaqua-Grossi D. Reliability and Validity of the Belt-Stabilized Handheld Dynamometer in Hip- and Knee-Strength Tests. J Athl Train. 2017;52(9):809-19.

54. Mentiplay BF, Perraton LG, Bower KJ, Adair B, Pua YH, Williams GP, et al. Assessment of Lower Limb Muscle Strength and Power Using Hand-Held and Fixed Dynamometry: A Reliability and Validity Study. PLoS One. 2015;10(10):e0140822.

55. Sung KS, Yi YG, Shin HI. Reliability and validity of knee extensor strength measurements using a portable dynamometer anchoring system in a supine position. BMC Musculoskelet Disord. 2019;20(1):320.

56. Romero-Franco N, Jiménez-Reyes P, Montaño-Munuera JA. Validity and reliability of a low-cost digital dynamometer for measuring isometric strength of lower limb. J Sports Sci. 2017;35(22):2179-84.

57. Whiteley R, Jacobsen P, Prior S, Skazalski C, Otten R, Johnson A. Correlation of isokinetic and novel hand-held dynamometry measures of knee flexion and extension strength testing. J Sci Med Sport. 2012;15(5):444-50.

58. Toonstra J, Mattacola CG. Test-retest reliability and validity of isometric knee-flexion and -extension measurement using 3 methods of assessing muscle strength. J Sport Rehabil. 2013;22(1).

59. Florencio LL, Martins J, da Silva MRB, da Silva JR, Bellizzi GL, Bevilaqua-Grossi D. Knee and hip strength measurements obtained by a hand-held dynamometer stabilized by a belt and an examiner demonstrate parallel reliability but not agreement. Phys Ther Sport. 2019;38:115-22.

60. Kim WK, Kim DK, Seo KM, Kang SH. Reliability and validity of isometric knee extensor strength test with hand-held dynamometer depending on its fixation: a pilot study. Ann Rehabil Med. 2014;38(1):84-93.

61. Romero-Franco N, Jimenez-Reyes P, Montano-Munuera JA. Validity and reliability of a low-cost digital dynamometer for measuring isometric strength of lower limb. J Sports Sci. 2017;35(22):2179-84.

62. Schache MB, McClelland JA, Webster KE. Reliability of measuring hip abductor strength following total knee arthroplasty using a hand-held dynamometer. Disabil Rehabil. 2016;38(6):597-600.

63. Mong Y, Teo TW, Ng SS. 5-repetition sit-to-stand test in subjects with chronic stroke: reliability and validity. Arch Phys Med Rehabil. 2010;91(3):407-13.

64. Souweine JS, Boudet A, Chenine L, Leray H, Rodriguez A, Mourad G, et al. Standardized Method to Measure Muscle Force at the Bedside in Hemodialysis Patients. J Ren Nutr. 2017;27(3):194-200.

65. Lu TW, Chien HL, Chang LY, Hsu HC. Enhancing the examiner's resisting force improves the validity of manual muscle strength measurements: application to knee extensors and flexors. J Strength Cond Res. 2012;26(9):2364-71.

66. Schurr K, Sherrington C, Wallbank G, Pamphlett P, Olivetti L. The minimum sit-to-stand height test: reliability, responsiveness and relationship to leg muscle strength. Clin Rehabil. 2012;26(7):656-63.

67. Kollock R, Van Lunen BL, Ringleb SI, Onate JA. Measures of functional performance and their association with hip and thigh strength. J Athl Train. 2015;50(1):14-22.

68. Arnold CM, Warkentin KD, Chilibeck PD, Magnus CR. The reliability and validity of handheld dynamometry for the measurement of lower-extremity muscle strength in older adults. J Strength Cond Res. 2010;24(3):815-24.

69. Bohannon RW, Kindig J, Sabo G, Duni AE, Cram P. Isometric knee extension force measured using a handheld dynamometer with and without belt-stabilization. Physiother Theory Pract. 2012;28(7):562-8.

70. Azegami M, Ohira M, Miyoshi K, Kobayashi C, Hongo M, Yanagihashi R, et al. Effect of single and multi-joint lower extremity muscle strength on the functional capacity and ADL/IADL status in Japanese community-dwelling older adults. Nurs Health Sci. 2007;9(3):168-76.

71. Kamiya K, Mezzani A, Hotta K, Shimizu R, Kamekawa D, Noda C, et al. Quadriceps isometric strength as a predictor of exercise capacity in coronary artery disease patients. Eur J Prev Cardiol. 2014;21(10):1285-91.

72. Mahony K, Hunt A, Daley D, Sims S, Adams R. Inter-tester reliability and precision of manual muscle testing and hand-held dynamometry in lower limb muscles of children with spina bifida. Phys Occup Ther Pediatr. 2009;29(1):44-59.

73. Ng SS, Ng HH, Chan KM, Lai JC, To AK, Yeung CW. Reliability of the 12-step ascend and descend test and its correlation with motor function in people with chronic stroke. J Rehabil Med. 2013;45(2):123-9.

74. Bohannon RW. Body weight-normalized knee extension strength explains sit-to-stand independence: a validation study. J Strength Cond Res. 2009;23(1):309-11.

75. Freilich RJ, Kirsner RL, Byrne E. Isometric strength and thickness relationships in human quadriceps muscle. Neuromuscul Disord. 1995;5(5):415-22.

76. Messina C, Maffi G, Vitale JA, Ulivieri FM, Guglielmi G, Sconfienza LM. Diagnostic imaging of osteoporosis and sarcopenia: a narrative review. Quant Imaging Med Surg. 2018;8(1):86-99.

77. Franchi MV, Longo S, Mallinson J, Quinlan JI, Taylor T, Greenhaff PL, et al. Muscle thickness correlates to muscle cross-sectional area in the assessment of strength training-induced hypertrophy. Scand J Med Sci Sports. 2018;28(3):846-53.

78. Katari Y, Srinivasan R, Arvind P, Hiremathada S. Point-of-Care Ultrasound to Evaluate Thickness of Rectus Femoris, Vastus Intermedius Muscle, and Fat as an Indicator of Muscle and Fat Wasting in Critically Ill Patients in a Multidisciplinary Intensive Care Unit. Indian J Crit Care Med. 2018;22(11):781-8.

79. Hughes MA, Myers BS, Schenkman ML. The role of strength in rising from a chair in the functionally impaired elderly. J Biomech. 1996;29(12):1509-13.

80. Ploutz-Snyder LL, Manini T, Ploutz-Snyder RJ, Wolf DA. Functionally relevant thresholds of quadriceps femoris strength. J Gerontol A Biol Sci Med Sci. 2002;57(4):B144-52.

81. Baldwin CE, Paratz JD, Bersten AD. Diaphragm and peripheral muscle thickness on ultrasound: intra-rater reliability and variability of a methodology using non-standard recumbent positions. Respirology. 2011;16(7):1136-43.

82. Pardo E, El Behi H, Boizeau P, Verdonk F, Alberti C, Lescot T. Reliability of ultrasound measurements of quadriceps muscle thickness in critically ill patients. BMC Anesthesiol. 2018;18(1):205.

83. Sahathevan S, Khor BH, Yeong CH, Tan TH, Meera Mohaideen AK, Ng HM, et al. Validity of Ultrasound Imaging in Measuring Quadriceps Muscle Thickness and Cross-Sectional Area in Patients Receiving Maintenance Hemodialysis. JPEN J Parenter Enteral Nutr. 2021;45(2):422-6.

84. Vieira L, Rocha LPB, Mathur S, Santana L, Melo PF, Silva VZMD, et al. Reliability of skeletal muscle ultrasound in critically ill trauma patients. Rev Bras Ter Intensiva. 2019;31(4):464-73.

85. Mayer KP, Dhar S, Cassity E, Denham A, England J, Morris PE, et al. Interrater Reliability of Muscle Ultrasonography Image Acquisition by Physical Therapists in Patients Who Have or Who Survived Critical Illness. Phys Ther. 2020;100(9):1701-11.

86. Hadda V, Khilnani GC, Kumar R, Dhunguna A, Mittal S, Khan MA, et al. Intra- and Inter-observer Reliability of Quadriceps Muscle Thickness Measured with Bedside Ultrasonography by Critical Care Physicians. Indian J Crit Care Med. 2017;21(7):448-52.

87. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol. 2008;61(4):344-9.

88. Bohannon RW. Grip Strength: An Indispensable Biomarker For Older Adults. Clin Interv Aging. 2019;14:1681-91.

89. Turan Z, Topaloglu M, Ozyemisci Taskiran O. Medical Research Council-sumscore: a tool for evaluating muscle weakness in patients with post-intensive care syndrome. Crit Care. 2020;24(1):562.

90. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016;15(2):155-63.

91. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet. 2007;370(9596):1453-7.

92. Ishida H, Watanabe S. Influence of inward pressure of the transducer on lateral abdominal muscle thickness during ultrasound imaging. J Orthop Sports Phys Ther. 2012;42(9):815-8.

93. Ishida H, Suehiro T, Suzuki K, Yoneda T, Watanabe S. Influence of the ultrasound transducer tilt on muscle thickness and echo intensity of the rectus femoris muscle of healthy subjects. J Phys Ther Sci. 2017;29(12):2190-3.

94. Schober P, Boer C, Schwarte LA. Correlation Coefficients: Appropriate Use and Interpretation. Anesth Analg. 2018;126(5):1763-8.

95. Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. PLoS Med. 2007;4(10):e297.

96. Pinto-Ramos J, Moreira T, Costa F, Tavares H, Cabral J, Costa-Santos C, et al. Handheld dynamometer reliability to measure knee extension strength in rehabilitation patients-A cross-sectional study. PLoS One. 2022;17(5):e0268254.

97. Pinto-Ramos J, Santos C, Costa F, et-al. Reliability of Point-of-Care Ultrasound for Measuring Quadriceps Femoris Muscle Thickness. European Journal of Physical and Rehabilitation Medicine2022.

98. Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res. 2018;27(5):1147-57.

99. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. J Clin Epidemiol. 2000;53(5):459-68.

100. Weinel LM, Summers MJ, Chapple LA. Ultrasonography to measure quadriceps muscle in critically ill patients: A literature review of reported methodologies. Anaesth Intensive Care. 2019;47(5):423-34.

101. Tourel C, Burnol L, Lanoiselé J, Molliex S, Viallon M, Croisille P, et al. Reliability of standardized ultrasound measurements of quadriceps muscle thickness in

neurological critically ill patients: a comparison to computed tomography measures. J Rehabil Med. 2020;52(3):jrm00032.

102. Barotsis N, Tsiganos P, Kokkalis Z, Panayiotakis G, Panagiotopoulos E. Reliability of muscle thickness measurements in ultrasonography. Int J Rehabil Res. 2020;43(2):123-8.

103. Parry SM, Denehy L, Beach LJ, Berney S, Williamson HC, Granger CL. Functional outcomes in ICU – what should we be using? – an observational study. Crit Care. 2015;19:127.

104. Parry SM, Huang M, Needham DM. Evaluating physical functioning in critical care: considerations for clinical practice and research. Crit Care. 2017;21(1):249.

105. Stone CA, Nolan B, Lawlor PG, Kenny RA. Hand-held dynamometry: tester strength is paramount, even in frail populations. J Rehabil Med. 2011;43(9):808-11.

106. Wikholm JB, Bohannon RW. Hand-held Dynamometer Measurements: Tester Strength Makes a Difference. J Orthop Sports Phys Ther. 1991;13(4):191-8.

107. McMahon LM, Burdett R, Whitney SL. Effects of muscle group and placement site on reliability of hand-held dynamometry strength measurements. J Orthop Sports Phys Ther. 1992;15(5):236-42.

108. Toemen A, Dalton S, Sandford F. The intra-and inter-rater reliability of manual muscle testing and a hand-held dynamometer for measuring wrist strength in symptomatic and asymptomatic subjects. *Hand Therapy*. 2011;*16* (3):67-74.

109. Fragala MS, Alley DE, Shardell MD, Harris TB, McLean RR, Kiel DP, et al. Comparison of Handgrip and Leg Extension Strength in Predicting Slow Gait Speed in Older Adults. J Am Geriatr Soc. 2016;64(1):144-50.

110. Schmitt LC, Paterno MV, Hewett TE. The impact of quadriceps femoris strength asymmetry on functional performance at return to sport following anterior cruciate ligament reconstruction. J Orthop Sports Phys Ther. 2012;42(9):750-9.

111. Lord SR, Murray SM, Chapman K, Munro B, Tiedemann A. Sit-to-stand performance depends on sensation, speed, balance, and psychological status in addition to strength in older people. J Gerontol A Biol Sci Med Sci. 2002;57(8):M539-43.

112. Jones CJ, Rikli RE, Beam WC. A 30-s chair-stand test as a measure of lower body strength in community-residing older adults. Res Q Exerc Sport. 1999;70(2):113-9.

113. Vanhorebeek I, Latronico N, Van den Berghe G. ICU-acquired weakness. Intensive Care Med. 2020;46(4):637-53.

114. Huang CH, Hsu CC, Yu PC, Peng LN, Lin MH, Chen LK. Hospitalization-associated muscle weakness and functional outcomes among oldest old patients: A hospital-based cohort study. Exp Gerontol. 2021;150:111353.

115. Kumar R, Shah TH, Hadda V, Tiwari P, Mittal S, Madan K, et al. Assessment of quadriceps muscle thickness using bedside ultrasonography by nurses and physicians in the intensive care unit: Intra- and inter-operator agreement. World J Crit Care Med. 2019;8(7):127-34.

116. Gutiérrez M, Di Geso L, Rovisco J, Di Carlo M, Ariani A, Filippucci E, et al. Ultrasound learning curve in gout: a disease-oriented training program. Arthritis Care Res (Hoboken). 2013;65(8):1265-74.

117. Eftedal O, Brubakk AO. Agreement between trained and untrained observers in grading intravascular bubble signals in ultrasonic images. Undersea Hyperb Med. 1997;24(4):293-9.

118. Blehar DJ, Barton B, Gaspari RJ. Learning curves in emergency ultrasound education. Acad Emerg Med. 2015;22(5):574-82.

119. Abe T, Loenneke JP, Thiebaud RS. Morphological and functional relationships with ultrasound measured muscle thickness of the lower extremity: a brief review. Ultrasound. 2015;23(3):166-73.

120. Núñez M, Nuñez E, Moreno JM, Segura V, Lozano L, Maurits NM, et al. Quadriceps muscle characteristics and subcutaneous fat assessed by ultrasound and relationship with function in patients with knee osteoarthritis awaiting knee arthroplasty. J Clin Orthop Trauma. 2019;10(1):102-6.

121. Nishihara K, Kawai H, Kera T, Hirano H, Watanabe Y, Fujiwara Y, et al. Correlation of physical function with the thickness of multiple muscles of the quadriceps femoris in community-dwelling elderly individuals. Clin Interv Aging. 2018;13:1945-51.

122. Yoshiko A, Ogawa M, Shimizu K, Radaelli R, Neske R, Maeda H, et al. Chair sit-to-stand performance is associated with diagnostic features of sarcopenia in older men and women. Arch Gerontol Geriatr. 2021;96:104463.

123. Andrews JS, Trupin L, Schmajuk G, Barton J, Margaretten M, Yazdany J, et al. Muscle strength, muscle mass, and physical disability in women with systemic lupus erythematosus. Arthritis Care Res (Hoboken). 2015;67(1):120-7.

124. Mithal A, Bonjour JP, Boonen S, Burckhardt P, Degens H, El Hajj Fuleihan G, et al. Impact of nutrition on muscle mass, strength, and performance in older adults. Osteoporos Int. 2013;24(5):1555-66.

125. Kara M, Kaymak B, Frontera W, Ata AM, Ricci V, Ekiz T, et al. Diagnosing sarcopenia: Functional perspectives and a new algorithm from the ISarcoPRM. J Rehabil Med. 2021;53(6):jrm00209.

126. Kara M, Kaymak B, Ata AM, Özkal Ö, Kara Ö, Baki A, et al. STAR-Sonographic Thigh Adjustment Ratio: A Golden Formula for the Diagnosis of Sarcopenia. Am J Phys Med Rehabil. 2020;99(10):902-8.

127. Eschalier R, Massoullié G, Boirie Y, Blanquet M, Mulliez A, Tartière PL, et al. Sarcopenia in patients after an episode of acute decompensated heart failure: An underdiagnosed problem with serious impact. Clin Nutr. 2021;40(6):4490-9.

128. Herridge MS, Cheung AM, Tansey CM, Matte-Martyn A, Diaz-Granados N, Al-Saidi F, et al. One-year outcomes in survivors of the acute respiratory distress syndrome. N Engl J Med. 2003;348(8):683-93.

129. Kwon S, Hartzema AG, Duncan PW, Min-Lai S. Disability measures in stroke: relationship among the Barthel Index, the Functional Independence Measure, and the Modified Rankin Scale. Stroke. 2004;35(4):918-23.

7. Appendices

7.1 Outputs of the studies

7.1.1 Peer-reviewed scientific publications

1. Pinto-Ramos J, Moreira T, Costa F, Tavares H, Cabral J, Costa-Santos C, Barroso J, Sousa-Pinto B. Handheld dynamometer reliability to measure knee extension strength in rehabilitation patients - A cross-sectional study. PLoS One. 2022;17(5):e0268254.

2. Pinto-Ramos J, Costa-Santos C, Costa F, Tavares H, Cabral J, Moreira T, Brito R, Barroso J, Sousa-Pinto B. Reliability of Point-of-Care Ultrasound for Measuring Quadriceps Femoris Muscle Thickness. European Journal of Physical and Rehabilitation Medicine. 2022.

3. Pinto-Ramos J, Moreira T, Costa L, Costa F, Barroso J, Sousa-Pinto B. Association between knee extension strength and functional capacity after intensive care unit discharge: A six-month prospective cohort study. American Journal of Physical Medicine and Rehabilitation. 2022

7.1.2 Oral and Poster Presentations

1. Poster: Coeficiente de Correlação entre Dinamómetro de Mão e Isocinético. XLI Curso de Reumatologia - Ciência na Prática 2021; February 2021; Portugal

2. Poster: Relationship Between Muscle Imaging and Measures of Muscle Strenght. Virtual ISPRM 2021 Congress - Furthering Rehabilitation in a New World. June 2021

3. Poster: Relationship Between Dynamometry Measures of Muscle strength and Functional Capacity. Virtual ISPRM 2021 Congress - Furthering Rehabilitation in a New World. June 2021

4. Poster: Validade e Reprodutibilidade da Medição da Espessura do Reto Femoral e Quadricípete Femoral com Ecógrafo. XXI Congresso Nacional SPMFR. October 2021; Portugal

5. Poster: Observadores Experientes e Inexperientes têm a mesma Validade e Reprodutibilidade na Medição de Espessura do Quadricípete Femoral e Reto Femoral num Ecógrafo Portátil. XXI Congresso Nacional SPMFR. October 2021; Portugal

6. Oral Presentation: Correlação entre Espessura Muscular e Força Muscular do Músculo Recto Femoral: Um estudo Transversal. XXI Congresso Nacional SPMFR. October 2021; Portugal

7. Poster: Avaliação da Força Muscular com Dinamómetro de Mão: Um estudo Transversal. 19º Congresso Nacional de Medicina Legal e Ciências Forenses. November 2021; Portugal

8. Oral Presentation: Avaliação da Espessura Muscular com Ecografia Portátil – Aplicação Médico Legal. 19º Congresso Nacional de Medicina Legal e Ciências Forenses. November 2021; Portugal

7.2 Published studies

RESEARCH ARTICLE

Handheld dynamometer reliability to measure knee extension strength in rehabilitation patients—A cross-sectional study

João Pinto-Ramos[1,2]*, **Tiago Moreira**[1], **Frederico Costa**[1], **Helena Tavares**[1], **João Cabral**[1], **Cristina Costa-Santos**[2,3], **Joana Barroso**[1,4,5,6], **Bernardo Sousa-Pinto**[2,3]

1 Department of Physical Medicine and Rehabilitation, Centro Hospitalar Universitário São João, Porto, Portugal, **2** CINTESIS – Center for Health Technologies and Services Research, University of Porto, Porto, Portugal, **3** MEDCIDS – Department of Community Medicine, Information and Health Decision Sciences, Faculty of Medicine, University of Porto, Porto, Portugal, **4** Department of Biomedicine, Faculty of Medicine, University of Porto, Porto, Portugal, **5** i3s – Institute for Health Research and Innovation, University of Porto, Porto, Portugal, **6** Departments of Neuroscience and Physical Medicine and Rehabilitation, Northwestern University, Feinberg School of Medicine, Chicago, Illinois, United States of America

* jap_ramos@hotmail.com

Abstract

Introduction

The Handheld Dynamometer (HHD) has the potential to overcome some of the logistic and economic limitations of isokinetic dynamometers for measuring knee extension muscle strength. However, its reliability has not been fully assessed. The purpose of this study is to measure intra and inter-rater reliability of HHD for knee extension strength in patients receiving rehabilitation treatment, as well as to understand in which conditions is the reliability higher.

Methods

Twenty-nine patients admitted in an inpatient Physical Medicine and Rehabilitation unit were consecutively included in this cross-sectional study. Two experienced and two inexperienced physicians made two assessments of knee extension strength with HHD, separated by three hours. Intraclass Correlation Coefficients (ICC), absolute differences between assessments, and correlations between strength and functional variables were calculated.

Results

Intra and inter-rater ICC were overall high ($\geq$ 0.950 and 0.927, respectively). Higher values were found when average of two measurements were made for estimating intra-rater ICC (ICC = 0.978; 95%CI = 0.969–0.985) but not for inter-rater ICC. ICC were not statistically significantly different when calculated based on measurements performed by inexperienced physicians and experienced ones. There was a moderate correlation between strength and functional variables.

OPEN ACCESS

Citation: Pinto-Ramos J, Moreira T, Costa F, Tavares H, Cabral J, Costa-Santos C, et al. (2022) Handheld dynamometer reliability to measure knee extension strength in rehabilitation patients—A cross-sectional study. PLoS ONE 17(5): e0268254. https://doi.org/10.1371/journal.pone.0268254

Editor: Shazlin Shaharudin, Universiti Sains Malaysia, MALAYSIA

Received: February 20, 2022

Accepted: April 26, 2022

Published: May 17, 2022

Data Availability Statement: Data cannot be shared publicly due to ethical committee restrictions in sharing data from patients. Data are available from corresponding author or his academic department (contact via medcids@med.up.pt) for researchers who meet the criteria for access to confidential data.

Funding: This study was partly funded by the Doctoral Programme in Clinical and Health Services Research (Faculty of Medicine, University of Porto, Porto, Portugal) and by the National

Funds from Fundação para a Ciência e a Tecnologia Instituto Público, within Centro de Investigação em Tecnologias e Serviços de Saúde (CINTESIS), R&D unit (reference UIDB/4255/2020).

Competing interests: The authors have declared that no competing interests exist.

Conclusion

Handheld Dynamometer seems to be a reliable option to measure knee extension muscle strength, particularly when two measurements are performed and their average is reported.

Introduction

Functional disability is a major problem in global population, affecting mobility and independent living. According to the Centers for Disease Control and Prevention, one-quarter of all adults in the United States have some kind of disability; fourteen percent have severe difficulty walking and climbing stairs, and seven percent have difficulty doing daily tasks without assistance [1]. In Portugal, forty percent of population reported some kind of long-term disability [2].

Functional disability is difficult to measure, one of the reasons being its parameters are dependent on patients' experiences and expectations. Moreover, patient reported outcome measures (such as the 36-Item Short-Form Survey) [3] and functional tests (such as the 6 minutes walking test) [4] are time-consuming and its application challenging in most outpatient clinical rehabilitation settings due to lack of an easy clinical measurable outcome that can be obtained along clinical treatment. This prompts the need for methods simultaneously capable to (i) accurately and reliably assess patients' disability (and its evolution over time), and (ii) be easily implemented in the clinical practice, namely in the outpatient setting of rehabilitation programs. Such methods may include the measurement of muscle strength, since a negative correlation between muscle strength, functional disability and activity of daily living dependence has been shown in the past [5–7]. In this context, measuring knee extension strength may be particularly adequate, as it displays a key role for maintenance of functional capacity for activities as walking, sitting, dressing or having a shower [8–10].

The isokinetic dynamometer (ID) is the gold-standard method for evaluating muscle strength in knee extension, but it is difficult to apply in the clinical practice, since it is very expensive, non-portable and requires previous specialized training [11]. On the other hand, the handheld dynamometer (HDD) is a less expensive and portable device that can also be used to measure muscle strength; its use requires less training, potentially rendering it more applicable in clinical practice [11]. However, questions still remain on the accuracy of HDD. In fact, while some studies have shown a high correlation between measurements in ID and HDD [11–13], only few of them calculated sample size or used more than 2 observers [12, 14–19].

The purpose of this work is to measure the intra-observer and inter-observer reliability of HHD for the assessment of knee extension strength. In addition, this study aims to analyse if the reliability of HHD differs according to the experience of observers, the number of measurements, and if there is an association between knee extension strength measured by HHD and functional capacity.

Methods

Study design

Twenty-nine patients were consecutively selected for this cross-sectional study at the Inpatient unit, Physical Medicine and Rehabilitation Department, Centro Hospitalar e Universitário de São João, which is a tertiary care hospital in Northern Portugal. For each patient, four different observers, two previously experienced and two inexperienced in HHD use, made four measurements of knee extension peak force with an HHD. Do to schedule logistics and in order to systematize the measurements, two measurements were taken in early afternoon (measurement A and measurement B–assessment AB), and two three hours after (measurement C and measurement D–assessment CD) (Fig 1). All observers were physicians with at least two years of clinical experience. Intra-rater and inter-rater intraclass Correlation Coefficients (ICC), absolute differences between measurements within each assessment, and correlation with functional variables were calculated. The study was approved by Ethical Committee of the respective hospital (Research Project 289/20). Patients were asked for written consent. This article was written according to STROBE statement guidelines [20].

Setting and participants

We assessed participants which had already been included in a previous study from our research group, with sample size of thirty patients being calculated. (Pinto-Ramos J et al., Submitted) In brief we consecutively included all patients admitted at our Service from March to June 2021 which satisfied defined inclusion and exclusion criteria. Patients aged $\geq$ 18 years-old performing rehabilitation treatment for any cause fulfilled the inclusion criteria. Exclusion criteria were: patients unable to ambulate without assistance before hospitalization, morbid obesity, actual lower limb bone fractures, serious pressure or venous ulcers, cardiorespiratory instability, major psychiatric disorders, severe cognitive impairment or neurologic conditions such as multiple sclerosis, traumatic brain injury or severe stroke. Patients admitted to our acute rehabilitation service were hospitalized due to acute events which have impacted their functional capacity, having been transferred for personalized and intense rehabilitation programs. Each day, patients had a 24-hour rehabilitation nursery care with activities of daily living training sessions, two physical therapy sessions of one and half hours each, one occupational therapy session and, in case of need, psychology and speech therapy.

Variables and measurements

From each patient, knee extension strength was measured four times by each of the four observers with the HDD. That is, the four observers performed two knee extension peak force measurements with HHD (Micro FET®2 HHD; Hoggan Health Industries, Draper, UT, USA) at each patient (measurements A and B, comprising the assessment AB) followed by two other measurements three hours later (measurements C and D, comprising the assessment CD) (Fig 1). The three-hour interval was defined to allow some interval between assessments in order to reduce the risk of memory bias while allowing all assessments to be done in the same day (ensuring clinical stability of the patients during the interval between assessments). Patients were seated in the edge of the stretcher with the feet and hands suspended and knee flexed at 60º. The observer was in squatting position with the back against the wall for support and both arms extended to the patient dominant leg for stabilization, so that the results dependence on patient or observer strength would be minimized. The HHD was placed in the anterior leg of each patient, five centimetres above the distal part of the medial malleolus (Fig 2). Patients were asked to make and maintain maximum knee extension strength for five seconds. One-minute rest was used between measurements to decrease fatigue impact. Results displayed on HHD were registered by an independent observer, so that the patients and observers applying the HHD were blinded. Strength was measured in Newtons (N).

In order to standardize the procedures, every observer was given a thirty-minute training before the beginning of the study on how to perform HHD measurements correctly, even though two of the observers had more than two years of experience using the HHD.

From each patient, information was also collected (by a different researcher) on his/her sex and age as well as height, weight, body mass index, diagnosis, pain on visual analogue scale and current and previous modified Rankin Scale. Functional variables were also collected,

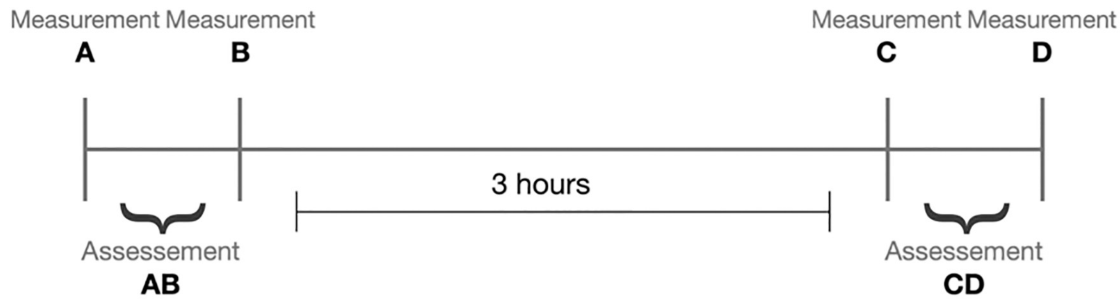


Fig 1. Schematic representation of the measurements performed by each observer in each participant.

https://doi.org/10.1371/journal.pone.0268254.g001

namely handgrip strength, Timed Up and Go (TUG) test, 1-minute sit-to-stand (STS) test and Medical Research Council Sum Score [21–24].

Study size

The sample size for this study was calculated simultaneously with that for the previous study. (Pinto-Ramos J et al., Submitted) Our primary effect size measure was the intra and inter-rater ICC. For us to estimate the sample size, we conducted a comprehensive literature search in order to identify previous ICC estimates. We identified seven studies providing such estimates [12, 14–19], which we pooled by random-effects meta-analysis, resulting in a pooled ICC of

Fig 2. Standardized evaluation method of the patients. Observer was squatting with back against the wall and stretch arms in order to increase stability.

https://doi.org/10.1371/journal.pone.0268254.g002

0.944 (95% confidence interval (95%CI) = [0.902;0.986]) for intra-rater reliability and 0.977 (95%CI = [0.959;0.995]) for inter-rater reliability. If more than one ICC were calculated in the same study (i.e., in left and right limbs), the lower ICC value was used in the meta-analysis. The minimum sample size estimated based on those meta-analytic values using a 95%CI and a semi-width of 5% was of 13 for intra-rater and 4 for inter-rater reliability. Since the required sample size calculated for the previous study was higher (30 patients), we ended up enrolling a larger number of patients.

Statistical analyses

For numerical (continuous) variables, we used means and standard-deviations (SD) for describing variables with normal distribution and medians and 25th and 75th percentiles (P25; P75) for variables with non-normal distribution. Categorical variables were described using absolute and relative frequencies (in percentage). Reliability of HHD measurements was estimated using ICC. For each set of two measurements (i.e., assessment AB, and assessment CD), we registered the average, maximum and first value. Two-way mixed-effects model ICCs were used to calculate intra-rater and inter-rater reliability. ICCs were calculated using single measurements per assessment set for maximum and first values. For intra-rater ICC, we calculated it taking in to account the measurements of all observers, comparing average, maximum and first values of the AB assessments with those of CD assessments. Individual intra-rater ICC were also presented for average, maximum and first values. For inter-rater ICC, we compared average, maximum and first values of AB assessments between observers (in an ancillary analysis, such comparisons were also performed for CD assessments). Experienced and inexperienced observers were also separately analysed for ICC comparison between different levels of training. Reliability was considered poor if ICC$<$0.50, moderate if 0.50$\leq$ICC$<$0.75, good if 0.75$\leq$ICC$\leq$0.90 and excellent if ICC$>$0.90 [25].

Since ICC values are dependent on the heterogeneity of the population [25], absolute differences between average, maximum and first measurements of each assessment were calculated, with differences being tested using the Friedman test. We also calculated relative differences, dividing the aforementioned absolute differences by average knee extension peak force of each participant. The correlation between absolute differences and average knee extension peak force was also calculated to assess if absolute differences between assessments depend on average knee extension peak force. If correlations were found to be sufficiently strong ($r>0.4$ or $r<-0.4$), univariable linear regression models were applied.

We also estimated the correlation between average knee extension peak force of the different participants with other functional variables. Such functional tests/variables include the handgrip strength, TUG test, STS test, Medical Research Council Sum Score, actual and previous Modified Rankin, and POCUS measured rectus and quadriceps femoris muscle thickness. If correlations were found to be sufficiently strong ($r>0.4$ or $r<-0.4$), univariable linear regression models were applied, with knee extension peak force being the independent variable and the dependent variables corresponding to the results of each functional variable.

Correlations were estimated using Spearman correlation coefficients. A p-value inferior to 0.05 was considered statistically significant. The Bonferroni correction was applied in order to control for multiple comparisons. ICC values were calculated using R software and the remaining statistical analyses were done in IBM SPSS Statistics 28 (IBM, Armonk, NY).

Results

Twenty-nine patients were included in this study. Participants' age ranged between 19 and 82 years old with mean age of 58.8 (SD = 14.1) years old; 75.9% of patients were male (Table 1).

Table 1. Demographic characteristics of the sample and descriptions of functional tests.

Variables	Patients (N = 29)
Age (years)–mean (SD)	58.8 (14.1)
Males–*n* (%)	22 (75.9)
BMI (Kg/m^2)–median (P25;P75)	23.1 (21.5;27.4)
Previous mRankin–median (P25;P75)	0 (0;1)
Current mRankin–median (P25;P75)	3 (2;4)
Medical Research Council Sum Score—median (P25;P75)	46 (43;47.5)
Handgrip Strength (kg)–median (P25;P75)	18.3 (11.7;23.3)
Timed Up and Go (s)–mean (SD)[1]	18.9 (11.6)
Sit to Stand test–mean (SD)[2]	17.2 (7.6)

BMI–Body Mass Index, P25;P75–25^{th} and 75^{th} Percentile, Kg–Kilogram, m–Meter, s–Seconds, SD–Standard Deviation;

[1]–Only 16 participants could complete the test,

[2]–Only 10 participants could complete the test

https://doi.org/10.1371/journal.pone.0268254.t001

Participants displayed a median body mass index of 23.1 kg/m^2 (P25;P75 = 21.5;27.4), and the current modified Rankin median score was of 3 (P25;P75 = 2;4). Nineteen patients (65.5%) were not able to stand up from a chair whereas thirteen patients could not ambulate (44.8%). Patients able to ambulate had a mean TUG test of 18.9 seconds (SD = 11.6) and patients able to stand from a sitting position had a mean STS test of 17.2 (SD = 7.6). Median muscle strength was of 189.8 N (P25;P75 = 130.7;274.4).

The most common admission diagnoses were Intensive Care Unit Acquired Weakness (ten patients) and non-severe (Initial National Institute of Health Stroke Scale (NIHSS) <15) Stroke (six patients).

Observers rapidly adapted to the use of HHD and measurements took less than one minute in the majority of times, although positioning the most disabled patients in the stretcher was often a challenge. Also, patients with higher knee extension strength were difficult to resist in order to measure their full strength.

Intra-rater reliability

The intra-rater ICC for knee extension strength of the 464 measurements (corresponding to 4 measurements done by each of the 4 observers to the 29 patients) was excellent with results higher than 0.950 when considering either average, maximum and first measurements. The intra-rater ICC was significantly higher when calculated based on the average of two measurements (average of AB versus average of CD) (ICC = 0.978, 95%CI = 0.969–0.985) than when based on the first measurement alone of each set (A and C) (ICC = 0.950, 95%CI = 0.928–0.965) (Table 2).

Inter-rater reliability

Considering measurements of the assessment AB, we observed an inter-rater ICC between 0.927 and 0.936 of knee extension strength measured with HHD. There were no substantial differences between ICC calculated based on average, maximum and first values of the assessment AB. (Table 2). Similar results were found from assessment CD with ICC between 0.943 and 0.952. (Table 2).

Table 2. Intra-rater and inter-rater intra-class correlation coefficients (ICC) of knee extension strength measured with hand-held dynamometers using average, maximum and first values within AB and CD assessments.

	ICC calculated based on average of measurements within each assessment (95% CI)	ICC calculated based on maximum of measurements within each assessment (95% CI)	ICC calculated based on first measurements within each assessment (95%CI)
Knee Extension Strength intra-rater ICC	0.978 (0.969–0.985)	0.961 (0.945–0.973)	0.950 (0.928–0.965)
Knee Extension Strength inter-rater ICC			
Assessment AB	0.932 (0.864–0.967)	0.936 (0.874–0.969)	0.927 (0.859–0.964)
Assessment CD	0.952 (0.908–0.976)	0.952 (0.911–0.976)	0.943 (0.899–0.971)

95%CI–95% Confidence Interval

Individual intra-rater ICC were similar or higher than 0.948 for the four observers, either using mean, maximum or first values (S1 Table).

https://doi.org/10.1371/journal.pone.0268254.t002

Absolute and relative differences

The median of the absolute differences between knee extension strength assessments (AB vs CD) ranged between 15.0 and 15.4 N (depending on whether average, maximum or first measurements in each assessment were being considered). No significant differences were found between median differences calculated based on average, maximum or first measurements within each assessment ($p = 0.23$). Relative differences between assessments ranged between 8.7% and 10.5%. (Table 3).

Average HHD strength were moderately correlated with absolute differences for average ($r = 0.477$ (95%CI = 0.318;0.610) [$p<0.001$]), maximum ($r = 0.415$ (95%CI = 0.246;0.559) [$p<0.001$]) and first measurements ($r = 0.414$ (95%CI = 0.245;0.558) [$p<0.001$]). Univariable linear regression coefficients were subsequently applied, with coefficients being of 0.112 (95% CI = 0.080;0.144) [$p<0.001$] for average measurements, 0.103 (CI95% = 0.070;0.136) [$p<0.001$] for maximum measurements, and 0.117 (95%CI = 0.080;0.155) [$p<0.001$] for first measurements. On the other hand, correlations between relative differences and average HHD strength were not significant, with Spearman correlation coefficients ranging between -0.063 and -0.109.

Experienced and inexperienced observers

Similar results were observed when comparing the reliability of measurements of experienced versus inexperienced observers. For experienced observers intra-rater ICCs were of 0.981, 0.963, 0.949, depending on whether such ICCs were estimated based on average, maximum or first measurements, respectively. For inexperienced observers, such values were of 0.976, 0.961, 0.950, respectively.

Table 3. Absolute difference (in Newton) of knee extension strength and percentual difference of absolute difference over average muscle strength of all assessments with Hand Held Dynamometer (HHD) using mean, maximum and first measurements between assessments AB and CD for the same observer.

	Average of measurements within each assessment—Median (P25;P75)	Maximum of measurements within each assessment—Median (P25;P75)	First measurement within each assessment—Median (P25;P75)
Absolute Difference (N) Between Knee Extension Strength	15.2 (7.9;34.7)	15.0 (6.2;33.0)	15.4 (8.9;38.0)
Percentual Difference Between Absolute Difference and Average Knee Extension Strength	8.7% (5.1;16.8)	8.8% (4.5;17.4)	10.5% (4.9;18.3)

HHD–Hand Held Dynamometer, P25;P75–25th and 75th Percentiles, N–Newton;

https://doi.org/10.1371/journal.pone.0268254.t003

Muscle strength and functional outcomes

Except for the Previous Modified Rankin, muscle strength measured by knee extension peak force in the HHD was moderately correlated (correlation coefficient ≥0.4) with all functional variables, including the current modified Rankin, the TUG test, the STS test, the Medical Research Council Sum Score, the Handgrip Strength and the Quadriceps and Rectus Femoris Muscle Thickness. (Table 4) Correlations after Bonferroni correction were statistically significant for all functional tests, except the Previous Modified Rankin, the Medical Research Council Sum Score and the TUG. Linear regression coefficients ranged from -0.062 (association between knee extension peak force and TUG) to 0.045 (association between knee extension peak force and STS). (Table 4).

Discussion

Our study shows that the HHD may be a reliable tool for estimating knee extensors muscle strength in rehabilitation patients, with both intra and inter-rater ICC being higher than 0.9 across all observers. Using an average of two measurements increases reliability when compared with estimates based on a single measurement, which suggests a random error in the measurements. The learning curve of HHD use is short, with experienced and inexperienced observers showing similar results.

This study has some limitations worth noting, including the fact that this is a single center study, and caution is required when generalising its results to other populations. Also, as the patients were hospitalized to perform a rehabilitation program in an acute tertiary care Hospital, results may not apply to other type of patients. Patients were observed in two assessments by four observers, so that at each set of measurements the patient was subject to eight maximum knee strength evaluations performed within a short period (being subject to sixteen of such evaluations within a period of three hours). As a result, patient fatigue may have modified the results between measurements and lead to underestimation of results. Authors tried to manage this limitation by setting a minimum one-minute rest between each measurement, so that maximum peak force could be reached every time. An additional limitation may concern the possibility of memory biases (given the three-hour gap between the two sets of

Table 4. Correlation and linear regression coefficient between Handheld Dynamometer and different functional variables.

Functional Variables	Correlation coefficient (95%CI) [p-value]	Linear regression coefficient (95%CI) [p value]
Previous Modified Rankin	-0.378 (-0.660; -0.002) [0.043]*	**
Current Modified Rankin	-0.565 (-0.776; -0.239) [0.001]	-0,005 (-0,008; -0.002) [0.004]
TUG test	-0.612 (-0.854; -0.151) [0.012]*	-0,062 (-0.106; -0.019) [0.009]
STS test	0.499 (0.151; 0.737) [0.006]	0.045 (0.018; 0.073) [0.002]
Medical Research Council Sum Score	0.484 (0.132; 0.728) [0.008]*	0.019 (-0.005; 0.043) [0.109]
Handgrip Strength	0.545 (0.213; 0.765) [0.002]	0.030 (0.004; 0.055) [0.023]
Quadriceps Femoris Muscle Thickness	0.511 (0.167; 0.744) [0.005]	0.002 (0.000; 0.003) [0.012]
Rectus Femoris Muscle Thickness	0.536 (0.201; 0.759) [0.003]	0.002 (0.000; 0.005) [0.027]

95%CI–95% Confidence Interval, TUG–Timed Up and Go; STS–1-minute sit-to-stand;

*non-significant after Bonferroni correction;

** Not applicable since r<0.40

https://doi.org/10.1371/journal.pone.0268254.t004

measurements). This may particularly concern the possibility of psychomotor memory bias from patients. On the other hand, we believe that memory bias from the observers may not have had a relevant impact, as each day there were seven to nine patients to be assessed by each independent observer (resulting in at least 56 measurements to remember before assessment CD). Finally, since ICC values are contingent on the homogeneity of the population, results might be overestimated as a result of the heterogeneity of studied population [25]. To overcome this limitation, and in order to estimate the actual differences, absolute and percentual differences between assessments were calculated.

This work has also important strengths. Firstly, this study was designed using more than two observers, differentiating it from most other studies assessing the reliability of the HHD. Also, this study compared different ways of calculating the strength of the patient knee extension, namely using two measurements or a single measurement, in order to estimate the most reliable assessment of muscle strength. Having four observers also allowed us to include observers with different experience levels, making us able to assess whether experience associated with different results. Finally, we ensured that this study was adequately powered to obtain precise estimates.

The HHD was found to be a reliable way of measuring muscle strength, with results showing excellent reliability, in accordance with what has been previously reported in previous studies assessing other types of patients [12, 14–19].

Absolute differences between assessments were relatively low, despite being higher for patients with higher muscle strength. By contrast, the percentual differences did not vary with increase of muscle strength. This might occur because knee extension of stronger patients is possibly more difficult to resist, leading to higher errors, specially between observers with less capability to resist knee extension strength. In fact, other studies already suggested that stronger observers tend to report higher values on HHD and also that external stabilization of the HHD, namely with a belt, can be more reliable than human-resisted evaluations [19, 26–28]. Nevertheless, belt stabilization is more difficult and time-consuming, possibly rendering it impractical to apply in clinical practice.

Results were similar when considering measurements performed by both experienced and previously inexperienced observers, which suggest that the use of HHD to measure muscle strength could be generalized across physicians if a short course of training is previously done. Since this study was conducted throughout a period of four months, we cannot estimate if reliability would be this high for longer periods if observers were not trained again for standardized evaluations. Other studies with inexperienced subjects with a short course training already showed that good reliabilities could be achieved with HDD but in different muscular groups [29, 30].

Knee extension strength on HHD was moderately correlated with the majority of analysed functional variables, confirming that muscle strength, including for knee extension, may be important for these tasks, as it was also suggested by other studies [31–34]. In these participants, correlations between functional tests and muscle strength were higher than those observed between functional tests and muscle thickness (except for hand grip strength). (Pinto-Ramos J et al., Submitted) This suggests that, compared with muscle thickness, muscle strength correlates better with functional variables for this specific population, which can be justified by many factors, such as muscular or polyneuropathic changes due to hospitalization, which can affect differently thickness and strength [35, 36].

The HHD overcomes many of the limitations associated with the use of isokinetic dynamometers (considering the gold-standard for the assessment of strength). In fact, while the latter require patients to move to a different location to be tested, well-trained users and expensive initial investment, HHD is portable, fast to use (with measurements done in less than two minutes during consultation/medical evaluation) and displays lower costs. This may propel the measurement of muscle strength in clinical practice which has been very limited so far. However, this requires not only that HHD are demonstrated to be reliable tools, but also that they are useful to estimate patients' functional level and evolution over time (which requires prospective studies to confirm this theory).

In conclusion, knee extension strength measured with HHD seems to be a reliable resource to objectively measure strength on patients in rehabilitation programs. This method seems feasible to apply, even for physicians with no previous contact with the HHD. Some questions are still to answer, including if method is adequate to objectively measure patients' functional evolution over time.

Supporting information

S1 Table. Individual intra-rater intra-class correlation coefficients (ICC) of knee extension strength measured with hand-held dynamometers using average, maximum and first values within AB and CD assessments for all the observers.
(DOCX)

Author Contributions

Conceptualization: João Pinto-Ramos, Bernardo Sousa-Pinto.

Data curation: João Pinto-Ramos, Tiago Moreira, Frederico Costa, Helena Tavares, João Cabral.

Formal analysis: João Pinto-Ramos, Cristina Costa-Santos, Joana Barroso, Bernardo Sousa-Pinto.

Methodology: João Pinto-Ramos, Cristina Costa-Santos, Joana Barroso, Bernardo Sousa-Pinto.

Supervision: Bernardo Sousa-Pinto.

Writing – original draft: João Pinto-Ramos.

Writing – review & editing: Cristina Costa-Santos, Joana Barroso, Bernardo Sousa-Pinto.

References

1. https://www.cdc.gov/ncbddd/disabilityandhealth/infographic-disability-impacts-all.html.
2. Raposo Dos Santos JR, Cascão Torres AR, Neto M, Namorado S. Prevalence of long-term disability in Portugal, 2014: Evidence of variation by personal and contextual factors. Disabil Health J. 2021; 14 (2):101043. Epub 20201119. https://doi.org/10.1016/j.dhjo.2020.101043 PMID: 33250390.
3. Hayes V, Morris J, Wolfe C, Morgan M. The SF-36 health survey questionnaire: is it suitable for use with older adults? Age Ageing. 1995; 24(2):120–5. https://doi.org/10.1093/ageing/24.2.120 PMID: 7793333.
4. Enright PL. The six-minute walk test. Respir Care. 2003; 48(8):783–5. PMID: 12890299.
5. Rantanen T. Muscle strength, disability and mortality. Scand J Med Sci Sports. 2003; 13(1):3–8. https://doi.org/10.1034/j.1600-0838.2003.00298.x PMID: 12535311.
6. Wang DXM, Yao J, Zirek Y, Reijnierse EM, Maier AB. Muscle mass, strength, and physical performance predicting activities of daily living: a meta-analysis. J Cachexia Sarcopenia Muscle. 2020; 11(1):3–25. Epub 20191201. https://doi.org/10.1002/jcsm.12502 PMID: 31788969.
7. Jung H, Tanaka S, Iwamoto Y, Kawano T, Yamasaki M, Tanaka R. Reductions in Muscle Strength and Range of Motion Cause Locomotion Disability via Locomotion-Related Functional Limitation in Japanese Older Adults: A Cross-Sectional Study. J Aging Res. 2021; 2021:6627767. Epub 20210707. https://doi.org/10.1155/2021/6627767 PMID: 34306754.

8. Arnold AS, Anderson FC, Pandy MG, Delp SL. Muscular contributions to hip and knee extension during the single limb stance phase of normal gait: a framework for investigating the causes of crouch gait. J Biomech. 2005; 38(11):2181–9. Epub 20041123. https://doi.org/10.1016/j.jbiomech.2004.09.036 PMID: 16154404.

9. Corrigan D, Bohannon RW. Relationship between knee extension force and stand-up performance in community-dwelling elderly women. Arch Phys Med Rehabil. 2001; 82(12):1666–72. https://doi.org/10.1053/apmr.2001.26811 PMID: 11733880.

10. Kojima N, Kim H, Saito K, Yoshida H, Yoshida Y, Hirano H, et al. Association of knee-extension strength with instrumental activities of daily living in community-dwelling older adults. Geriatr Gerontol Int. 2014; 14(3):674–80. Epub 20131112. https://doi.org/10.1111/ggi.12158 PMID: 24215603.

11. Stark T, Walker B, Phillips JK, Fejer R, Beck R. Hand-held dynamometry correlation with the gold standard isokinetic dynamometry: a systematic review. PM R. 2011; 3(5):472–9. https://doi.org/10.1016/j.pmrj.2010.10.025 PMID: 21570036.

12. Sung KS, Yi YG, Shin HI. Reliability and validity of knee extensor strength measurements using a portable dynamometer anchoring system in a supine position. BMC Musculoskelet Disord. 2019; 20(1):320. Epub 20190708. https://doi.org/10.1186/s12891-019-2703-0 PMID: 31286912.

13. Kim WK, Kim DK, Seo KM, Kang SH. Reliability and validity of isometric knee extensor strength test with hand-held dynamometer depending on its fixation: a pilot study. Ann Rehabil Med. 2014; 38(1):84–93. Epub 20140225. PMID: 24639931.

14. Martins J, da Silva JR, da Silva MRB, Bevilaqua-Grossi D. Reliability and Validity of the Belt-Stabilized Handheld Dynamometer in Hip- and Knee-Strength Tests. J Athl Train. 2017; 52(9):809–19. Epub 20170808. https://doi.org/10.4085/1062-6050-52.6.04 PMID: 28787180.

15. Mentiplay BF, Perraton LG, Bower KJ, Adair B, Pua YH, Williams GP, et al. Assessment of Lower Limb Muscle Strength and Power Using Hand-Held and Fixed Dynamometry: A Reliability and Validity Study. PLoS One. 2015; 10(10):e0140822. Epub 20151028. https://doi.org/10.1371/journal.pone.0140822 PMID: 26509265.

16. Romero-Franco N, Jiménez-Reyes P, Montaño-Munuera JA. Validity and reliability of a low-cost digital dynamometer for measuring isometric strength of lower limb. J Sports Sci. 2017; 35(22):2179–84. Epub 20161124. https://doi.org/10.1080/02640414.2016.1260152 PMID: 27882825.

17. Whiteley R, Jacobsen P, Prior S, Skazalski C, Otten R, Johnson A. Correlation of isokinetic and novel hand-held dynamometry measures of knee flexion and extension strength testing. J Sci Med Sport. 2012; 15(5):444–50. Epub 20120315. https://doi.org/10.1016/j.jsams.2012.01.003 PMID: 22424705.

18. Toonstra J, Mattacola CG. Test-retest reliability and validity of isometric knee-flexion and -extension measurement using 3 methods of assessing muscle strength. J Sport Rehabil. 2013; 22(1). Epub 20130218. https://doi.org/10.1123/jsr.2013.TR7 PMID: 22951307.

19. Florencio LL, Martins J, da Silva MRB, da Silva JR, Bellizzi GL, Bevilaqua-Grossi D. Knee and hip strength measurements obtained by a hand-held dynamometer stabilized by a belt and an examiner demonstrate parallel reliability but not agreement. Phys Ther Sport. 2019; 38:115–22. Epub 20190425. https://doi.org/10.1016/j.ptsp.2019.04.011 PMID: 31091492.

20. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol. 2008; 61(4):344–9. https://doi.org/10.1016/j.jclinepi.2007.11.008 PMID: 18313558.

21. Bohannon RW. Grip Strength: An Indispensable Biomarker For Older Adults. Clin Interv Aging. 2019; 14:1681–91. Epub 20191001. https://doi.org/10.2147/CIA.S194543 PMID: 31631989.

22. Kear BM, Guck TP, McGaha AL. Timed Up and Go (TUG) Test: Normative Reference Values for Ages 20 to 59 Years and Relationships With Physical and Mental Health Risk Factors. J Prim Care Community Health. 2017; 8(1):9–13. Epub 20160725. https://doi.org/10.1177/2150131916659282 PMID: 27450179.

23. Bohannon RW, Crouch R. 1-Minute Sit-to-Stand Test: SYSTEMATIC REVIEW OF PROCEDURES, PERFORMANCE, AND CLINIMETRIC PROPERTIES. J Cardiopulm Rehabil Prev. 2019; 39(1):2–8. https://doi.org/10.1097/HCR.0000000000000336 PMID: 30489442.

24. Turan Z, Topaloglu M, Ozyemisci Taskiran O. Medical Research Council-sumscore: a tool for evaluating muscle weakness in patients with post-intensive care syndrome. Crit Care. 2020; 24(1):562. Epub 20200918. https://doi.org/10.1186/s13054-020-03282-x PMID: 32948221.

25. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016; 15(2):155–63. Epub 20160331. https://doi.org/10.1016/j.jcm.2016.02.012 PMID: 27330520.

26. Stone CA, Nolan B, Lawlor PG, Kenny RA. Hand-held dynamometry: tester strength is paramount, even in frail populations. J Rehabil Med. 2011; 43(9):808–11. https://doi.org/10.2340/16501977-0860 PMID: 21826388.

27. Wikholm JB, Bohannon RW. Hand-held Dynamometer Measurements: Tester Strength Makes a Difference. J Orthop Sports Phys Ther. 1991; 13(4):191–8. https://doi.org/10.2519/jospt.1991.13.4.191 PMID: 18796845.

28. Bohannon RW, Kindig J, Sabo G, Duni AE, Cram P. Isometric knee extension force measured using a handheld dynamometer with and without belt-stabilization. Physiother Theory Pract. 2012; 28(7):562–8. Epub 20111222. https://doi.org/10.3109/09593985.2011.640385 PMID: 22191509.

29. McMahon LM, Burdett R, Whitney SL. Effects of muscle group and placement site on reliability of hand-held dynamometry strength measurements. J Orthop Sports Phys Ther. 1992; 15(5):236–42. https://doi.org/10.2519/jospt.1992.15.5.236 PMID: 18796778.

30. Toemen A, Dalton S, Sandford F. The intra-and inter-rater reliability of manual muscle testing and a hand-held dynamometer for measuring wrist strength in symptomatic and asymptomatic subjects. *Hand Therapy*. 2011; 16 (3):67–74.

31. Fragala MS, Alley DE, Shardell MD, Harris TB, McLean RR, Kiel DP, et al. Comparison of Handgrip and Leg Extension Strength in Predicting Slow Gait Speed in Older Adults. J Am Geriatr Soc. 2016; 64 (1):144–50. https://doi.org/10.1111/jgs.13871 PMID: 26782864.

32. Schmitt LC, Paterno MV, Hewett TE. The impact of quadriceps femoris strength asymmetry on functional performance at return to sport following anterior cruciate ligament reconstruction. J Orthop Sports Phys Ther. 2012; 42(9):750–9. Epub 20120719. https://doi.org/10.2519/jospt.2012.4194 PMID: 22813542.

33. Lord SR, Murray SM, Chapman K, Munro B, Tiedemann A. Sit-to-stand performance depends on sensation, speed, balance, and psychological status in addition to strength in older people. J Gerontol A Biol Sci Med Sci. 2002; 57(8):M539–43. https://doi.org/10.1093/gerona/57.8.m539 PMID: 12145369.

34. Jones CJ, Rikli RE, Beam WC. A 30-s chair-stand test as a measure of lower body strength in community-residing older adults. Res Q Exerc Sport. 1999; 70(2):113–9. https://doi.org/10.1080/02701367.1999.10608028 PMID: 10380242.

35. Vanhorebeek I, Latronico N, Van den Berghe G. ICU-acquired weakness. Intensive Care Med. 2020; 46(4):637–53. Epub 20200219. https://doi.org/10.1007/s00134-020-05944-4 PMID: 32076765.

36. Huang CH, Hsu CC, Yu PC, Peng LN, Lin MH, Chen LK. Hospitalization-associated muscle weakness and functional outcomes among oldest old patients: A hospital-based cohort study. Exp Gerontol. 2021; 150:111353. Epub 20210420. https://doi.org/10.1016/j.exger.2021.111353 PMID: 33892132.

European Journal of Physical and Rehabilitation Medicine
EDIZIONI MINERVA MEDICA

Reliability of Point-of-Care Ultrasound for Measuring Quadriceps Femoris Muscle Thickness

João PINTO-RAMOS, Cristina SANTOS, Frederico COSTA, Helena TAVARES, João CABRAL, Tiago MOREIRA, Rui BRITO, Joana BARROSO, Bernardo SOUSA-PINTO

European Journal of Physical and Rehabilitation Medicine2022 Sep 01
DOI: 10.23736/S1973-9087.22.07432-9

Article type: Original Article

Supplementary material available online at http://www.minervamedica.it
Article first published online: September 1, 2022
Manuscript accepted: August 24, 2022
Manuscript revised: April 27, 2022
Manuscript received: January 11, 2022

Reliability of Point-of-Care Ultrasound for Measuring Quadriceps Femoris Muscle Thickness

Point-of-Care Ultrasound for Quadriceps

João Pinto-Ramos, MD* [1,2], Cristina Costa-Santos, MD, PhD [2,3], Frederico Costa, MD [1], Helena Tavares, MD [1], João Cabral, MD [1], Tiago Moreira, MD [1], Rui Brito, MD [4], Joana Barroso, MD, PhD [1,5,6,7], Bernardo Sousa-Pinto, MD, PhD [2,3]

1 – Department of Physical Medicine and Rehabilitation, Centro Hospitalar Universitário São João, Porto, Portugal

2 – CINTESIS – Center for Health Technologies and Services Research, University of Porto, Porto, Portugal

3 – MEDCIDS – Department of Community Medicine, Information and Health Decision Sciences, Faculty of Medicine, University of Porto, Porto, Portugal

4 – Department of Physical Medicine and Rehabilitation, Centro Hospitalar do Porto, Porto, Portugal

5 – Department of Biomedicine, Faculty of Medicine, University of Porto, Porto, Portugal

6 – i3s – Institute for Health Research and Innovation, University of Porto, Porto, Portugal

7 – Departments of Neuroscience and Physical Medicine and Rehabilitation, Northwestern University, Feinberg School of Medicine, Chicago, IL, United States of America

Correspondence to: João Pinto-Ramos, Physical Medicine and Rehabilitation Service, Centro Hospitalar Universitário São João, Alameda Prof. Hernâni Monteiro, Porto, Portugal | E-mail: jap_ramos@hotmail.com | Telephone: +351 225 513 622

Author Disclosures: No conflicts of interest related to the manuscript have been reported by the authors. No funding or equipment was provided for the project from any source. There are no financial benefits to the authors. This manuscript has not been previously submitted to any other publication. Authorship has been granted only to those individuals who have contributed substantially to the manuscript. All authors have read and approved the manuscript.

ABSTRACT

BACKGROUND: Point-of-care ultrasound can be used to assess muscle thickness. However, its reliability has not been fully evaluated.

AIM: This study aims to assess the intra-rater and inter-rater reliability of point-of-care ultrasound for the estimation of quadriceps and rectus femoris thickness in patients from a rehabilitation setting.

DESIGN: Cross-sectional study.

SETTING: Physical Medicine and Rehabilitation Department of a tertiary care hospital.

POPULATION: Twenty-nine inpatients consecutively selected after admission.

METHODS: Four observers, two trained and two untrained, used point-of-care ultrasound to measure quadriceps femoris and rectus femoris thickness. Each observer performed two measurements followed by a second set of two measurements three hours later. Intraclass correlation coefficients (ICC) were then calculated.

RESULTS: Both intra-rater and inter-rater ICC were higher than 0.888 for both quadriceps and rectus femoris measurements. Reliability was highest when ICC were calculated based on the average of two measurements, with the intra-rater ICC being of 0.956 (95%CI=0.937-0.970) for rectus femoris and of 0.966 (95%CI=0.951-0.976) for quadriceps femoris, and with the inter-rater ICC being of 0.919 (95%CI=0.863-0.957) for rectus femoris and 0.945 (95%CI=0.907- 0.971) for quadriceps femoris. Trained and untrained observers did not have significantly different ICC values.

CONCLUSION: These results suggest that point-of-care ultrasound is a reliable option to measure muscle thickness of knee extensors by the same or different observers.

CLINICAL REHABILITATION IMPACT: Measuring knee extensors thickness may aid to adequately modulate treatment choices in patients with disability. This study suggests that quadriceps and rectus femoris muscle thickness measured after a short training course, by either an experienced or inexperienced clinician, presents high reliability. Reliability can be increased if the average of two measurements is used. Besides being inexpensive and portable, point-of-care ultrasound is a reliable tool for measuring knee extensors' thickness, rendering it potentially adequate to be used in clinical practice.

KEY WORDS: Quadriceps Muscle; Validation Study; Ultrasonography

INTRODUCTION

Disability is defined by the World Health Organization as an "umbrella term for impairments, activity limitations and participation restrictions".[6] Overcoming disability is the main purpose of Physical Medicine and Rehabilitation, which acts, among others, to promote normal interaction between the musculoskeletal and neurological systems.[7] However, to be able to provide the most adequate therapeutic options (including on nutritional support and training parameters) and to modify them according to patients' progression over time, disability biomarkers are needed. Muscle thickness has been hypothesized as one of such biomarkers – for example, in knee extensors (which play a key role in important functional tasks such as walking, standing up from a sitting position and climbing stairs[8, 9]) muscle thickness appears to have a correlation with muscle strength.[10, 11] Therefore, measuring knee extensors thickness is conceivable as a simple, easy to obtain and reliable metric, which may aid to adequately modulate treatment choices in patients with disability, while also being useful as diagnostic criteria for conditions such as sarcopenia.[1, 2] Sarcopenia is one of the most common conditions leading to disability in the elderly; it is defined as "age-related loss of muscle mass and muscle function", primarily targeting the anterior tight muscles.[2]

Muscle mass and thickness are usually measured using Dual-energy X-ray absorptiometry and magnetic resonance imaging.[11-13] Nevertheless, these techniques are not easily accessible during outpatient assessments, which makes them difficult to apply at point-of-care. In addition, magnetic resonance imaging is an expensive procedure. By contrast, point-of-care ultrasound (POCUS) is easier to apply in an outpatient setting and is becoming more affordable, rendering it a potentially useful tool to measure muscle thickness.[14] On the other hand, there are still several unanswered questions regarding the reliability of this method, since the majority of studies done with either ultrasound (US) or POCUS included only 2 observers, did not calculate sample size and assessed patients outside of the rehabilitation setting.[15-19]

Therefore, the purpose of this study is to assess the intra-rater and inter-rater reliability of POCUS for assessing muscle thickness in a rehabilitation patient population. In addition, this study aims to assess if reliability increases when measurements are performed by trained physicians, and if there is an association between muscle thickness and disability.

METHODS

Study Design

This single-centre cross-sectional study was conducted at the Physical Medicine and Rehabilitation Department of Centro Hospitalar e Universitário de São João, a tertiary care hospital in Northern Portugal. Twenty-nine patients were assessed by four different observers using a portable ultrasound to evaluate muscle thickness of rectus femoris and quadriceps femoris. Each observer performed four measurements on each patient, two in the early afternoon (assessments A and B) and two in the late afternoon (3 hours after the first measurements; assessments C and D) (Figure 1). Intra and Inter-rater intraclass correlation coefficients (ICC) were calculated, as well as absolute differences between assessment AB and CD values. The study was approved by the institution's Ethical committee and written consent was obtained from each patient. This article was written according to the STROBE statement for cross-sectional studies.[20]

Setting and Participants

Patients admitted to the Physical Medicine and Rehabilitation Department of Centro Hospitalar e Universitário de São João from March to June 2021 were consecutively included in the study if they fulfilled all the eligibility criteria. To fulfil the inclusion criteria, patients ought to be aged ≥ 18 years old and to have performed rehabilitation treatment for any cause. We excluded patients who were unable to walk independently before starting rehabilitation treatment, as well as patients with severe cognitive impairment, neurological/neuromuscular conditions (such as multiple sclerosis, traumatic brain injury or amyotrophic lateral sclerosis), major psychiatric disorders, traumatic bone fractures of the lower limb, morbid obesity (body mass index ≥ 35), serious pressure or venous ulcers, or cardiorespiratory instability.

Variables and measurements

For all participants we assessed the current and previous to admission modified Rankin (mRankin) scale, the current handgrip strength, and current pain intensity (by means of visual analogue scale). Handgrip strength was measured using Camry® Electronic Hand Dynamometer (Camry Scale, USA).
Patients were asked to sit with their elbows flexed at 90º and to produce and maintain maximum handgrip strength for five seconds with their dominant hand. The maximum values of each pair of attempts were recorded. In addition, we performed Timed Up and Go (TUG) and 1-minute sit-to-stand (STS) tests, Medical Research Council Sum Score and Knee Extension Peak force measured with a Hand Held Dynamometer. Participants’ demographic characteristics (sex and age), height, weight, body mass index and diagnosis at admission were also retrieved.
For each patient four different observers performed two measurements of rectus femoris and quadriceps femoris thickness in the early afternoon (assessments A and B) and two measurements

of the same muscles' thickness 3 hours after (assessments C and D) (Figure 1). Measurements were performed using a handheld portable POCUS (Butterfly iQ+, Butterfly Network, Inc., Guilford, CT) configured for musculoskeletal image acquisition. (Figure 2) The probe was applied 15 cm above the superior pole of the patella of the dominant leg while the patients were in supine position with both legs stretched. Minimal hand force was applied at the ultrasound probe in order to avoid the effect of pressure on muscle thickness.[21] The probe was positioned perpendicular to muscle fibers in order to avoid increases in muscle thickness due to angulation.[22] Maximal distances between the cortex of femur and the most superficial muscular fascia were measured.

Two of the four observers (observer 1 and 2) had previously received formal training and had already been familiar with musculoskeletal ultrasound imaging for at least two years; nonetheless one hour of training was provided for all observers. The four observers were blinded for each other's results, and a fifth researcher statistically analysed the data.

Study size

This study primary effect size measure consisted of the ICC for intra-rater and inter-rater reliability. In order to estimate the adequate sample size, we performed a comprehensive literature search, followed by random-effects meta-analysis of ICC values published in identified studies.[15-19] The estimated meta-analytical intra-rater ICC value for quadriceps femoris muscle thickness POCUS measurements was 0.91 (95% confidence interval [CI]=0.82;1.00), while the meta-analytical inter-rater ICC was of 0.95 (95%CI=0.89;1.00). Using the meta-analytical value, and in order to obtain 95%CI with a semi-width of 5%, with a power of 80%, and considering four observers we estimated a minimum sample size of thirty patients for intra-rater reliability ICC and ten for inter-rater reliability ICC. Thirty-five patients were selected although we were only able to assess twenty-nine patients, due to loss of follow-up.

Statistical analyses

Means and standard-deviations (SD) were used to describe variables with normal distribution, and medians and 25th and 75th percentile (P25-P75) were used to describe asymmetrical variables.

For each set of two assessments (i.e., assessments AB, and assessments CD), we registered the average, maximum, minimum and first value. Two-way model intraclass correlation coefficients were then calculated in order to estimate the reliability between average, maximum, minimum and first AB versus CD assessments from the same observer (intra-rater ICC). In addition, ICC were calculated to estimate the reliability among average, maximum, minimum and first AB measurements of the four different observers (inter-rater ICC).[23] A subanalysis was performed with ICC values being separately estimated for trained and untrained observers. The ICC was considered poor for values < 0.50, moderate for values between 0.50–0.75, good for values between 0.75–0.90, and excellent for values >0.90.[23]

We calculated absolute and relative percent differences (absolute differences divided by the average value of the four assessments from the four observers) between the average, maximum, minimum and first measurements of AB assessments versus those of CD assessments. Spearman correlation coefficients were calculated to assess the correlation between computed differences and average muscle thickness, so as to assess if increased thickness associated with higher absolute differences.

Finally, as a post hoc analysis, we computed the Spearman correlation coefficient between average muscle thickness of rectus femoris and quadriceps femoris from all observers and functional variables measured by a fifth independent researcher (namely current and previous mRankin scale, handgrip strength, Medical Research Council Sum Score, Timed Up and Go and Sit to Stand tests) and Knee Extension Peak force measured with an Hand Held Dynamometer.[24]

Statistical significance was considered at p-value <0.05. All analyses were performed using IBM SPSS Statistics 27 software and the R software.

RESULTS

A total of 29 patients were included in the study, of whom 22 (75.9%) were male. (Table 1) Participants' mean age was of 58.8 years old (SD=14.1). The median mRankin scale was 0 (P25-P75 0-1) before hospitalization, and of 3 (P25-P75 2-4) at time of observation. The most common admission diagnosis was intensive care unit acquired weakness (ICUAW) (10 patients; 34.5%), followed by ischemic stroke (6 patients; 20.7%). At the time of assessment, thirteen (44.8%) patients were unable to walk, and 19 (65.5%) patients were not capable of standing up from a sitting position.

Intra-rater ICC

In all assessments, intra-rater ICCs for rectus femoris displayed values over 0.90 (Table 2). Intraclass correlation coefficients were significantly higher when calculated based on the average measurements of AB and CD (ICC=0.956; 95%CI=0.937-0.970) than when calculated based on maximum, minimum or first measurements. (Table 2)

Similar results were observed for quadriceps femoris, with all ICC values being superior to 0.90 (Table 2). Intraclass correlation coefficients calculated based on the average values (ICC=0.966; 95%CI=0.951–0.976) were significantly higher than those estimated based on singular measurements (except for maximum values). All ICC values for quadriceps femoris were higher than those for rectus femoris measurements, but differences were not statistically significant. (Table 2)

Inter-rater ICC

Inter-rater ICC for recuts femoris assessments was higher when estimated based on average measurements (ICC=0.919; 95%CI=0.963–0.957) than single measurements, although differences were not statistically significant. (Table 2) Similar results were observed for ICC for quadriceps femoris, with estimates based on average measurements (ICC=0.945; 95%CI=0.907–0.971) having a non-significantly higher ICC value than when based on single measurements. All ICC values were found to be over 0.90. All quadriceps femoris ICC values were higher than those for rectus femoris, although such differences were not significant (Table 2). Similar results were observed when calculating ICC based on CD values. (table e1)

Muscle Thickness Absolute Difference

Median absolute differences in measurements of rectus femoris muscle thickness from the same observer ranged between 0.9 and 1.2 mm (corresponding to an average percent difference of 5.1-9.2%). (Tables 3-4) Differences in muscle thickness measurements were not correlated with average muscle thickness with all correlations being inferior to 0.037.

Median values of the absolute differences between the same observer muscle thickness measurements in different assessments were higher for the quadriceps femoris than for rectus femoris, with results ranging from 1.4 to 1.5 mm (corresponding to an average percent difference of 5.9-6.8%). (Tables 3-

4) Differences in muscle thickness were weakly correlated with average muscle thickness with all correlations showing values inferior to 0.210.

Trained vs Untrained Observers

Differences were not statistically significant between ICC values of observers already trained for the use of POCUS and observers with no previous experience, either for rectus or quadriceps femoris thickness assessments.

Correlation between Muscle Thickness and Other Variables

We observed that rectus femoris thickness was significantly correlated with Knee Extension Peak Force measured with an HHD (r=0.536; p=0.003) and handgrip strength (r=0.574; p<0.001), but not for current and previous mRankin (r=-0.325; p= 0.086 and r=0.063; p= 0.746, respectively), Medical Research Council Sum Score (r=0.070; p=0.718), Timed up Go test on patients able to walk (r=-0.297; p=0.264), or Sit to Stand test on patients able to stand up from a sitting position (r=0.433; p=0.211).

We observed a moderate correlation between quadriceps femoris muscle thickness and current mRankin score (r=-0.376; p=0.044), Knee Extension Peak Force measured with an HHD (r=0.511; p=0.005), and handgrip strength (r=0.588; p<0.001). Correlations were weaker and not significant for other functional variables such as previous mRankin (r=0.141; p=0.466) and Medical Research Council Sum Score (r=0.064; p=0.742). Timed Up Go (n=16; r=-0.247; p=0.356) and Sit to Stand test (n=10; r=0.482; p=0.159) had a poor correlation for patients that could perform them.

DISCUSSION

In this study, we observed that POCUS is a reliable tool to measure muscle thickness in rehabilitation patients. Both intra-class and inter-class ICC displayed overall very high values across the four observers, which suggests that POCUS is a reliable method to use in the clinical setting. Performing two measurements for each assessment (with subsequent computation of the average value) appears to result in increased reliability, suggesting a random error in the measurements.

This work has some limitations worth noting. In fact, this is a single centre study, so caution must be used when applying these results to different populations. In addition, the fact that this study was performed in rehabilitation patients admitted to an acute tertiary care Hospital can limit the generalisability of our results. Furthermore, for simplicity purposes, the probe was applied 15 centimetres above the superior pole of the patella, which may not correspond to the thickest part of the muscle. Nevertheless, other previous studies took the same approach as we did.(4, 5) Also, because patients' muscular conditions were heterogeneous (as patients displayed different pathologies and stages of disease) and muscle thickness variability was substantial in the assessed sample, higher ICC values may have resulted. In fact, lower values could have been obtained if a more homogenous population had been assessed.(23) We tried to overcome this limitation by estimating absolute and percentual differences between observers, which depend less on the homogeneity of the population.(23) Finally, we performed each pair of measurements separated by a three-hour interval, so that there may be hourly changes in patients' muscle thickness that may have resulted in a reduced reliability of measurements. However, previous studies did not find significant differences in rectus femoris thickness when measurements were done with 1, 6 and 24 hours intervals.(25)

We assessed the use of POCUS, an affordable method that can be used in the outpatient setting. We applied a strong methodological design, with the inclusion of both trained and untrained observers (allowing us to explore differences in reliability according to users' experience), the comparison of estimates based on one or two measurements (in order to explore if measuring more than once can increase reliability), the calculation of sample size and the use of more than 2 observers to collect muscle thickness data.

Our study showed the reliability of POCUS to measure rectus femoris and quadriceps femoris muscle thickness to be very high, which is in line with previous studies.(15-19) These results also suggest that measurements are reliable even when done by different observers and at different time points, which suggests that POCUS can possibly be a reliable tool to estimate evolution over time, although prospective studies should be performed to assess this hypothesis.

The learning curve for this procedure appears to be faster than with other US techniques/metrics, since trained and untrained observers had similar results. Absence of significant differences between trained and untrained observers with the use of POCUS were also observed in other fields, such as rheumatology, vascular and emergency medicine as well as for quadriceps muscle thickness measured in intensive care units (which also did not appear to depend on clinical experience).(26-29) This suggests that this technique can be used across a wider set of physicians with a single and short training course in order to standardize the data collection. The high reliability of POCUS and the possibility of its wider use across physicians is especially important for the diagnosis of conditions such as sarcopenia which, according to some algorithms, requires the measurement of quadriceps femoris muscle thickness.[2, 30] This is all the more relevant as sarcopenia is an underdiagnosed condition; the wider use of POCUS may curb such underdiagnosis.[3]

The absolute difference between same and different observers had a poor correlation with muscle thickness, which suggests that between-measurement differences are usually low, independently of the muscle thickness differences. This is opposite to what might have been intuitively thought that assessments in thicker muscles would have higher errors.

Muscle thickness was moderately correlated with knee extension peak force, sit to stand test and handgrip strength, and weakly correlated with other functional variables. Previous studies found similar results; this may be due to the fact that thickness does not perfectly correlate with the muscle cross-sectional area and, more importantly, with muscle quality, such as fat and fibrosis levels.[31-34] Possibly, correlation is stronger when muscle thickness variation and functional variation are measured over time than with a simple cross-sectional correlation between thickness and function, which depends on other variables like muscle quality.[35] Further longitudinal studies should be conducted to estimate this correlation prospectively.

CONCLUSION

In conclusion, POCUS may be a reliable resource for physicians to measure muscle thickness of patients in the rehabilitation setting. A short training course seems adequate to allow reproducibility of the procedure even when physicians had no previous experience in US imaging. Further studies may be performed to further explore the use of POCUS in the rehabilitation setting, including (i) whether differences in muscle thickness as measured by the POCUS correlate with functional differences assessed in a prospective way, and (ii) whether generalised use of POCUS is cost-effective.

References

1. Mithal A, Bonjour JP, Boonen S, Burckhardt P, Degens H, El Hajj Fuleihan G, et al. Impact of nutrition on muscle mass, strength, and performance in older adults. Osteoporos Int. 2013;24(5):1555-66.

2. Kara M, Kaymak B, Frontera W, Ata AM, Ricci V, Ekiz T, et al. Diagnosing sarcopenia: Functional perspectives and a new algorithm from the ISarcoPRM. J Rehabil Med. 2021;53(6):jrm00209.

3. Eschalier R, Massoullié G, Boirie Y, Blanquet M, Mulliez A, Tartière PL, et al. Sarcopenia in patients after an episode of acute decompensated heart failure: An underdiagnosed problem with serious impact. Clin Nutr. 2021;40(6):4490-9.

4. Weinel LM, Summers MJ, Chapple LA. Ultrasonography to measure quadriceps muscle in critically ill patients: A literature review of reported methodologies. Anaesth Intensive Care. 2019;47(5):423-34.

5. Tourel C, Burnol L, Lanoiselé J, Molliex S, Viallon M, Croisille P, et al. Reliability of standardized ultrasound measurements of quadriceps muscle thickness in neurological critically ill patients: a comparison to computed tomography measures. J Rehabil Med. 2020;52(3):jrm00032.

6. [Available from: http://www.emro.who.int/health-topics/disabilities/index.html.

7. Rantanen T. Muscle strength, disability and mortality. Scand J Med Sci Sports. 2003;13(1):3-8.

8. Hughes MA, Myers BS, Schenkman ML. The role of strength in rising from a chair in the functionally impaired elderly. J Biomech. 1996;29(12):1509-13.

9. Ploutz-Snyder LL, Manini T, Ploutz-Snyder RJ, Wolf DA. Functionally relevant thresholds of quadriceps femoris strength. J Gerontol A Biol Sci Med Sci. 2002;57(4):B144-52.

10. Parry SM, El-Ansary D, Cartwright MS, Sarwal A, Berney S, Koopman R, et al. Ultrasonography in the intensive care setting can be used to detect changes in the quality and quantity of muscle and is related to muscle strength and function. J Crit Care. 2015;30(5):1151.e9-14.

11. Freilich RJ, Kirsner RL, Byrne E. Isometric strength and thickness relationships in human quadriceps muscle. Neuromuscul Disord. 1995;5(5):415-22.

12. Messina C, Maffi G, Vitale JA, Ulivieri FM, Guglielmi G, Sconfienza LM. Diagnostic imaging of osteoporosis and sarcopenia: a narrative review. Quant Imaging Med Surg. 2018;8(1):86-99.

13. Franchi MV, Longo S, Mallinson J, Quinlan JI, Taylor T, Greenhaff PL, et al. Muscle thickness correlates to muscle cross-sectional area in the assessment of strength training-induced hypertrophy. Scand J Med Sci Sports. 2018;28(3):846-53.

14. Katari Y, Srinivasan R, Arvind P, Hiremathada S. Point-of-Care Ultrasound to Evaluate Thickness of Rectus Femoris, Vastus Intermedius Muscle, and Fat as an Indicator of Muscle and Fat Wasting in Critically Ill Patients in a Multidisciplinary Intensive Care Unit. Indian J Crit Care Med. 2018;22(11):781-8.

15. Pardo E, El Behi H, Boizeau P, Verdonk F, Alberti C, Lescot T. Reliability of ultrasound measurements of quadriceps muscle thickness in critically ill patients. BMC Anesthesiol. 2018;18(1):205.

16. Sahathevan S, Khor BH, Yeong CH, Tan TH, Meera Mohaideen AK, Ng HM, et al. Validity of Ultrasound Imaging in Measuring Quadriceps Muscle Thickness and Cross-Sectional Area in Patients Receiving Maintenance Hemodialysis. JPEN J Parenter Enteral Nutr. 2021;45(2):422-6.

17. Vieira L, Rocha LPB, Mathur S, Santana L, Melo PF, Silva VZMD, et al. Reliability of skeletal muscle ultrasound in critically ill trauma patients. Rev Bras Ter Intensiva. 2019;31(4):464-73.

18. Mayer KP, Dhar S, Cassity E, Denham A, England J, Morris PE, et al. Interrater Reliability of Muscle Ultrasonography Image Acquisition by Physical Therapists in Patients Who Have or Who Survived Critical Illness. Phys Ther. 2020;100(9):1701-11.

19. Hadda V, Khilnani GC, Kumar R, Dhunguna A, Mittal S, Khan MA, et al. Intra- and Inter-observer Reliability of Quadriceps Muscle Thickness Measured with Bedside Ultrasonography by Critical Care Physicians. Indian J Crit Care Med. 2017;21(7):448-52.

20. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet. 2007;370(9596):1453-7.

21. Ishida H, Watanabe S. Influence of inward pressure of the transducer on lateral abdominal muscle thickness during ultrasound imaging. J Orthop Sports Phys Ther. 2012;42(9):815-8.

22. Ishida H, Suehiro T, Suzuki K, Yoneda T, Watanabe S. Influence of the ultrasound transducer tilt on muscle thickness and echo intensity of the rectus femoris muscle of healthy subjects. J Phys Ther Sci. 2017;29(12):2190-3.

23. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016;15(2):155-63.

24. Schober P, Boer C, Schwarte LA. Correlation Coefficients: Appropriate Use and Interpretation. Anesth Analg. 2018;126(5):1763-8.

25. Barotsis N, Tsiganos P, Kokkalis Z, Panayiotakis G, Panagiotopoulos E. Reliability of muscle thickness measurements in ultrasonography. Int J Rehabil Res. 2020;43(2):123-8.

26. Kumar R, Shah TH, Hadda V, Tiwari P, Mittal S, Madan K, et al. Assessment of quadriceps muscle thickness using bedside ultrasonography by nurses and physicians in the intensive care unit: Intra- and inter-operator agreement. World J Crit Care Med. 2019;8(7):127-34.

27. Gutiérrez M, Di Geso L, Rovisco J, Di Carlo M, Ariani A, Filippucci E, et al. Ultrasound learning curve in gout: a disease-oriented training program. Arthritis Care Res (Hoboken). 2013;65(8):1265-74.

28. Eftedal O, Brubakk AO. Agreement between trained and untrained observers in grading intravascular bubble signals in ultrasonic images. Undersea Hyperb Med. 1997;24(4):293-9.

29. Blehar DJ, Barton B, Gaspari RJ. Learning curves in emergency ultrasound education. Acad Emerg Med. 2015;22(5):574-82.

30. Kara M, Kaymak B, Ata AM, Özkal Ö, Kara Ö, Baki A, et al. STAR-Sonographic Thigh Adjustment Ratio: A Golden Formula for the Diagnosis of Sarcopenia. Am J Phys Med Rehabil. 2020;99(10):902-8.

31. Abe T, Loenneke JP, Thiebaud RS. Morphological and functional relationships with ultrasound measured muscle thickness of the lower extremity: a brief review. Ultrasound. 2015;23(3):166-73.

32. Núñez M, Nuñez E, Moreno JM, Segura V, Lozano L, Maurits NM, et al. Quadriceps muscle characteristics and subcutaneous fat assessed by ultrasound and relationship with function in patients with knee osteoarthritis awaiting knee arthroplasty. J Clin Orthop Trauma. 2019;10(1):102-6.

33. Nishihara K, Kawai H, Kera T, Hirano H, Watanabe Y, Fujiwara Y, et al. Correlation of physical function with the thickness of multiple muscles of the quadriceps femoris in community-dwelling elderly individuals. Clin Interv Aging. 2018;13:1945-51.

34. Yoshiko A, Ogawa M, Shimizu K, Radaelli R, Neske R, Maeda H, et al. Chair sit-to-stand performance is associated with diagnostic features of sarcopenia in older men and women. Arch Gerontol Geriatr. 2021;96:104463.

35. Andrews JS, Trupin L, Schmajuk G, Barton J, Margaretten M, Yazdany J, et al. Muscle strength, muscle mass, and physical disability in women with systemic lupus erythematosus. Arthritis Care Res (Hoboken). 2015;67(1):120-7.

Authors' contributions

João Pinto-Ramos has given substantial contributions to the conception, design of the manuscript, statistical analysis, interpretation of the data and writing. Bernardo Sousa-Pinto, Cristina Costa-Santos and Joana Barroso contributed to the design, statistical analysis and revision of the manuscript. Frederico Costa, Helena Tavares, João Cabral, Tiago Moreira and Rui Brito contributed to acquisition of data.

All authors contributed equally to the manuscript and read and approved the final version of the manuscript.

Tables

Table 1: Demographic characteristics of included patients

	Included participants (N=29)
Age (years) – mean (SD)	58.8 (14.1)
Males – *n* (%)	22 (75.9)
Weight (Kg) – mean (SD)	72.7 (13.4)
Height (m) – mean (SD)	1.72 (0.10)
BMI (Kg/m^2) – median (P25-P75)	23.1 (21.5-27.4)
Previous mRankin – median (P25-P75)	0 (0 - 1)
Current mRankin – median (P25-P75)	3 (2 - 4)
Medical Research Council Sum Score - median (P25-P75)	46 (43 – 47.5)
Handgrip Strength (kg) – median (P25-P75)	18.3 (11.7 – 23.3)
Timed Up and Go test (s) – mean (SD)[a]	18.9 (11.6)
Sit to Stand test – mean (SD)[b]	17.2 (7.6)

BMI – Body Mass Index, P25-P75 – 25th and 75th percentile, Kg – Kilogram, m – Meter, s – Seconds, SD – Standard Deviation; a – Only 16 participants could complete the test, b - Only 10 participants could complete the test

Table 2: Intra-rater and Inter-rater ICC of rectus femoris and quadriceps femoris thickness measurements using Average, Maximum, Minimum and First values between AB and CD assessment and between AB measurements, respectively

	ICC calculated based on			
	Average of measurements (95%CI)	**Maximum of measurements (95%CI)**	**Minimum of measurements (95%CI)**	**First measurements (95%CI)**
Intra-rater:				
Rectus Femoris	0.956 (0.937-0.970)	0.908 (0.870-0.935)	0.911 (0.874-0.937)	0.901 (0.861-0.9319)
Quadriceps Femoris	0.966 (0.951-0.976)	0.932 (0.904-0.953)	0.927 (0.896-0.949)	0.922 (0.889-0.945)
Inter-rater:				
Rectus femoris	0.919 (0.863-0.957)	0.912 (0.853-0.953)	0.908 (0.847-0.951)	0.888 (0.815-0.940)
Quadriceps femoris	0.945 (0.907-0.971)	0.941 (0.900-0.969)	0.941 (0.900-0.969)	0.937 (0.892-0.966)

95%CI – 95% Confidence Interval, ICC - Intraclass Correlation Coefficient.

Table 3: Absolute difference [in millimetres] of rectus femoris and quadriceps femoris muscle thickness mean, maximum, minimum and first measurement between assessment AB and CD for same observer

	Absolute difference [mm] between average of measurements (P25-P75)	**Absolute difference [mm] between maximum of measurements (P25-P75)**	**Absolute difference [mm] between minimum of measurements (P25-P75)**	**Absolute difference [mm] between first measurement (P25-P75)**
Rectus Femoris	1.0 (0.4 – 1.7)	0.9 (0.4 – 1.7)	1.1 (0.4 – 1.9)	1.2 (0.5 – 2.0)
Quadriceps Femoris	1.5 (0.6 – 2.3)	1.4 (0.5 – 2.7)	1.5 (0.7 – 2.6)	1.5 (0.7 – 2.7)

P25-P75 – 25th and 75th percentile, mm - millimetres.

Table 4: Percentual difference of absolute difference in millimetres of rectus femoris and quadriceps femoris muscle thickness average, maximum, minimum and first measurement over average rectus femoris and quadriceps femoris (respectively) muscle thickness between assessment AB and CD for same observer

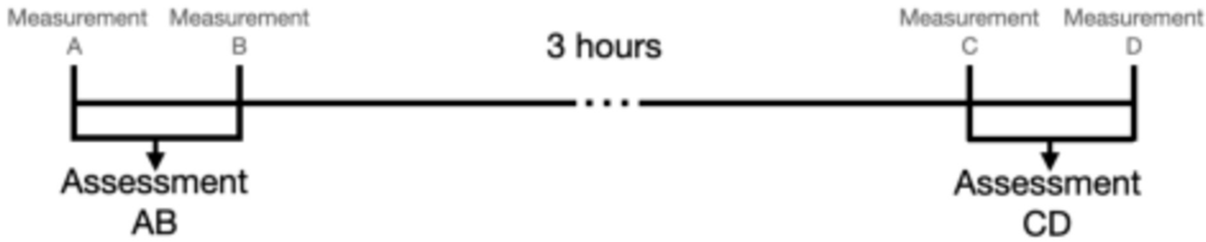
	Percentual difference between average measurement and total thickness (P25-P75)	**Percentual difference between maximum measurement and total thickness (P25-P75)**	**Percentual difference between minimum measurement and total thickness (P25-P75)**	**Percentual difference between first measurement and total thickness (P25-P75)**
Rectus Femoris	7.1% (2.2 – 11.8)	5.1% (2.2 – 11.6)	6.9% (2.6 – 13.2)	9.2% (3.2 – 16.0)
Quadriceps Femoris	6.7% (3.7 – 9.9)	6.1% (3.4 – 10.3)	6.8% (3.5 – 8.6)	5.9% (3.8 – 10.3)

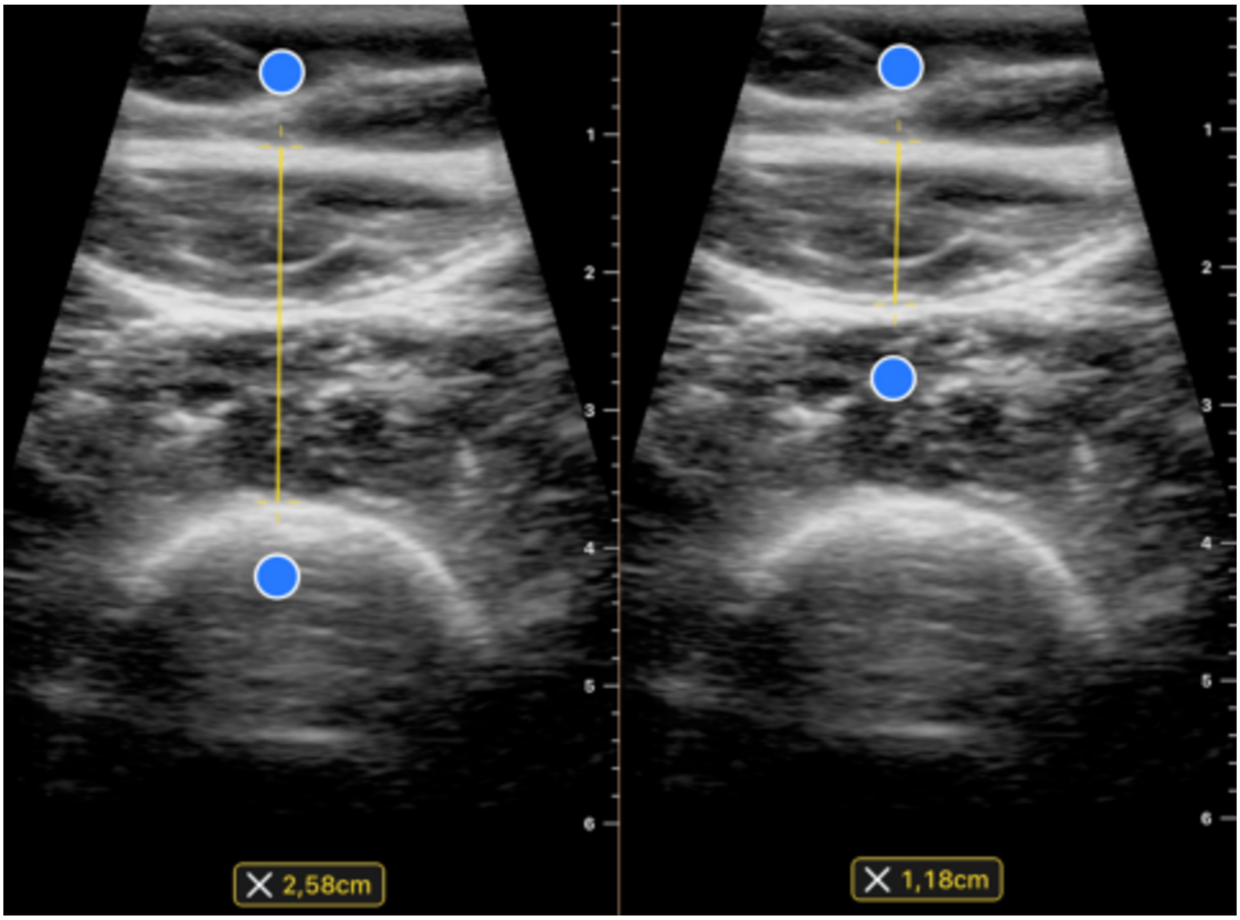
P25-P75 – 25th and 75th percentile.

Figures

Figure 1 – Schematic representation of the measurements performed by each observer in each participant

Figure 2 – Ultrasound measurements of quadriceps femoris (left) and rectus femoris (right); Yellow line represent the measurement of muscle thickness.


Measurement A
Measurement B
3 hours
Measurement C
Measurement D
Assessment AB
Assessment CD



1
2
3
4
5
6
2,58cm
1,18cm

Supplementary Digital Material

Download supplementary material file: Eur J Phys Rehabil Med-7432_Supplementary Digital Material3_V1_2022-04-27.docx

Association Between Knee Extension Strength and Functional Capacity After Intensive Care Unit Discharge

A 6-Mo Prospective Cohort Study

AQ1 João Pinto-Ramos, MD, Tiago Moreira, MD, Liliana Costa, MD, Frederico Costa, MD, Joana Barroso, MD, PhD, and Bernardo Sousa-Pinto, MD, PhD

Introduction: Assessing functional improvement after intensive care unit discharge is particularly challenging. The aim of this study was to measure the association between (1) changes in knee extension muscle strength or quadriceps femoris and rectus femoris muscle thickness and (2) changes in functionality/function-related measurements in post–intensive care unit patients.

Methods: This prospective cohort study included adult patients without previous disability, consecutively selected after intensive care unit discharge. Some parameters, such as Short-Form 36, 6-min walking test, 1-min sit-to-stand, and Short Physical Performance Battery, were measured at baseline and 3 and 6 mos after discharge. Correlations were assessed and regression models were built to assess the association between evolution in knee extension strength or muscle thickness and evolution in functional tests.

Results: Thirty patients completed the follow-up. Moderate correlation was found between knee extension strength change and Short-Form 36 physical functioning (correlation coefficient [ρ] = 0.53), 6-min walking test (ρ = 0.38), 1-min sit-to-stand (ρ = 0.52), and Short Physical Performance Battery (ρ = 0.38). Baseline values and changes in knee extension strength moderately predicted evolution in Short-Form 36 physical functioning (r^2 = 0.32, P = 0.006). Changes in muscle thickness were overall not associated with changes in functional variables.

Conclusion: Changes in knee extension muscle strength may inform on functional progression over time after intensive care unit discharge, although confirmatory studies are needed.

Key Words: Quadriceps Muscle, Muscle Strength Dynamometer, Ultrasonography, Intensive Care Units, Disability Evaluation

(*Am J Phys Med Rehabil* 2022;00:00–00)

What Is Known

- Patients in intensive care units lose muscle mass, strength, and functional capacity. These parameters are progressively recovered over the subsequent months.

What Is New

- Weight-adjusted knee extension strength change over 6 mos seems to be correlated with changes in function-related parameters, especially when combined with information on baseline functionality or strength.
- Quadriceps and rectus femoris muscle thickness changes do not seem to correlate with function-related parameters.
- Changes in knee extension strength and changes in functional parameters tend to display better correlation in the first 3 mos after discharge than in the following months.

Patients in intensive care units (ICUs) tend to lose muscle mass and strength and, concomitantly, functional capacity. During the first days of an ICU admission, an average of 3%–4% of muscle mass is lost per day.[1] Antigravity muscles are the most affected with immobilization—the rectus femoris loses an average of 9% of its thickness and cross-sectional area after three ICU admission days, and 30% after 10 days. The loss of muscle mass of the quadriceps muscle occurs at an even faster pace.[2] In the months after discharge, most patients tend to recover the muscle mass, strength, and functional capacity lost during ICU admission.

Particularly regarding functionality, there are some functional and function-related tests and questionnaires that can be used in the clinical practice in patients after ICU hospitalization to allow physicians to estimate patients' recovery. The Short Form-36 (SF-36) is the most frequently used function-related questionnaire in the intensive care setting,[3] where it has been shown to have an acceptable validity and reliability.[4] Other frequently used questionnaires and tests that have been validated to measure functional capacity and function-related parameters include the Barthel Index, handgrip strength, 6-min walking test (6MWT), 1-min sit-to-stand (1MSTS), timed up and go (TUG)

From the Department of Physical Medicine and Rehabilitation, Centro Hospitalar Universitário São João, Porto, Portugal (JP-R, TM, FC, JB); CINTESIS—Center for Health Technologies and Services Research, University of Porto, Porto, Portugal (JP-R, BS-P); Intensive Care Unit, Centro Hospitalar Universitário São João, Porto, Portugal (LC); Department of Biomedicine, Faculty of Medicine, University of Porto, Porto, Portugal (JB); i3s—Institute for Health Research and Innovation, University of Porto, Porto, Portugal (JB); Departments of Neuroscience and Physical Medicine and Rehabilitation, Northwestern University, Feinberg School of Medicine, Chicago, Illinois (JB); and MEDCIDS—Department of Community Medicine, Information and Health Decision Sciences, Faculty of Medicine, University of Porto, Porto, Portugal (BS-P).

All correspondence should be addressed to: João Pinto-Ramos, MD, Physical Medicine and Rehabilitation Service, Centro Hospitalar Universitário São João, Alameda Prof. Hernâni Monteiro, Porto, Portugal.
AQ2

This manuscript has not been previously submitted to any other publication. Authorship has been granted only to those individuals who have contributed substantially to the manuscript. All authors have read and approved the manuscript.

This article was supported by National Funds through FCT—Fundação para a Ciência e a Tecnologia, I.P., within CINTESIS, R&D Unit (reference UIDP/4255/2020).

Financial disclosure statements have been obtained, and no conflicts of interest have been reported by the authors or by any individuals in control of the content of this article.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.ajpmr.com).

ISSN: 0894-9115
DOI: 10.1097/PHM.0000000000002137

test, and Short Physical Performance Battery (SPPB).[5–9] Some of AQ3 these parameters (such as SF-36 physical functioning, 6MWT, 1MSTS, and SPPB) are predictors of mortality and health-related quality of life.[10–13]

However, although these different tests provide valuable and complementary information, there is not a clear single gold standard test to estimate functional capacity in post-ICU patients. In addition, the most commonly used tests are time-consuming, and their application is challenging in the outpatient setting because of logistic problems and time restrictions. This prompts the need for easily applicable, reliable, and valid tools for estimating functionality and improvement over time. Possibilities for such tools include the measurement of strength and muscle mass with handheld dynamometer (HHD) and point-of-care ultrasound (POCUS), respectively.[14,15] These are not only reliable and simple tools but also portable, inexpensive, and noninvasive, rendering them potentially easily applicable in the clinical practice.[14–18] The rationale for their potential use lies on the fact that functional capacity, muscle mass, and strength tend to improve after discharge, so there might be an association between the evolution of these variables. That is, HHD and POCUS might be useful to assess the evolution of functionality.[6,19,20]

The purpose of this work was to do an exploratory assessment on the association between changes in knee extensor muscle thickness and strength (measured using POCUS and HHD) and correspondent changes in SF-36 physical functioning in ICU patients. Additional aims include the measurement of the association of changes in muscle thickness and strength with changes in other functional ability scales, namely, the Barthel Index, handgrip strength, 6MWT, 1MSTS, TUG test, and SPPB. Therefore, this study was devised considering the potential of muscle strength and thickness measurements in the clinical practice, even though it does not aim to fully clarify their role in the clinical setting (as that would include, among others an assessment of their feasibility and cost effectiveness).

METHODS

Study Design

This study was written according to STROBE statement guidelines[21] (see Supplementary Checklist, Supplemental Digital Content 1, http://links.lww.com/PHM/B880). In this prospective study, a consecutive sample of participants after ICU discharge was assessed. Patients were initially assessed during hospitalization within 3 days after recovering walking capability (ability to walk 10 m in the corridor without assistance)—assessment 1 (A1). Patients were subsequently assessed 3 mos (assessment 2 F1 [A2]) and 6 mos (assessment 3 [A3]) after A1 (Fig. 1). For each assessment, one physical medicine and rehabilitation resident measured the patient's knee extension strength and quadriceps femoris thickness with HHD and POCUS, respectively, and applied a set of questionnaires (SF-36 physical functioning, Barthel Index, and International Physical Activity Questionnaire [IPAQ]) and physical tests (6MWT, 1MSTS, TUG test, handgrip strength, Medical Research Council sum score [MRCSS] and SPPB).

This study was approved by the ethical committee of the hospital where this study was conducted. Patients provided written informed consent before participation to be included in the study.

Setting and Participants

Patients from a single center (Centro Hospitalar Universitário de São João, a tertiary care hospital in Northern Portugal) meeting eligibility criteria were consecutively included after ICU discharge from end March to mid of August 2021. Patients 18 yrs or older who had undergone more than 5 days of ICU hospitalization and were previously able to ambulate without assistance before hospitalization were included. Exclusion criteria were severe cognitive impairment, major neurologic/neuromuscular conditions such as traumatic brain injury or stroke (but not ICU-acquired weakness), major psychiatric conditions, unrecoverable disease, traumatic lesions with long bone fractures, severe obesity (body mass index ≥35 kg/m^2), full thickness pressure ulcers, cardiorespiratory instability, and current or planned home ventilation.

Variables and Measurements

Each participant was assessed three times, namely, (1) during admission (within 3 days after walking capability was recovered—A1), (2) 3 mos after A1 (A2), and (3) 6 mos after A1 (A3). Each assessment included two HHD (Micro FET®2 HHD, Hoggan Health Industries, Draper, UT) measurements of knee extension strength. The HHD was placed 5 cm above the distal portion of the medial malleolus of the dominant leg, with the patient seated in the edge of the stretcher.[17] In addition, each assessment encompassed two measurements of both rectus femoris and quadriceps femoris muscle thickness using a POCUS (Butterfly iQ+, Butterfly Network, Inc, Guilford, CT), configured to musculoskeletal image acquisition. Both rectus femoris and quadriceps femoris were assessed because there is no consensus on which one is preferable to assess knee extension. The POCUS was placed 15 cm above the superior pole of the patella of the same leg of the HHD measurements, with the patient in supine position with both legs stretched and minimal hand force applied perpendicular to muscle fibers. The distance between the cortex of the femur and the most superficial muscular fascia was measured.[18] Figure 2 illustrates a representative scanned image. F2

In each assessment, a set of questionnaires (namely, the SF-36 physical functioning, the Barthel Index, and IPAQ) and physical function tests (namely, the 6MWT, 1MSTS, TUG test, handgrip strength, MRCSS, and SPPB) were applied to all participants. In addition, at A1, patient-reported IPAQ and Barthel Index before hospital admission (Fig. 3) were assessed. F3 Supplementary Table 1 (Supplemental Digital Content 2, http://links.lww.com/PHM/B881) summarizes the function-related parameters evaluated in this study as well as the time to perform them and their properties. Questionnaires and physical function tests were applied by one of three physical medicine and rehabilitation residents (according to schedule availability) with at least 1 yr of experience in physical test performance and both HHD and POCUS experience. Residents were blinded for previous results from each patient or from other patients.

Information on the participants' demographic characteristics (sex and age), height, weight, body mass index, diagnosis at admission, Acute Physiology Assessment and Chronic Health Evaluation II (Apache II), Simplified Acute Physiology Score II

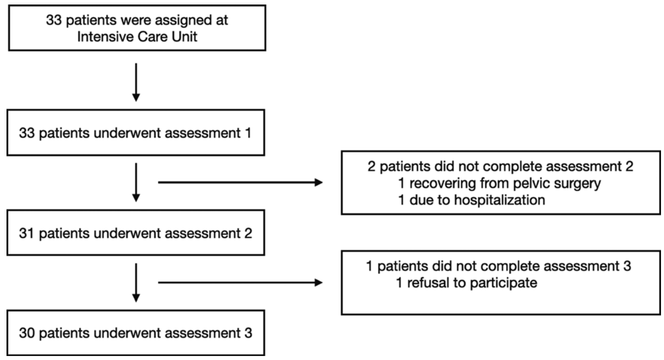


FIGURE 1. Flow diagram of patients included in the study and withdrawals at each stage.

(SAPS II), and hospitalization in ICU was retrieved from their electronic health records. Apache II and SAPS II are both classification systems of the severity of disease, with higher scores indicating higher chances of mortality (ranging from 0 to 163 for Apache II and 0 to 71 for SAPS II).

The assessments were performed in different settings, with A1 having been performed during hospitalization in the ward where the patient had been admitted, and A2 and A3 having been performed during physical medicine and rehabilitation outpatient visits.

Study Size

Sample size calculation was made to detect a significant correlation coefficient between HHD or POCUS variation and SF-36 physical functioning variation of at least 0.5 (defined by AQ4 the COSMIN guidelines as indicator of good correlation).[22] A sample size of 33 was found based on a power of 80% and a level of significance of 0.05 and considering a 15% patient withdrawal.

Statistical Analysis

Means and standard deviations were used for describing continuous variables with normal distributions, whereas medians and 25th and 75th percentiles (P25;P75) were used for variables with non-normal distributions. Absolute and relative frequencies were used to describe categorical variables.

Changes (arithmetic difference between results at different assessment periods) between the results of HHD and POCUS were calculated and questionnaires and physical tests were applied (with such differences henceforth indicated as "Δ"). The correlations—using Spearman correlation coefficients—between Δ HHD or Δ POCUS at each time interval and (1) Δ SF-36 physical functioning, (2) Δ Barthel Index, (3) Δ IPAQ, (4) Δ 6MWT, (5) Δ MRCSS, (6) Δ 1MSTS, (7) Δ TUG, (8) Δ SPPB chair stand test, (9) Δ SPPB balance tests, (10) Δ SPPB Gait Speed Test, (11) Δ SPPB, and (12) Δ handgrip strength were estimated. Whenever Spearman correlation coefficients were higher than 0.30 (considered as a cut-off for moderate correlation in some studies),[23,24] multiple linear regression models were built using either baseline and Δ muscle strength (as measured with HHD) or muscle thickness (POCUS) as independent variables and Δ of each of the questionnaires and physical function tests as dependent variable. This allowed obtaining models predicting change in functionality based on two variables: (1) baseline muscle strength or thickness and (2) respective change in muscle strength or thickness; for each of these models, the r^2 and the omnibus *P* value were retrieved. Alternative models were built adjusting for the baseline dependent variable (instead of the baseline muscle

Fig 2 4/C

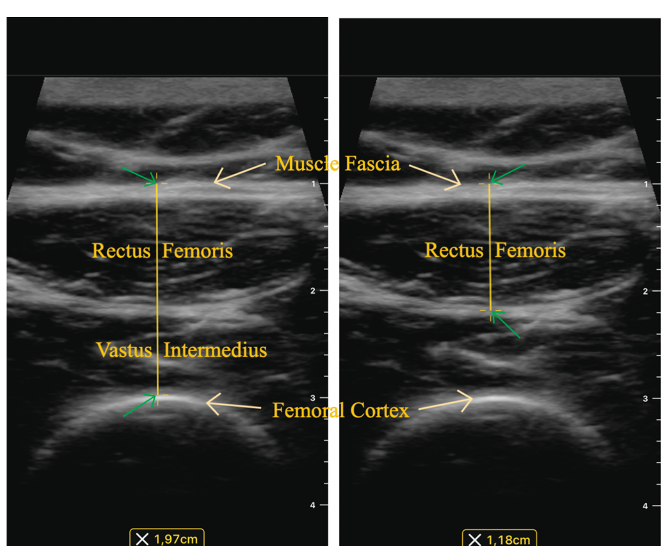


FIGURE 2. Representative scanned image of quadriceps (left) and rectus (right) femoris. Distance was measured between yellow line edges (green arrows).

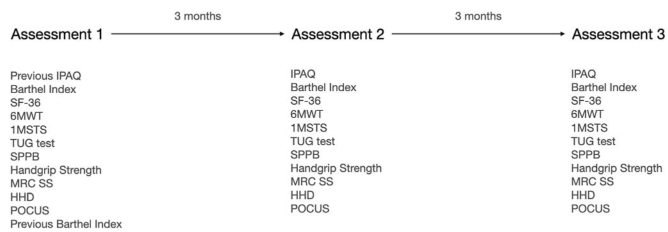


FIGURE 3. Variables collected at each assessment.

strength and thickness), as suggested by Husted et al.[25] Further covariates were not included in the models because of sample size restrictions. In addition, for each patient and for each of the used main function-related parameters (namely, the SF-36 physical functioning, the 1MSTS, the 6MWT, the SPPB, and the TUG tests), the number of minimal clinically important differences (MCIDs) that occurred in between the last and the first assessments were calculated (e.g., for 6MWT, the MCID is 50 m, and a patient whose 6MWT improves by 100 m displays a recovery corresponding to 2 MCID). The average of the number of MCIDs registered for the different parameters was then calculated and correlated with HHD weight-adjusted strength variation in the same period.

Statistical significance was defined at $P < 0.05$. All analyses were performed using the software IBM SPSS Statistics 28 (IBM, Armonk, NY).

RESULTS

A total of 33 participants who agreed to participate in this study and completed A1 were included. Six months' follow-up was accomplished for 30 participants, with assessments being finished in February 2022. Two patients did not complete A2 and A3 because of hospitalization and pelvic surgery, respectively, and one patient did not complete A3 because of refusal to participate (Fig. 1). Eleven patients underwent rehabilitation sessions during hospitalization.

Among patients who completed their follow-up, 23 were male (76.7%), mean (SD) participant age was 60.4 (14.2) yrs, and average body mass index was 25.5 kg/m^2 (SD, 4.2 kg/m^2) (Table 1). Participants were hospitalized in the ICU for a median of 10 days (P25;P75, 5.75;17.25), with 16 (53.3%) patients having received invasive ventilation, 14 (46.7%) noninvasive ventilation, and 2 (6.7%) extracorporeal membrane oxygenation. The ICU gravity score SAPS II mean (SD) value was 41.8 (18.2), and the Apache II score median value was 19 (P25;P75, 12.25;26.25). Before admission, patients' median Barthel Index was 100 (P25;P75, 100;100), and the median IPAQ was 627 (P25;P75, 99;2973).

Variables Progression During Follow-up

Tests assessing physical function, such as gait speed test, TUG, 6MWT, 1MSTS, hand grip strength, and SPPB, had a greater improvement from A1 to A2 than from A2 to A3 (Figs. 4 and 5). The same pattern was seen in SF-36 physical functioning, whereas improvements in the SF-36 general health remained stable over the follow-up period. Patients' median level of activity reported by IPAQ before admission (594 MET-min/week; P25;P75, 165;2976) was not achieved during the follow-up period, with an A3 median value of 396 MET-min/week (P25; P75, 99;1640). On the other hand, patients' disability measured by the median Barthel Index matched preadmission levels after 3 mos at visit A2, with patients reaching previous levels on basic functions.

Both quadriceps and rectus femoris muscle thickness increased steadily during the follow-up period, whereas knee extension muscle strength had a larger increase between A1 and A2 than between A2 and A3 (Fig. 5).

Association Between Variations of Functional Variables and Muscle Strength or Thickness

Moderate correlation was found between Δ knee extension muscle strength and Δ SF-36 physical functioning at Δ A1–A3 and Δ A1–A2 but not at Δ A2–A3 (Spearman correlation coefficients = 0.53 [P = 0.003], 0.45 [P = 0.017], and −0.01 [P = 0.967], respectively) (Table 2). In multiple linear regression models, Δ knee extension muscle strength was shown to be significantly associated with Δ in SF-36 physical functioning in the A1–A3 period (r^2 = 0.32, P = 0.006) (Table 3). The Δ knee extension muscle strength was also correlated with other variables, with the strongest correlations being those with Δ 6MWT, Δ 1MSTS, and Δ handgrip strength (the latter one only for Δ A1–A2). These variables were also those associated with the best-performing multiple linear regression models, particularly at Δ A1–A2. Overall, correlations tended to be stronger and regression models tended to display a better performance

TABLE 1. Demographic and clinical characteristics of the assessed sample

Variables	Patients Who Completed Follow-Up (N = 30)
Age, mean (SD), years	60.4 (14.3)
Males, n (%)	23 (76.7)
Weight, mean (SD), kg	73.3 (14.9)
Height, mean (SD), m	1.69 (0.07)
BMI, mean (SD), kg/m^2	25.5 (4.2)
ICU hospitalization, median (P25;P75), days	10 (5.75;17.25)
Previous Barthel Index, median (P25;P75)	100 (100;100)
Previous IPAQ, median (P25;P75)	627 (99;2973)
Apache, median (P25;P75)	19 (12.25;26.25)
SAPS II, mean (SD)	41.8 (18.2)
Noninvasive ventilation, n (%)	14 (46.7)
Invasive ventilation, n (%)	16 (53.3)
ECMO, n (%)	2 (6.7)
Tracheostomy, n (%)	16 (53.3)

BMI indicates body mass index; ECMO, extracorporeal membrane oxygenation.

Fig 4 4/C

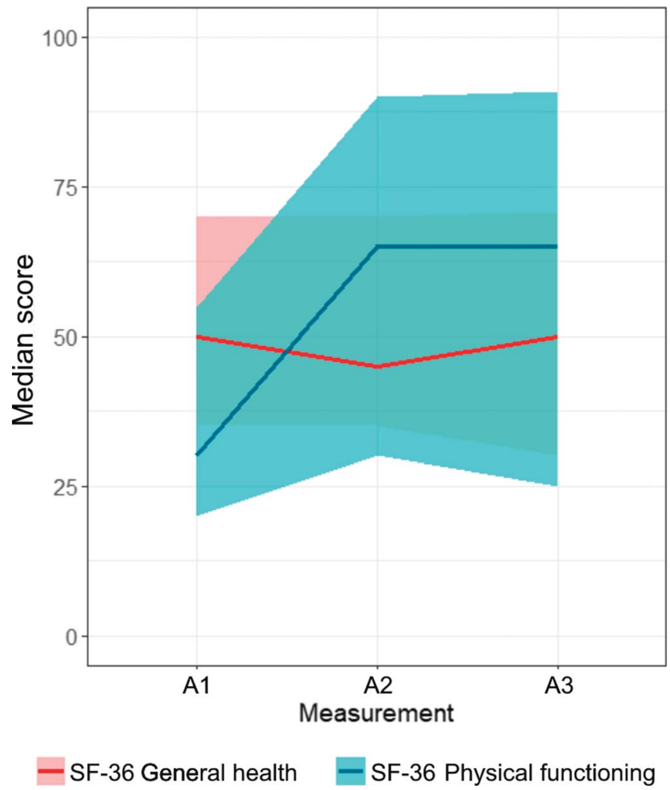


FIGURE 4. Median and 25th and 75th percentile values of SF-36 physical functioning and general health at baseline, 3 mos, and 6 mos.

when considering the Δ A1–A3 or Δ A1–A2 periods than the Δ A2–A3 periods. For the association between muscle strength change and that composite measure, a correlation coefficient of 0.56 (P = 0.002) was obtained, and the linear regression coefficient was 1.79 (95% confidence interval, 0.63–2.94; P = 0.004, r^2 = 0.29).

Δ Rectus femoris muscle thickness and Δ quadriceps femoris muscle thickness displayed weak correlations (correlation coefficient <0.30) with most Δ functional variables, with the exception of SF-36 physical functioning, 1MSTS, 6MWT, SPPB chair stand test, and handgrip strength (Table 2). Multiple linear regression showed statistically significant results for 6MWT at Δ A2–A3 for both rectus and quadriceps femoris and 1MSTS at Δ A1–A2 for POCUS rectus femoris only (Table 4). T4

A sensitivity analysis, building multiple linear regression models adjusting for each of the baseline dependent variables, was performed. For knee extension muscle strength, those models tended to display better performance than those based on baseline and Δ strength (Supplementary Table 2, Supplemental Digital Content 3, http://links.lww.com/PHM/B882). For muscle thickness, this trend was also observed, although differences in model performance tended to be much smaller (Supplementary Table 3, Supplemental Digital Content 4, http://links.lww.com/PHM/B883).

DISCUSSION

In this exploratory study, weight-adjusted HHD knee extension strength change over 6 mos was found to be correlated with changes in physical function tests and questionnaires such as SF-36 physical functioning, 6MWT, 1MSTS, and SPPB. Prediction ability was higher when information on strength change was combined with that on baseline functionality than with that on baseline strength. By contrast, changes in quadriceps and rectus femoris muscle thickness were not correlated with physical function tests. In addition, it was also found that physical function tests and questionnaires indexes had a larger improvement in the first 3 mos of the follow-up period, with correlations between changes in knee extension strength and functionality tending to be better when considering the first 3 mos after discharge than the subsequent months.

This study has some limitations worth noting; this is a single-center study, and generalization to other populations must be taken carefully. In particular, this study was conducted in a tertiary care hospital, where patients with higher Apache II and SAPS II scores are expected to be overrepresented because more severely diseased patients from other hospitals are transferred to this center. Although all patients were from the ICU setting, there was heterogeneity in their severity scores and admission periods, limiting the results' generalizability. Patients were evaluated in the ward, 3 days after starting ambulation, and some may have overestimated their real capabilities when answering the SF-36 physical functioning scale because they had a short time to perceive their limitations. This may justify why some patients reported deterioration of physical function in SF-36 from A1 to A3 while improving in functional tests at the same time. This potential information bias may have led to an underestimation of the association between knee extension muscle strength change and SF-36 change. Patients were assessed

Fig 5 4/C

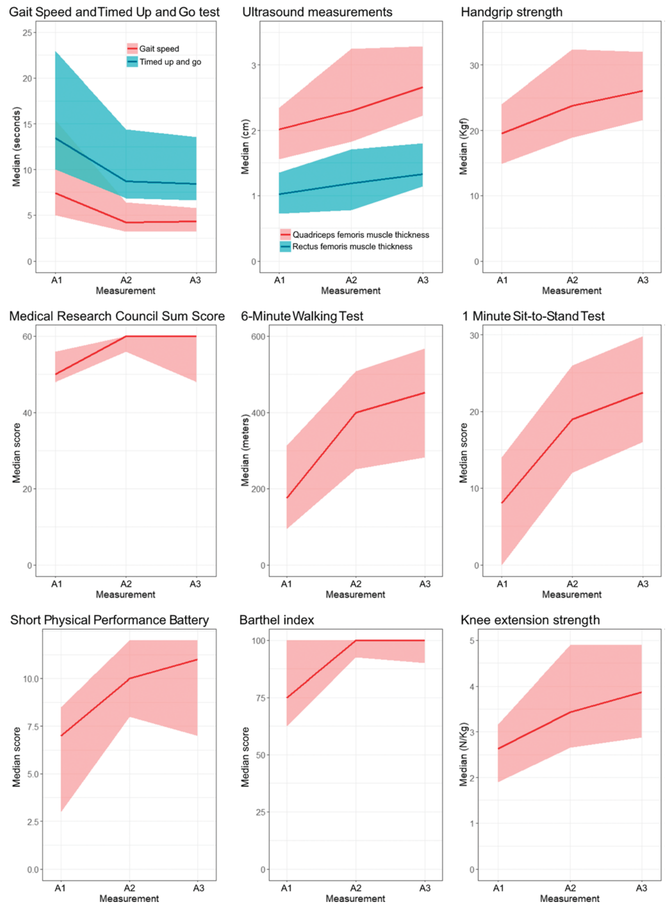


FIGURE 5. Median and 25th and 75th percentile values of different functional variables at baseline, 3 mos, and 6 mos.

in different locations, with A1 assessments having been performed in the ward during admission, and A2 and A3 in the outpatient setting. Although this may have influenced the performed measurements, the effect of this phenomenon is probably nondifferential (i.e., probably not associated with patients' specific baseline characteristics). Another limitation concerns the absence of a single gold standard comparator assessing functionality—although the authors were able to compare changes in knee extension muscle strength or thickness with changes in several function-related tests, none of these tests is a perfect gold standard for functionality and associations might have been stronger if there were a better way to measure their change over time.[3,26,27] Finally, the assessed sample size precluded inclusion of further covariates in linear regression models, including the simultaneous testing of knee extension muscle strength and muscle thickness changes or other factors affecting physical function, muscle mass or strength recovery, or the relationship between these variables. Inclusion of such variables could have resulted in models explaining a greater amount of variability of functionality changes.

This study has also some important strengths worth noting. First, this is, to the authors' knowledge, the first prospective study to evaluate the correlation of the change in functional tests, questionnaires, muscle strength, and muscle thickness after ICU hospitalization. These functional tests and questionnaires have been validated in different settings, whereas HDD and POCUS were found to display strong intrarater and interrater reliability.[17,18] In addition, these methods are portable, fast to use, and inexpensive, rendering them easily used in the clinical practice. It was also ensured that the study was adequately powered to achieve reliable results. The consistency of the obtained results also seems to be an important strength of our study.

It was observed that patient progression was faster in the first than in the second 3 mos after starting ambulation. In fact, previous studies had already pointed that improvement is faster in the first months after hospitalization, probably because functional parameters are furthest away from patients' previous state, which gives them more functional potential.[19,28] Quadriceps and rectus femoris thicknesses were the only exception, with steady muscle growth not necessarily meaning similar functional improvement. This might happen because muscle thickness does not entirely correlate with muscle volume and because strength also depends on muscle quality, such as levels of fibrosis and fat, which was not evaluated.[29,30] Also, another observation was that the Barthel Index reached levels before hospitalization at 3 mos postdischarge. In fact, at that time point, the maximum score was achieved, probably because of the ceiling effect of this test, which limits differentiation of patients with higher levels of functional capacity.[31] On the other hand, patients' previous IPAQ was never reached, even after 6 mos after hospitalization, probably owing to loss of capacity or fear of overexertion, although memory bias can also be a factor.

Knee extensor muscle strength measured with HHD may be helpful for estimating functional gains over same period.

TABLE 2. Spearman correlation coefficients for the associations between changes in knee extension muscle strength or muscle thickness and change in functional variables for the same period

Functional Variable	Change Between A1 and A3 (*P*)	Change Between A1 and A2 (*P*)	Change Between A2 and A3 (*P*)
Correlations with knee extension muscle strength change			
SF-36 physical functioning	0.53 (0.003)	0.45 (0.017)	−0.01 (0.967)
Barthel Index	0.30 (0.113)	−0.05 (0.816)	0.08 (0.686)
6MWT	0.38 (0.045)	0.48 (0.009)	0.43 (0.024)
MRCSS	0.09 (0.623)	0.23 (0.224)	0.10 (0.597)
1MSTS	0.52 (0.003)	0.81 (<0.001)	0.60 (<0.001)
TUG	−0.28 (0.133)	−0.39 (0.035)	−0.41 (0.029)
SPPB chair stand test	−0.04 (0.880)	−0.15 (0.573)	−0.13 (0.528)
SPPB balance tests	−0.02 (0.920)	0.19 (0.307)	−0.03 (0.873)
SPPB gait speed test	0.13 (0.483)	0.05 (0.790)	−0.21 (0.280)
SPPB	0.38 (0.041)	0.40 (0.027)	0.26 (0.158)
Hand Grip Strength	0.19 (0.312)	0.80 (<0.001)	0.29 (0.118)
POCUS rectus femoris	0.34 (0.062)	0.46 (0.009)	0.35 (0.056)
POCUS quadriceps femoris	0.32 (0.090)	0.47 (0.008)	0.33 (0.071)
Correlations with rectus femoris muscle thickness change			
SF-36 physical functioning	0.03 (0.886)	0.15 (0.461)	0.34 (0.077)
Barthel Index	0.17 (0.370)	−0.15 (0.442)	−0.01 (0.948)
6MWT	0.28 (0.155)	0.10 (0.628)	0.37 (0.059)
MRCSS	0.16 (0.396)	−0.08 (0.672)	−0.01 (0.956)
1MSTS	0.24 (0.193)	0.46 (0.009)	0.33 (0.077)
TUG	−0.29 (0.124)	−0.09 (0.639)	0.12 (0.528)
SPPB chair stand test	0.33 (0.217)	0.14 (0.599)	−0.34 (0.086)
SPPB balance tests	−0.03 (0.877)	−0.18 (0.330)	−0.15 (0.438)
SPPB gait speed test	−0.25 (0.189)	0.09 (0.622)	−0.17 (0.374)
SPPB	0.15 (0.445)	−0.07 (0.716)	0.23 (0.215)
Hand grip strength	0.02 (0.931)	0.37 (0.043)	0.17 (0.381)
Correlations with quadriceps femoris muscle thickness change			
SF-36 physical functioning	0.11 (0.564)	0.19 (0.344)	0.36 (0.064)
Barthel Index	−0.01 (0.971)	−0.21 (0.285)	−0.05 (0.818)
6MWT	0.14 (0.482)	−0.05 (0.806)	0.43 (0.024)
MRCSS	0.14 (0.478)	0.06 (0.737)	−0.05 (0.787)
1MSTS	0.23 (0.213)	0.41 (0.024)	0.31 (0.091)
TUG	−0.26 (0.104)	−0.09 (0.656)	0.10 (0.602)
SPPB chair stand test	0.23 (0.399)	0.10 (0.701)	−0.34 (0.083)
SPPB balance tests	−0.01 (0.950)	0.10 (0.611)	−0.16 (0.397)
SPPB gait speed test	−0.20 (0.296)	0.20 (0.296)	−0.24 (0.201)
SPPB	0.10 (0.609)	0.01 (0.958)	0.26 (0.172)
Hand grip strength	0.22 (0.234)	0.41 (0.022)	0.09 (0.632)

This was particularly suggested by the results observed in its association with SF-36 physical functioning, 6MWT, and 1MSTS, as well as the moderate correlation even with a composite outcome reflecting different function-related tests. In case of SF-36 physical functioning, such association was not found for variation between 3 and 6 mos' assessments, which might be explained by a difficulty of this variable to measure minor functional changes that are seen in this period. For other functional tests, weaker associations between their change and that of knee extension muscle strength were observed. This difference may occur because knee extension strength is more critical for some functional tasks than others, namely, 1MSTS, whereas some depend more on endurance, balance, or coordination. This diversity of results and the absence of a single gold standard to assess functionality prompt the need for further studies assessing how changes in strength predict functionality gains. Our results suggest that changes in strength can at least inform on the evolution of some dimensions of functionality in an easy and less time-consuming way. In fact, preparing the patient, assessing, and scoring the SF-36 physical functioning, the 6MWT, and the 1MSTS could take more than 20 mins: (1) the SF-36 may be difficult to use for some patients and time-consuming to score by the observers and (2) the 6MWT and 1MSTS usually require more than the theoretically expected 7 mins to be performed, given the time needed to provide patients with instructions on how to execute the tests and the time needed to relocate them. On the other hand, the HHD assessment could take less than 2 mins because a direct

TABLE 3. Results of the linear regression models predicting change in functional tests and muscle thickness based on change and baseline knee extension muscle strength

	Change Between A1 and A3		Change Between A1 and A2		Change Between A2 and A3	
Functional Variable	r^2	**Predicting Formula [*P*]**	r^2	**Predicting Formula [*P*]**	r^2	**Predicting Formula [*P*]**
SF-36 physical functioning	0.32	$y = 17.7x - 3.8h + 8.3$ [0.006]	0.21	$y = 11.7x + 0.5h + 0.5$ [0.053]	–	–
Barthel Index	0.20	$y = 2.5x - 7.8h + 32.0$ [0.053]	–	–	–	–
6MWT	0.17	$y = 27.9x - 36.9h + 253.1$ [0.103]	0.33	$y = 54.4x - 26.2h + 164.5$ [0.006]	0.15	$y = 20.9x - 4.1h + 57.0$ [0.142]
1MSTS	0.22	$y = 4.6x - 0.1h + 7.9$ [0.038]	0.62	$y = 7.1x + 0.2h + 2.2$ [<0.001]	0.40	$y = 3.6x + 0.6h + 0.2$ [0.001]
TUG	–	–	0.16	$y = -2.4x + 11.4h - 43.6$ [0.100]	0.08	$y = 0.2x + 1.1h - 5.2$ [0.345]
SPPB	0.14	$y = 0.5x - 0.5h + 3.8$ [0.133]	0.14	$y = 0.7x - 0.3h + 2.9$ [0.117]	–	–
Hand grip strength	–	–	0.54	$y = 5.5x + 2.5h - 8.2$ [<0.001]	–	–
POCUS rectus femoris	0.23	$y = 0.12x - 0.10h + 0.48$ [0.028]	0.20	$y = 0.14x - 0.10h - 0.19$ [0.047]	0.26	$y = 0.07x - 0.10h + 0.48$ [0.019]
POCUS quadriceps femoris	0.19	$y = 0.18x - 0.14h + 0.87$ [0.056]	0.23	$y = 0.28x - 0.18h - 0.25$ [0.027]	0.27	$y = 0.09x - 0.15h - 0.76$ [0.014]

y is the functional variable change between assessments; *x* is the rectus femoris muscle thickness change between assessments, in centimeters; and *h* is knee extension muscle strength at first of the two assessments, in Newton per kilogram.

value is obtained after a couple of knee extension tests. This is particularly relevant for patients' follow-up in the outpatient setting (in fact, the possibility of collecting functional variables at baseline can improve the estimation of functional improvement, complementing information on strength variation). Having a practical method to systematically estimate functional parameters (some of which are predictors of mortality and quality of life) can give physicians crucial information on patients' improvement. Besides, HHD seems feasible and reliable to apply even for previously unexperienced clinicians, which can contribute to its universal application.[17] Nevertheless, more studies should take place to confirm this hypothesis, namely, to understand whether larger improvements in knee extension strength leads to better functional and vital prognosis. In addition, before the introduction of HHD usage in the clinical practice, future studies should evaluate its feasibility, acceptability, and resources required and even compare it with newer methods of patient assessment such as video-monitored physical performance testing.

Measuring the muscle thickness of quadriceps and rectus femoris using POCUS does not seem to be a consistently valid method of estimating changes in functional variables, and its use may be dispensable in clinical practice for assessment of those changes. Besides, muscle mass gains are consistent over time and not necessarily correlated with functional improvements. A possible explanation is that neural drive adaptations may occur earlier during recovery than hypertrophic changes.[32] This challenges the assumption that patients are improving solely based on improvement of amyotrophy.

In conclusion, it was observed that knee extension strength change measured with HHD has a moderate association with changes in functional parameters (notably SF-36 physical functioning,

TABLE 4. Results of the linear regression models predicting change in functional tests based on change and baseline muscle thickness

	Change Between A1 and A3		Change Between A1 and A2		Change Between A2 and A3	
Functional Variable	r^2	**Predicting Formula [*P*]**	r^2	**Predicting Formula [*P*]**	r^2	**Predicting Formula [*P*]**
Models involving rectus femora muscle thickness						
SF-36 physical functioning	–	–	–	–	0.21	$y = 10.0x - 5.0r + 11.8$ [0.053]
6MWT	–	–	–	–	0.24	$y = 113.5x + 30.1r - 12.9$ [0.036]
1MSTS	–	–	0.22	$y = 12.9x - 1.2r + 9.2$ [0.033]	0.07	$y = 3.5x - 0.7r - 3.7$ [0.381]
SPPB chair stand test	0.09	$y = 5.9x + 3.8r - 14.6$ [0.549]	–	–	0.15	$y = -3.4x - 0.2r - 0.4$ [0.139]
Hand grip strength	–	–	0.14	$y = 8.4x - 1.1r + 3.6$ [0.120]	–	–
Models involving quadriceps femoris muscle thickness						
SF-36 physical functioning	–	–	–	–	0.21	$y = 10.5z - 1.3q + 7.4$ [0.055]
6MWT	–	–	–	–	0.35	$y = 103.8z + 28.4q - 57.7$ [0.006]
1MSTS	–	–	0.16	$y = 6.3z - 2.0q + 12.0$ [0.084]	0.09	$y = 3.3z - 0.3q + 3.3$ [0.271]
SPPB chair stand test	–	–	–	–	0.17	$y = -2.8z - 0.2q + 0.2$ [0.111]
Hand grip strength	–	–	0.16	$y = 5.1z - 1.6q + 5.2$ [0.085]	–	–

y is the functional variable change between assessments; *x* is rectus femoris muscle thickness change between assessments, in centimeters; *r* is rectus femora muscle thickness at first of the two assessments, in centimeters; *z* is quadriceps femora muscle thickness change between assessments, in centimeters; and *q* is quadriceps femora muscle thickness at first of the two assessments, in centimeters.

6MWT, and 1MSTS), rendering HHD as an easily applicable tool to estimate changes in functionality in patients after ICU discharge. Quadriceps and rectus femoris muscle thickness measured with POCUS did not correlate well with functional variables and may not be as useful in clinical practice for this purpose. Nevertheless, further confirmatory studies—with larger samples, more frequent assessments, and/or longer follow-up periods—are needed to confirm the findings obtained in this exploratory study.

ACKNOWLEDGMENTS

The authors acknowledge Dr João Barroso and Dr Fernando Parada, as current and previous heads of the Department of Physical Medicine and Rehabilitation of Centro Hospitalar Universitário São João, where this work was conducted, and Dr Maria José Festas for their logistic support. The authors also acknowledge Dr Inês Machado Vaz for her contribution to manuscript design.

REFERENCES

1. Puthucheary ZA, Phadke R, Rawal J, et al: Qualitative ultrasound in acute critical illness muscle wasting. *Crit Care Med* 2015;43:1603–11
2. Parry SM, El-Ansary D, Cartwright MS, et al: Ultrasonography in the intensive care setting can be used to detect changes in the quality and quantity of muscle and is related to muscle strength and function. *J Crit Care* 2015;30:1151.e9–14
3. Oeyen SG, Vandijck DM, Benoit DD, et al: Quality of life after intensive care: A systematic review of the literature. *Crit Care Med* 2010;38:2386–400
4. Heyland DK, Hopman W, Coo H, et al: Long-term health-related quality of life in survivors of sepsis. Short Form 36: A valid and reliable measure of health-related quality of life. *Crit Care Med* 2000;28:3599–605
5. Silveira LTYD, Silva JMD, Soler JMP, et al: Assessing functional status after intensive care unit stay: The Barthel Index and the Katz Index. *Int J Qual Health Care* 2018;30:265–70
AQ5 6. Parry SM, Nalamalapu SR, Nunna K, et al: Six-minute walk distance after critical illness: A systematic review and meta-analysis. *J Intensive Care Med* 2021;36:343–51
7. Bohannon RW, Crouch R: 1-minute sit-to-stand test: Systematic review of procedures, performance, and clinimetric properties. *J Cardiopulm Rehabil Prev* 2019;39:2–8
8. Kear BM, Guck TP, McGaha AL: Timed up and go (TUG) test: Normative reference values for ages 20 to 59 years and relationships with physical and mental health risk factors. *J Prim Care Community Health* 2017;8:9–13
9. Fisher S, Ottenbacher KJ, Goodwin JS, et al: Short Physical Performance Battery in hospitalized older adults. *Aging Clin Exp Res* 2009;21:445–52
10. Grundtvig M, Eriksen-Volnes T, Ørn S, et al: 6 min walk test is a strong independent predictor of death in outpatients with heart failure. *ESC Heart Fail* 2020;7:2904–11
11. Crook S, Frei A, Ter Riet G, et al: Prediction of long-term clinical outcomes using simple functional exercise performance tests in patients with COPD: A 5-year prospective cohort study. *Respir Res* 2017;18:112
12. Pavasini R, Guralnik J, Brown JC, et al: Short Physical Performance Battery and all-cause mortality: Systematic review and meta-analysis. *BMC Med* 2016;14:215
13. Lowrie EG, Curtin RB, LePain N, et al: Medical outcomes study Short Form-36: A consistent and powerful predictor of morbidity and mortality in dialysis patients. *Am J Kidney Dis* 2003; 41:1286–92
14. Wang DXM, Yao J, Zirek Y, et al: Muscle mass, strength, and physical performance predicting activities of daily living: A meta-analysis. *J Cachexia Sarcopenia Muscle* 2020;11:3–25
15. Burtin C, Clerckx B, Robbeets C, et al: Early exercise in critically ill patients enhances short-term functional recovery. *Crit Care Med* 2009;37:2499–505
16. Grimm A, Teschner U, Porzelius C, et al: Muscle ultrasound for early assessment of critical illness neuromyopathy in severe sepsis. *Crit Care* 2013;17:R227
17. Pinto-Ramos J, Moreira T, Costa F, et al: Handheld dynamometer reliability to measure knee extension strength in rehabilitation patients—A cross-sectional study. *PLoS One* 2022;17: e0268254
18. Pinto-Ramos J, Costa-Santos C, Costa F, et al: Reliability of point-of-care ultrasound for measuring quadriceps femoris muscle thickness. *Eur J Phys Rehabil Med* 2022;58:767–73
19. Denehy L, Skinner EH, Edbrooke L, et al: Exercise rehabilitation for patients with critical illness: A randomized controlled trial with 12 months of follow-up. *Crit Care* 2013;17:R156
20. Herridge MS, Tansey CM, Matté A, et al: Functional disability 5 years after acute respiratory distress syndrome. *N Engl J Med* 2011;364:1293–304
21. Vandenbroucke JP, von Elm E, Altman DG, et al: Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and elaboration. *PLoS Med* 2007;4:e297
22. Prinsen CAC, Mokkink LB, Bouter LM, et al: COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018;27:1147–57
23. Akoglu H: User's guide to correlation coefficients. *Turk J Emerg Med* 2018;18:91–3
24. Ratner B: The correlation coefficient: Its values range between +1/−1, or do they? *J Target Meas Anal Market* 2009;17:139–42
25. Husted JA, Cook RJ, Farewell VT, et al: Methods for assessing responsiveness: A critical review and recommendations. *J Clin Epidemiol* 2000;53:459–68
26. Parry SM, Denehy L, Beach LJ, et al: Functional outcomes in ICU—What should we be using?—An observational study. *Crit Care* 2015;19:127
27. Parry SM, Huang M, Needham DM: Evaluating physical functioning in critical care: Considerations for clinical practice and research. *Crit Care* 2017;21:249
28. Herridge MS, Cheung AM, Tansey CM, et al: One-year outcomes in survivors of the acute respiratory distress syndrome. *N Engl J Med* 2003;348:683–93
29. Núñez M, Nuñez E, Moreno JM, et al: Quadriceps muscle characteristics and subcutaneous fat assessed by ultrasound and relationship with function in patients with knee osteoarthritis awaiting knee arthroplasty. *J Clin Orthop Trauma* 2019;10:102–6
30. Nishihara K, Kawai H, Kera T, et al: Correlation of physical function with the thickness of multiple muscles of the quadriceps femoris in community-dwelling elderly individuals. *Clin Interv Aging* 2018;13:1945–51
31. Kwon S, Hartzema AG, Duncan PW, et al: Disability measures in stroke: Relationship among the Barthel Index, the Functional Independence Measure, and the Modified Rankin Scale. *Stroke* 2004;35:918–23
32. Maeo S, Shan X, Otsuka S, et al: Neuromuscular adaptations to work-matched maximal eccentric versus concentric training. *Med Sci Sports Exerc* 2018;50:1629–40