

FEASIBILITY, SAFETY AND RESPONSIVENESS OF HEALTH-RELATED FIELD-BASED PHYSICAL FITNESS TESTS IN ADULTS: THE ADULT-FIT PROJECT

VIABILIDAD, SEGURIDAD Y SENSIBILIDAD AL CAMBIO
DE RENDIMIENTO DE TEST DE CAMPO PARA EVALUAR
LA CONDICIÓN FÍSICA RELACIONADA CON LA SALUD EN
ADULTOS. PROYECTO ADULT- FIT

Tesis Doctoral Internacional



Programa oficial de doctorado en Ciencias de la Salud

Facultad de Ciencias de la Educación

Universidad de Cádiz

Carolina Cruz León

24 de Mayo de 2024

INTERNATIONAL DOCTORAL THESIS / TESIS DOCTORAL INTERNACIONAL

PROGRAMA DE DOCTORADO EN CIENCIAS DE LA SALUD

UNIVERSIDAD DE CÁDIZ

FEASIBILITY, SAFETY AND RESPONSIVENESS OF HEALTH-RELATED FIELD-BASED
PHYSICAL FITNESS TESTS IN ADULTS: THE ADULT-FIT PROJECT



VIABILIDAD, SEGURIDAD Y SENSIBILIDAD AL CAMBIO DE
RENDIMIENTO DE TEST DE CAMPO PARA EVALUAR LA CONDICIÓN
FÍSICA RELACIONADA CON LA SALUD EN ADULTOS. PROYECTO
ADULT- FIT

CAROLINA CRUZ LEÓN

24 de Mayo de 2024



DEPARTAMENTO DE LA EDUCACIÓN FÍSICA, PLÁSTICA Y MUSICAL
FACULTAD DE CIENCIAS DE LA EDUCACIÓN
UNIVERSIDAD DE CÁDIZ

**FEASIBILITY, SAFETY AND RESPONSIVENESS OF HEALTH-RELATED FIELD-BASED
PHYSICAL FITNESS TESTS IN ADULTS: THE ADULT-FIT PROJECT**

VIABILIDAD, SEGURIDAD Y SENSIBILIDAD AL CAMBIO DE RENDIMIENTO
DE TEST DE CAMPO PARA EVALUAR LA CONDICIÓN FÍSICA
RELACIONADA CON LA SALUD EN ADULTOS. PROYECTO ADULT- FIT

CAROLINA CRUZ LEÓN

Directores de Tesis [Thesis Supervisors]

Dr. José Castro Piñero

Catedrático de Universidad
Universidad de Cádiz
España

Dra. M^a Magdalena Cuenca García

Profesora Titular de Universidad
Universidad de Cádiz
España

En Puerto Real, 24 de Mayo de 2024



Dr. D. José Castro Piñero

Catedrático

Departamento de la Educación Física, Plástica y Musical

Facultad de Ciencias de la Educación

Universidad de Cádiz

JOSÉ CASTRO PIÑERO, CATEDRÁTICO DE LA UNIVERSIDAD DE CÁDIZ,
CERTIFICA:

Que la Tesis Doctoral titulada “VIABILIDAD, SEGURIDAD Y SENSIBILIDAD AL CAMBIO DE RENDIMIENTO DE TEST DE CAMPO PARA EVALUAR LA CONDICIÓN FÍSICA RELACIONADA CON LA SALUD EN ADULTOS. PROYECTO ADULT- FIT” que presenta Dña. Carolina Cruz León al superior juicio que del tribunal que designe la Universidad de Cádiz, ha sido realizada bajo mi dirección durante los años 2019-2024, siendo expresión de la capacidad técnica e interpretativa de su autora, en condiciones tan aventajadas que le hacen merecedora del Título de Doctora, siempre y cuando así lo considere el citado Tribunal.

A handwritten signature in blue ink, consisting of a large, stylized 'J' and 'C' followed by a horizontal line.

Fdo. José Castro Piñero

En Puerto Real, a 24 de mayo de 2024

En Puerto Real, a 24 de mayo de 2024



Dra. Dña. M^a Magdalena Cuenca García
Profesora Titular de Universidad

Departamento de la Educación Física, Plástica y Musical
Facultad de Ciencias de la Educación
Universidad de Cádiz

MARÍA MAGDALENA CUENCA GARCÍA, PROFESORA TITULAR DE LA
UNIVERSIDAD DE CÁDIZ, CERTIFICA:

Que la Tesis Doctoral titulada “VIABILIDAD, SEGURIDAD Y SENSIBILIDAD AL CAMBIO DE RENDIMIENTO DE TEST DE CAMPO PARA EVALUAR LA CONDICIÓN FÍSICA RELACIONADA CON LA SALUD EN ADULTOS. PROYECTO ADULT- FIT” que presenta Dña. Carolina Cruz León al superior juicio que del tribunal que designe la Universidad de Cádiz, ha sido realizada bajo mi dirección durante los años 2019-2024, siendo expresión de la capacidad técnica e interpretativa de su autora, en condiciones tan aventajadas que le hacen merecedora del Título de Doctora, siempre y cuando así lo considere el citado Tribunal.

A handwritten signature in blue ink, appearing to read 'M. Magdalena Cuenca García', with a stylized flourish at the end.

Fdo. M^a Magdalena Cuenca García

En Puerto Real, a 24 de mayo de 2024

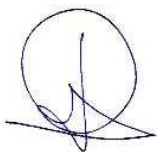
En Puerto Real, a 24 de mayo de 2024

Tanto la doctoranda, Carolina Cruz León, como sus directores de tesis, D. José Castro Piñero y Dña. M^a Magdalena Cuenca García:

Garantizamos que la presente Tesis Doctoral ha sido redactada bajo la supervisión de los directores de tesis y, hasta donde nuestro conocimiento alcanza, respetando la autoría de los estudios citados, al utilizarse sus resultados o publicaciones.

Directores de tesis

Doctoranda



Fdo. José Castro Piñero



Fdo. Carolina Cruz León



Fdo. M^a Magdalena Cuenca García

En Puerto Real, a 24 de mayo de 2024

En Puerto Real, a 24 de mayo de 2024

“Si he logrado ver más lejos, ha sido porque he subido a hombros de gigantes”

-Isaac Newton.

TABLE OF CONTENTS

Research projects and funding	17
Baseline studies	21
List of tables and figures	25
Abbreviations	30
ABSTRACT / RESUMEN	33
INTRODUCTION	42
HYPOTHESES	51
AIMS / OBJETIVOS	55
MATERIAL AND METHODS	59
Participants (study I-II-III)	61
Procedures (study I-II-III)	65
Statistical analysis (study I-II-III)	79
RESULTS	84
Study I	87
Study II	129
Study III	136
DISCUSSION	146
Study I	149
Study II	167
Study III	173
LIMITATIONS AND STRENGTHS	182
CONCLUSIONS / CONCLUSIONES	185
FUTURE RESEARCH DIRECTIONS	196
ANNEXES	203
Curriculum Vitae	205
Acknowledgments / Agradecimientos	212

RESEARCH PROJECTS AND FUNDING

RESEARCH PROJECTS AND FUNDING

The present International Doctoral Thesis was performed under the framework of the ADULT-FIT project*, which received the following funding:

- Ministry of Economy, Industry and Competitiveness in the 2017, call for R&D Projects of the State Program for Research, Development and Innovation Oriented to the Challenges of the Company; National Plan for Scientific and Technical Research and of Innovation 2017-2020 (DEP2017-88043-R).
- Ministry of Science, Innovation and Universities in the 2018, call for the Procedure for the granting of Grants for the acquisition of Scientific-Technical Equipment corresponding to the State Subprogram for Research Infrastructures and Scientific-Technical Equipment (State Plan R&D&I 2017-2020). Acquisition of Isokinetic equipment for Functional Assessment, Prevention, Rehabilitation and Treatment in Healthy People and People with Pathologies (EQC2018-004475-P).

*Visit <https://adultfitstudy.uca.es/> for more information.

BASELINE STUDIES

BASELINE STUDIES

The present International Doctoral Thesis are based on the following studies. These three studies are presented and grouped according to each content section throughout the Doctoral Thesis document, in the following order:

Study I. Feasibility and Safety of Field-Based Physical Fitness Test: A Systematic Review.

Cruz-León, C., Expósito-Carrillo, P., Sánchez-Parente, S., Jiménez-Iglesias, J., Borges Cosic, M., Cuenca-García, M. & Castro-Piñero, J. (2023). Feasibility and Safety of Field-Based Physical Fitness Test: A Systematic Review. *Sports Medicine - Open*. Under review, Q1.

Study II. Feasibility and Safety in health-related field-based physical fitness test in adult population: The ADULT-FIT project.

Cruz-León, C., Marín-Jiménez, N., Sánchez-Parente, S., Borges Cosic, M., Grao-Cruces, A., Castro-Piñero, J. & Cuenca-Garcia, M. (2024). Feasibility and Safety in health-related field-based physical fitness test in adult population: The ADULT-FIT project. *Journal of Science and Medicine in Sport*. Submitted, Q1.

Study III. Responsiveness of health-related field-based physical fitness tests in adults: The ADULT-FIT project.

Cuenca-Garcia, M., **Cruz-León, C.**, Jiménez-Iglesias, J., Sánchez-Parente, S., Víctor Jiménez-Segura., Francisco B. Ortega. & Castro-Piñero, J. (2024). Responsiveness of health-related field-based physical fitness tests in adults: The ADULT-FIT project. *British Journal of Sports Medicine*. Submitted, Q1.

LIST OF TABLES
AND
FIGURES

LIST OF TABLES

Table 1. Inclusion and exclusion criteria of the studies of the present Doctoral Thesis.

Table 2. Quality assessment list for field-based fitness test feasibility studies (*study I*).

Table 3. Quality assessment list for field-based fitness test safety studies (*study I*).

Table 4. Quality assessment of field-based fitness test feasibility studies (*study I*).

Table 5. Quality assessment of field-based fitness test safety studies (*study I*).

Table 6. Overview of field-based fitness test feasibility studies by study quality (*study I*).

Table 7. Overview of field-based fitness test safety studies by study quality (*study I*).

Table 8. Feasibility of the measurement protocol of field-based body composition and cardiorespiratory fitness tests (*study II*).

Table 9. Feasibility of the measurement protocol of field-based muscular and motor fitness tests (*study II*).

Table 10. Safety of the measurement protocol of field-based fitness test answered by evaluators and participants (*study II*).

Table 11. Participants baseline characteristics (*study III*).

Table 12. Pre-post differences for each groups and percentage of the population that is expected to respond to the intervention (*study III*).

Table 13. Levels of evidence of feasibility in field-based fitness tests by age group (*study I*).

Table 14. Levels of evidence of safety in field-based fitness tests by age group (*study I*).

Table 15. Items to establish feasibility reported by evaluators and participants (*study I*).

Table 15. (continued). Items to establish feasibility reported by evaluators and participants (*study I*).

Table 16. Items to establish safety reported by evaluators and participants (*study I*).

LIST OF FIGURES

Figure 1. PRISMA flow-chart showing the selection of original studies (*study I*).

Figure 2. Flow-chart of the studies II and III derived from ADULT-FIT project.

Figure 3. Participants categorized into responders / no responders to intervention in weight, BMI and waist circumference variables based on individual change (Δ), 80%CI and Smallest Worthwhile Change (SWC). *Arrow next to y-axis title represents direction of the positive effect for each variable. Participants were classified as likely responder if the CI lies completely over or below the SWC (horizontal dotted line) depending of the interpretation of positive effect. (study III).*

Figure 4. Participants categorized into responders / no responders to intervention in 20-m shuttle run, 6-min walk and 2-km walk tests based on individual change (Δ), 80%CI and Smallest Worthwhile Change (SWC). *Arrow next to y-axis title represents direction of the positive effect for each variable. Participants were classified as likely responder if the CI lies completely over or below the SWC (horizontal dotted line) depending of the interpretation of positive effect. (study III).*

Figure 5. Participants categorized into responders / no responders to intervention in 6-m gait speed, 2.45-m time up & go and 4×10-m shuttle run tests based on individual change (Δ), 80%CI and Smallest Worthwhile Change (SWC). *Arrow next to y-axis title represents direction of the positive effect for each variable. Participants were classified as likely responder if the CI lies completely over or below the SWC (horizontal dotted line) depending of the interpretation of positive effect. (study III).*

Figure 6. Participants categorized into responders / no responders to intervention in handgrip strength, standing long jump, 30-s sit-to-stand and front plank tests based on individual change (Δ), 80%CI and Smallest Worthwhile Change (SWC). *Arrow next to y-axis title represents direction of the positive effect for each variable. Participants were classified as likely responder if the CI lies completely over or below the SWC (horizontal dotted line) depending of the interpretation of positive effect. (study III).*

Figure 7. Proposed items for assessing the feasibility and the safety of field-based physical fitness tests in people of all ages (*study I*).

Figure 8. Proposal for an evidence-based health-related field physical fitness test battery for adults: The ADULT-FIT battery.

ABBREVIATIONS

ABBREVIATIONS

%BF: Body Fat percentage

%Change: Percentage of Change

%CI: Percentage of Confidence Interval

85%HRmax: 85 Percent from maximum heart rate

ACSM: American College of Sport Medicine

AFISAL: Actividad Física y Salud para Adultos

ALPHA: Assessing Levels of Physical Activity

AMRAP: As Many Repetitions as Possible

ANOVA: Analysis of Variance

BMI: Body Mass Index

CDC: Center for Disease Control and Prevention

CHMS: Canadian Health Measures Survey

CI: Confidence Interval

CPAFLA: Canadian Physical Activity, Fitness and Lifestyle Appraisal

CVD: Cardiovascular Disease

DOMS: Delayed-onset Muscle Soreness

HIIT: High Intensity Interval Training

HR: Heart Rate

HRmax: Maximum Heart Rate

IMC: Indice de Masa Corporal

ISAK: International Society for the Advancement of Kinanthropometry

MD: Mean Difference

MDC: Minimal Detectable Change

MeSH: Medical Subject Heading

MICT: Moderate-Intensity Continuous Training

PAR-Q: Physical Activity Readiness Questionnaire

PREFIT: Field-based FITness testing in PRESchool children

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analysis

PROSPERO: International Prospective Register of Systematic Reviews

QoL: Quality of Life

RM: Repetition Maximum

ROM: Range of Movement

RPE: Rated Perceived Exertion

SD: Standard Deviation

SPSS: Statistical Package for the Social Sciences

SWC: Smallest Worthwhile Change

TE: Typical Error

TRX: Total-body Resistance Exercise

VO₂: Volume of Oxygen Uptake

VO_{2max}: Maximum Volume of Oxygen Uptake

WHO: World Health Organisation

ABSTRACT/RESUMEN

ABSTRACT

Physical fitness is recognised as an important indicator of health in all age groups, such as children, adolescents, adults and the elderly. Research has indicated an inverse relation between cardiorespiratory, muscular fitness and motor fitness, with various diseases, morbidity and all-cause mortality. Therefore, the assessment of physical fitness is of substantial value as a means of prevention and identification. Although laboratory tests are accurate reference points for measuring fitness, their feasibility is restricted due to high costs, complexity of equipment, need for skilled personnel and time constraints, especially in studies of larger populations. Conducting field-based physical fitness tests offers a practical substitute, as these tests are simple to administer, require minimal equipment, are cost-effective and allow the assessment of numerous participants in a relatively short-term and simultaneous period of time. It is crucial that any assessment tool maintains qualities such as validity, reliability, feasibility, safety and responsiveness. Recent studies have added some of these qualities to existing health-related fitness field test batteries for preschoolers (PREFIT study) and children and adolescents (ALPHA study). However, there is no evidence-based field test battery that assesses physical fitness in adults. It is possible to find studies that include some of the mentioned qualities in various field tests, but not a complete battery. This background exposes the need to propose a valid, reliable, feasible, safe and responsive health-related field-based physical fitness test battery in adults to be used in a consensual way. This is how the proposal for the ADULT-FIT Project [Plan nacional I+D +i 2017-2020, (ref.: DEP2017-88043-R)] was born, whose **general aim** was: to propose a field-based test battery to assess health-related physical fitness in adults based on scientific evidence, valid, reliable, feasible, safe and responsive, according to sex, age and level of physical fitness.

This International Doctoral Thesis is part of the ADULT-FIT project and its **general aim** was: to propose a battery of field-based test to assess health-related physical fitness in adults based on scientific evidence, feasible, safe and responsive, according to sex, age and physical activity level.

Likewise, the **specific aims** were: (i) to comprehensively analyze the **feasibility** and **safety** of the existing health-related field-based physical fitness tests in adults, through a systematic review of the literature (*study I*); (ii) develop methodological **feasibility** and **safety** studies of those field-based physical fitness tests identified in the reviews that do not present sufficient or strong scientific evidence on their feasibility and safety in adults,

according to sex, age, and physical activity level (*study II*); and (iii) to determine the **responsiveness** of the health-related field-based physical fitness tests used in adults (*study III*).

The **main results** showed that there is **strong evidence** indicated that i) the 2-km walk and that 20-m shuttle run tests were feasible to assess cardiorespiratory fitness in adults and children and adolescents, respectively; ii) the handgrip strength test was **feasible** to assess isometric muscular strength in children and adolescents; iii) the standing long jump test was **feasible** to assess lower body explosive strength in children and adolescents; iv) only the 2-km walk test have shown strong evidence on **safety** in adults (*study I*). v) **Feasibility** items presented a successful answer of 100% in all tests (i.e., body composition (i.e., weight, height, neck, waist, and hip circumferences and triceps and subscapular skinfolds), cardiorespiratory fitness (i.e., 6-min walk, 2-km walk and 20-m shuttle run tests), muscular fitness (i.e., handgrip strength, standing long jump, 30-s sit-to-stand and front plank tests) and, motor fitness (i.e., 6-m gait speed, 2.45-m time up & go, 4x10 shuttle run and single-leg stand tests); vi) “participants evaluated” were the 100% of the sample, with an adequate “ratio participants/evaluators” and “time of preparation” (i.e., ~1 minute); vii) **safety** items presented a successful answer of 99% in all tests; viii) “heart rate” (HR) was evaluated only in cardiorespiratory fitness tests; the 97% of the participants reached the “85% HRmax” in the 20-m shuttle run test; ix) for “rate of perceived exertion”, the 51% and 88% of the participants reported a moderate and hard to maximum effort in the 30-s sit-to-stand and the front plank tests, respectively; x) for “tibial pain”, the 32% of the participants reported pain in the 2-km walk test and; xi) for “delayed-onset muscle soreness”, the 45% of the participants expressed some degree of muscle soreness, with the 95% of them reporting that cardiorespiratory fitness tests could be the cause (*study II*). xii) All tests were found to be **responsive** (all $P < 0.01$) after the exercise program, except the waist circumference measurement ($P = 0.09$); with an effects size of moderate to large (*Cohen’s d* > 0.50), except to the weight, body mass index, waist circumference and handgrip tests with trivial effect sizes ($P > 0.05$, *Cohen’s d* < 0.20); xiii) the highest proportion of individual **responders** based on the standard deviation was observed in the 4x10-m shuttle run, 30-s sit-to-stand, 6-m gait speed and 2.45-m time up & go tests and; xiv) the proportion of the population that was expected to **respond** to the intervention was higher than 85% in most of the field-based physical fitness tests evaluated, but $\leq 33\%$ in the weight, body mass index, waist circumference and handgrip

strength tests (*study III*).

This International Doctoral Thesis provides scientific and methodological evidence to propose **a battery of health-related field-based physical fitness tests in adults based on feasibility, safety and responsiveness**, according to sex, age, and physical activity level.

Previous systematic reviews and methodological studies developed in an International Doctoral Thesis, have provided scientific evidence on the validity (predictive and criterion-related) and reliability of health-related field-based physical fitness tests in adults according to sex, age and physical activity level.

RESUMEN

La condición física es reconocida como un importante indicador de salud en todos los grupos de edad, como niños, adolescentes, adultos y ancianos. La evidencia científica muestra una relación inversa entre la capacidad cardiorespiratoria, la fuerza muscular y la capacidad motora con diversas enfermedades, morbilidad y mortalidad por todas las causas.

Por lo tanto, la evaluación de la condición física tiene un valor sustancial como medio de prevención e identificación. Aunque las pruebas de laboratorio son un referente o *gold standard* para medir la condición física, su viabilidad se ve limitada por los elevados costes, la complejidad del equipo, la necesidad de personal cualificado y las limitaciones de tiempo, especialmente en estudios de poblaciones más numerosas. La realización de test de campo para evaluar la condición física, ofrece una alternativa más práctica, ya que estos test son sencillos de administrar, requieren un equipamiento mínimo, son rentables y permiten evaluar a numerosos participantes en un periodo de tiempo relativamente corto y de forma simultánea. Es crucial que cualquier herramienta de evaluación mantenga cualidades como la validez, la fiabilidad, la viabilidad, la seguridad y sensibilidad al cambio de rendimiento. Estudios recientes han añadido algunas de estas cualidades a las baterías de test de campo para evaluar la condición física relacionada con la salud existentes en preescolares (estudio PREFIT) y, en niños y adolescentes (estudio ALPHA). Sin embargo, no existe una batería de test de campo basada en evidencia científica que evalúe la condición física en adultos. Actualmente, es posible encontrar estudios que incluyan algunas de las cualidades mencionadas anteriormente en diversos test de campo, pero no una batería completa. Estos antecedentes exponen la necesidad de proponer una batería de test de campo válida, fiable, viable, segura y sensible al cambio de rendimiento en adultos para ser utilizada de forma consensuada. Es así, como nace la propuesta del Proyecto ADULT-FIT [Plan nacional I+D+i 2017-2020, (ref.: DEP2017-88043-R)], cuyo **objetivo general** fue: proponer una batería de test de campo para evaluar la condición física relacionada con la salud en adultos basada en la evidencia científica, válida, fiable, segura, viable y sensible al cambio de rendimiento, en función del sexo, la edad y el nivel de condición física.

La presente Tesis Doctoral Internacional se engloba dentro del proyecto ADULT-FIT y su **objetivo general** es: proponer una batería de test de campo para evaluar la

condición física relacionada con la salud en adultos basada en la evidencia científica, viable, segura y sensible al cambio de rendimiento, en función del sexo, la edad y el nivel de actividad física.

Asimismo, los **objetivos específicos** son: (i) analizar exhaustivamente la **viabilidad y seguridad** de los test de campo para evaluar la condición física relacionada con la salud existentes en adultos relacionados, mediante una revisión sistemática de la literatura (*estudio I*); (ii) desarrollar estudios metodológicos de **viabilidad y seguridad** de aquellos test de campo para evaluar la condición física relacionada con la salud identificados en las revisión que no presentan evidencia científica suficiente o sólida sobre su **viabilidad y seguridad** en adultos, en función del sexo, la edad y el nivel de actividad física (*estudio II*); y (iii) determinar la **sensibilidad al cambio de rendimiento** de los test de campo para evaluar la condición física relacionada con la salud utilizados en adultos (*estudio III*).

Los principales resultados mostraron que existe **fuerte evidencia** indicando que i) el test de 2-km andando y test de 20-m ida y vuelta, son **viables** para evaluar la capacidad cardiorrespiratoria en adultos, niños y adolescentes, respectivamente; ii) el test de fuerza de prensión manual es **viable** para evaluar la fuerza muscular isométrica en niños y adolescentes; iii) el test de salto de longitud a pies juntos es **viable** para evaluar la fuerza explosiva del tren inferior en niños y adolescentes; iv) sólo el test de 2-km andando ha mostrado **fuerte evidencia** sobre la **seguridad** en adultos (*estudio I*). v) Los ítems de **viabilidad** presentaron una respuesta satisfactoria del 100% en todos los test evaluados (es decir, composición corporal (es decir, peso, altura, circunferencias de cuello, cintura y cadera y, pliegues cutáneos del tríceps y subescapular), capacidad cardiorrespiratoria (es decir, los test de 6-min andando, 2-km andando y 20-m ida y vuelta), capacidad muscular (es decir, los test de fuerza de prensión manual, salto de longitud a pies juntos, 30-s sentarse y levantarse y plancha frontal) y, capacidad motora (es decir, los test de velocidad de la marcha de 6-m, 2.45-m levantarse ir y volver, 4x10-m ida y vuelta y, equilibrio en una pierna; vi) los "participantes evaluados" fueron el 100% de la muestra, con una "ratio participantes/evaluadores" y un "tiempo de preparación" adecuados (es decir, ~1 minuto); vii) los ítems de **seguridad** presentaron una respuesta satisfactoria del 99% en todos los test; viii) la "frecuencia cardiaca" sólo se evaluó en los test de capacidad cardiorrespiratoria; el 97% de los participantes alcanzó el "85% de FC_{máx}" en el test de 20-m ida y vuelta; ix) en cuanto al "índice de esfuerzo percibido", el 51% y el 88% de los

participantes declararon haber realizado un esfuerzo moderado y; de duro a máximo en el test de 30-s sentarse y levantarse y de plancha frontal, respectivamente; x) en cuanto al "dolor tibial", el 32% de los participantes manifestó dolor en el test de 2-km andando y; xi) en cuanto al "dolor muscular de aparición tardía", el 45% de los participantes manifestó cierto grado de dolor muscular, y el 95% de ellos señaló que los test de capacidad cardiorrespiratoria podían ser la causa (*estudio II*). xii) Se observó que todos los test mostraron **sensibilidad al cambio de rendimiento** (todos $P < 0,01$) después de realizar el programa de ejercicios, excepto la medición de la circunferencia de cintura ($P = 0,09$); con un tamaño del efecto de moderado a grande (d de Cohen $> 0,50$), excepto los test de peso, índice de masa corporal, circunferencia de cintura y fuerza de prensión manual, con un tamaño del efecto trivial ($P > 0,05$, d de Cohen $< 0,20$); xiii) la mayor proporción de individuos que mostraron **sensibilidad al cambio de rendimiento**, según la desviación estándar, se observó en los test de 4x10-m ida y vuelta, 30-s sentarse y levantarse, velocidad de 6-m andando y 2.45-m levantarse ir y volver y; xiv) la proporción de la población que se esperaba que mostrara **sensibilidad al cambio de rendimiento** fue superior al 85% en la mayoría de los test de campo para evaluar la condición física, salvo $\leq 33\%$ en los test de peso, índice de masa corporal, circunferencia de cintura y fuerza de prensión manual (*estudio III*).

La presente Tesis Doctoral Internacional aporta evidencia científica y metodológica para proponer una **batería de test de campo para evaluar la condición física relacionada con la salud en adultos viable, segura y sensible al cambio de rendimiento**, en función del sexo, edad y el nivel de actividad física.

Revisiones sistemáticas y estudios metodológicos previos desarrollados en una Tesis Doctoral Internacional, han proporcionado evidencia científica sobre la validez (predictiva y de criterio) y la fiabilidad de los test de campo para evaluar la condición física relacionada con la salud en adultos en función del sexo, edad y el nivel de actividad física.

INTRODUCTION

INTRODUCTION

Scientific evidence has shown that an adequate level of physical activity is a predictor of future health status as it helps to prevent and/or decrease the risk of developing certain diseases such as type 2 diabetes, hypertension, obesity, heart disease, some types of cancer, or clinically detected depression and several age-related diseases or pathologies in the adult population¹⁻⁴. The World Health Organisation (WHO) recognises physical inactivity as the fourth most important risk factor for mortality worldwide, accounting for 6-10% of all deaths. It is second only to high blood pressure (13%) and tobacco use (9%), and carries the same level of risk as hyperglycaemia (6%)^{5, 6}. Approximately 3.2 million people die each year due to insufficient physical activity⁷. Physical fitness is an integrated measure of all the functions and structures involved in performing physical activity⁸ is defined as “*a set of attributes or characteristics that people have or achieve that relates to the ability to perform physical activity*”. More specifically, health-related physical fitness is defined as “*those specific components of physical fitness that have a relationship with good health*”⁹. The American College of Sport Medicine (ACSM) compiles the components of health-related physical fitness as following: cardiorespiratory fitness, musculoskeletal fitness (related to muscular fitness and flexibility), motor fitness, and body composition⁹.

Cardiorespiratory fitness refers to the ability of the circulatory and the respiratory systems to supply oxygen during sustained physical activity⁹. It shows an inverse relationship with various health conditions, such as cardiovascular disease (CVD)^{10, 11}, osteoporosis¹², diabetes¹³, obesity^{14, 15}, metabolic syndrome¹⁶, and different forms of cancer^{17, 18}, and serves as a predictor of all-cause mortality^{19, 20}. Furthermore, from a psychological point of view, high cardiorespiratory fitness is associated with greater mental well-being^{21, 22}, better cognitive performance²³ and a decreased risk of Alzheimer's disease²⁴, as well as mental health problems such as anxiety and depression²⁵.

Musculoskeletal fitness (i.e., muscular fitness and flexibility) relates to the ability of the muscle to exert force²⁶, and to the component of physical fitness which refers to the ability to move a joint through its full range of motion (ROM) with ease²⁷. Muscular fitness also offers protection against all-cause mortality¹, improvement of bone health²⁸ and is inversely associated with weight gain, adiposity-related hypertension, metabolic syndrome onset²⁹ and, clinical manifestations of mental health problems^{4, 30, 31}. Furthermore, it is still

necessary found more evidence that support the health benefits of flexibility in adult population.

Motor fitness is considered as the performance aspect of physical fitness in daily activities which requires quick reactions, speed of movement, agility, coordination and balance³². Appropriate motor fitness level is associated with a minor decline in disability, less risk of falls, and reduced risk of frailty, depression³³, cognitive decline³⁴ and cardiovascular disease³⁵. The evidence supports that slower gait speed, worse agility and impaired balance are significantly associated with an increased risk of adverse health outcomes, such as falls and fear of falling, mobility disability, CVD risk or all-cause mortality in adult and older adults⁴. Based on this evidence, the assessment of physical fitness in all ages is relevant from a public health point of view.

Finally, body composition refers to the relative amount or percentage of body tissues (bone content, fat mass, muscular mass), that are related to health. The most common health-related measure is the total body fat percentage (%BF)⁹. In the adult population, the maintenance of proper levels of body composition, such as reduced body fat or increased muscle and bone mass, is associated with a better health outcome³⁶.

Therefore, assessing physical fitness is a vital method of disease diagnosis and prevention.

Physical fitness assessment

Health-related physical fitness can be assessed objectively by the use of laboratory devices (i.e., reference methods or gold standards) in a safe and controlled environment. Regardless, its use is limited due to the, qualified technicians, specific instruments and facilities are required. Health-related field-based physical fitness tests are shown to be a more efficient, low-cost, time-efficient option that allows to measure a large number of people at the same time and easy to administer³⁷.

A physical fitness battery that assesses health-related physical fitness in adults has to fulfil a series of fundamental criteria. The tests should be based on: their criterion-validity (i.e., output of the test correlates with the criterion measure, e.g., the gold standard)^{38, 39}, predictive validity (i.e., relationship with health outcomes)^{4, 40, 41}, reliability (i.e., reproducibility of values in repeated trials on the same individual)⁴², feasibility (i.e.,

degree of being conveniently done)⁴³, safety (i.e., number of health complications occurring during the testing procedure)⁴³, and responsiveness or longitudinal validity (i.e., ability of a test to detect changes over time) for use in adults⁴⁴. It is important that the health-related field-based physical fitness test battery should meet the criteria quality mentioned above to an efficient and appropriate use.

Based on scientific evidence (i.e., predictive and criterion-related validity, reliability, feasibility and safety), different field-based physical fitness batteries have been proposed to assess health-related physical fitness in preschoolers (i.e., the PREFIT battery⁴⁵), and in children and adolescents (i.e., the ALPHA battery⁴⁶). Also, a field-based physical fitness test battery has been proposed to assess health-related physical fitness for the elderly population (i.e., the Senior Fitness Test battery⁴⁷), although this battery is not based on scientific evidence, it is the most commonly used health-related field-based physical fitness battery for the elderly.

Currently, there are numerous test batteries that assess physical fitness in adults, such as the *UKK Health-Related Fitness Test Battery for Adults*⁴⁸, *The Canadian Physical Activity, Fitness and Lifestyle Appraisal (CPAFLA)*⁴⁹, the *Actividad Física y Salud para Adultos (AFISAL)*⁵⁰, and the *Eurofit Fitness Test Battery*⁵¹; however, there is no scientific evidence that they are valid, reliable, feasible, safe and responsive to be used in a consensual manner. To the best of our knowledge, there is no known battery of field-based physical fitness test that meets all these aforementioned criteria for the adult population.

The aim of the ADULT-FIT study supported by the Spain Ministry of Economy, Industry and Competitiveness was to propose a health-related field-based physical fitness test battery in adults, based on their predictive and criterion-related validity, reliability, feasibility, safety, and responsiveness.

There is available evidence on predictive validity^{7, 9, 37}, criterion-related validity⁵⁴ and reliability⁴⁴ of field-based physical fitness tests in adults. These studies have been developed in the ADULT-FIT study and have resulted in an International Doctoral Thesis, whose general aim was: to propose a battery of field-based tests to assess physical fitness related to health in adults based on scientific evidence, valid (i.e., predictive and criterion-related validity) and reliable, depending on sex, age, and physical activity level.

However, studies on feasibility, safety and responsiveness of health-related field-based physical fitness tests in adults are limited, with a lack of the methodological

homogeneity in the definition of the feasibility and safety items. Feasibility, safety and responsiveness are three important components when assessing physical fitness. However, several weaknesses have been identified, including scarcity of studies addressing the feasibility and safety of field-based physical fitness tests. Furthermore, those studies that do evaluate these aspects lack standardized criteria, potentially resulting in discrepancies across findings. Similarly, and regrettably, the assessment of responsiveness is not usually considered when proposing a health-related field-based physical fitness test in health adults.

Feasibility allows researchers to evaluate whether ideas and outcomes can be adapted to be relevant and sustainable⁵³. Otherwise, safety refers to the assessment of risk or harm that is likely to occur before, during and after the performance of the tests. A test is considered safe when it does not cause physical complications to participants⁵⁴. Finally, the responsiveness of a field-based physical fitness test is relevant in outcome studies as well as practical application in different settings (e.g., clinical, sport club, etc.). A health-related field-based physical fitness test battery can be deemed to be valid, reliable, feasible and safe but if the test is unable to detect the relatively small changes in response to an intervention cannot be considered useful and appropriate in practice.

Given that it is known which health-related field-based physical fitness tests are valid and reliable in adults, it would be advisable to also know which health-related field-based physical fitness tests are feasible, safe and responsive in this population, in order to be able to propose, in the future, a health-related field-based physical fitness test battery that meet the aforementioned precepts.

REFERENCES

1. García-Hermoso A, Cavero-Redondo I, Ramirez-Velez R, Ruiz JR, Ortega FB, Lee DC, Martínez-Vizcaino V. Muscular Strength as a Predictor of All-Cause Mortality in an Apparently Healthy Population: A Systematic Review and Meta-Analysis of Data From Approximately 2 Million Men and Women. *Arch Phys Med Rehabil.* 2018;99(10):2100-13 e5.
2. Organization WH. WHO guidelines on physical activity and sedentary behaviour. 2020.
3. Barry V, Baruth M, Beets MW, Durstine JL, Liu J, N. BS. Fitness vs. fatness on all-cause mortality: a meta-analysis. *Prog Cardiovasc Dis.* 2014;56:382-90.
4. Marín-Jiménez N, Cruz-León C, Perez-Bey A, Conde-Caveda J, Grao-Cruces A, Aparicio VA, ..., Cuenca-García M. Predictive Validity of Motor Fitness and Flexibility Tests in Adults and Older Adults: A Systematic Review. *Journal of Clinical Medicine.* 2022;11:328.
5. Prevention WHO. Report of the Formal Meeting of Member States to conclude the work on the comprehensive global monitoring framework, including indicators, and a set of voluntary global targets for the prevention and control of noncommunicable diseases. . Geneva: WHO. 2012:3-9.
6. Guthold R, Stevens GA, Riley LM, Bull FC. Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1·9 million participants. *The Lancet Global Health.* 2018;6(10):e1077-e86.
7. Santos AC, Willumsen J, Meheus F, Ilbawi A, Bull FC. The cost of inaction on physical inactivity to public health-care systems: a population-attributable fraction analysis. *The Lancet Global Health.* 2023;11(1):e32-e9.
8. Castillo-Garzón MJ, Ortega-Porcel FB. Improvement of physical fitness as anti-aging intervention. *Medicina clinica.* 2005;124(4):146-55.
9. Medicine ACoS. ACSM's health-related physical fitness assessment manual 2021.
10. Ortega FB, Lavie CJ, Blair SN. Obesity and cardiovascular disease. *Circulation research.* 2016;118(11):1752-70.
11. Barry VW, Caputo JL, Kang M. The joint association of fitness and fatness on cardiovascular disease mortality: a meta-analysis. *Progress in cardiovascular diseases.* 2018;61(2):136-41.
12. Howe TE, Rochester L, Neil F, Skelton DA, Ballinger C. Exercise for improving balance in older people. *Cochrane database of systematic reviews.* 2011(11).
13. Tarp J, Støle, A. P., Blond, K.,, Grøntved A. Cardiorespiratory fitness, muscular strength and risk of type 2 diabetes: a systematic review and meta-analysis. *Diabetologia.* 2019;62:1129-42.
14. McAuley PA, Beavers KM. Contribution of cardiorespiratory fitness to the obesity paradox. *Progress in cardiovascular diseases.* 2014;56:434-40.
15. Lavie CJ, McAuley PA, Church TS, Milani RV, Blair SN. Obesity and cardiovascular diseases: implications regarding fitness, fatness, and severity in the obesity paradox. *Journal of the American College of Cardiology.* 2014;63(14):1345-54.
16. LaMonte MJ, Barlow CE, Jurca R, Kampert JB, Church TS, Blair SN. Cardiorespiratory fitness is inversely associated with the incidence of metabolic syndrome: a prospective study of men and women. *Circulation.* 2005;112(4):505-12.
17. Ezzatvar Y, Ramírez-Vélez R, Saez de Asteasu ML, Martínez-Velilla N, Zambom-Ferraresi F, Lobelo F, et al. Cardiorespiratory fitness and all-cause mortality in adults diagnosed with cancer systematic review and meta-analysis. *Scandinavian Journal of Medicine & Science in Sports.* 2021;31(9):1745-52.
18. Zhu N, Jacobs, D. R., Schreiner, P. J., Launer, L. J., Whitmer, R. A., Sidney, S., ..., Bryan RN. Cardiorespiratory fitness and brain volume and white matter integrity: The CARDIA Study. *Neurology.* 2015;84:2347-53.

19. Blair SN, Kohl HW, Paffenbarger RS, Clark DG, Cooper KH, Gibbons LW. Physical fitness and all-cause mortality: a prospective study of healthy men and women. *Jama*. 1989;262(17):2395-401.
20. Farrell SW, Fitzgerald SJ, McAuley PA, Barlow CE. Cardiorespiratory fitness, adiposity, and all-cause mortality in women. *Medicine and science in sports and exercise*. 2010;42(11):2006-12.
21. Guo L, Chen J, Yuan W. The effect of HIIT on body composition, cardiovascular fitness, psychological well-being, and executive function of overweight/obese female young adults. *Frontiers in Psychology*. 2023;13:1095328.
22. Hallgren M, Kandola A, Stubbs B, Wallin P, Andersson G, Ekblom-Bak E. Associations of exercise frequency and cardiorespiratory fitness with symptoms of depression and anxiety-a cross-sectional study of 36,595 adults. *Mental Health and Physical Activity*. 2020;19:100351.
23. Ortega FB, Lee D-c, Sui X, Kubzansky LD, Ruiz JR, Baruth M, et al. Psychological well-being, cardiorespiratory fitness, and long-term survival. *American journal of preventive medicine*. 2010;39(5):440-8.
24. Boots EA, Schultz, S. A., Oh, J. M., Larson, J., Edwards, D., Cook, D., ..., Okonkwo OC. Cardiorespiratory fitness is associated with brain structure, cognition, and mood in a middle-aged cohort at risk for Alzheimer's disease. *Brain imaging and behavior*. 2015;9:639-49.
25. Willis BL, Gao A, Leonard D, DeFina LF, Berry JD. Midlife fitness and the development of chronic conditions in later life. *Archives of internal medicine*. 2012;172(17):1333-40.
26. Corbin CB, Pangrazi RP, Franks BD. Definitions: Health, fitness, and physical activity. President's Council on Physical Fitness and Sports Research Digest. 2000.
27. Bandy WD, Sanders B. Therapeutic exercise for physical therapist assistants: Lippincott Williams & Wilkins; 2007.
28. Torres-Costoso A, López-Muñoz, P., Martínez-Vizcaíno, V., Álvarez-Bueno, C.,, Cavero-Redondo I. Association between muscular strength and bone health from children to young adults: a systematic review and meta-analysis. *Sports Medicine*. 2020;50:1163-90.
29. De Lima TR, Martins PC, Guerra PH, Santos Silva DA. Muscular strength and cardiovascular risk factors in adults: A systematic review. *The Physician and Sportsmedicine*. 2021;49(1):18-30.
30. Tucker LA. Muscular strength and mental health. *Journal of Personality and Social Psychology*. 1983;45(6):1355.
31. Ortega FB, Silventoinen K, Tynelius P, Rasmussen F. Muscular strength in male adolescents and premature death: cohort study of one million participants. *Bmj*. 2012;345:e7279.
32. Goodway JD, Ozmun JC, Gallahue DL. Understanding motor development: Infants, children, adolescents, adults: Jones & Bartlett Learning; 2019.
33. Briggs R, Carey D, Claffey P, McNicholas T, Donoghue O, Kennelly SP, Kenny RA. Do differences in spatiotemporal gait parameters predict the risk of developing depression in later life? *Journal of the American Geriatrics Society*. 2019;67(5):1050-6.
34. Doi T, Tsutsumimoto K, Nakakubo S, Kim M-J, Kurita S, Hotta R, Shimada H. Physical performance predictors for incident dementia among Japanese community-dwelling older adults. *Physical Therapy*. 2019;99(9):1132-40.
35. Navarete-Villanueva D, Gómez-Cabello A, Marín-Puyalto J, Moreno LA, Vicente-Rodríguez G, Casajús JA. Frailty and physical fitness in elderly people: a systematic review and meta-analysis. *Sports Medicine*. 2021;51:143-60.
36. Abarca-Gómez L, Abdeen ZA, Hamid ZA, Abu-Rmeileh NM, Acosta-Cazares B, Acuin C, et al. Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128·9 million children, adolescents, and adults. *The lancet*. 2017;390(10113):2627-42.

37. Castro-Pinero J, Artero EG, Espana-Romero V, Ortega FB, Sjostrom M, Suni J, Ruiz JR. Criterion-related validity of field-based fitness tests in youth: a systematic review. *British Journal of Sports Medicine*. 2009;44(13):934-43.
38. Castro-Piñero J, Marin-Jimenez, N., Fernandez-Santos, J. R., Martin-Acosta, F., Segura-Jimenez, V., Izquierdo-Gomez, R., ... & Cuenca-Garcia, M. Criterion-Related Validity of Field-Based Fitness Tests in Adults: A Systematic Review. *Journal of clinical medicine*. 2021;10:3743.
39. Marin-Jimenez N, Cruz-Leon C, Sanchez-Oliva D, Jimenez-Iglesias J, Caraballo I, Padilla-Moledo C, Castro-Piñero J. Criterion-Related Validity of Field-Based Methods and Equations for Body Composition Estimation in Adults: A Systematic Review. . *Current Obesity Reports*. 2022;1-14.
40. Harber M, Kaminsky L, Arena R, Blair S, Franklin B, Myers J, al. e. Impact of cardiorespiratory fitness on all-cause and disease-specific mortality: advances since 2009. . *Progress in cardiovascular diseases*. 2017;60:11-20.
41. Soysal P, Hurst C, Demurtas J, Firth J, Howden R, Yang L, et al. Handgrip strength and health outcomes: Umbrella review of systematic reviews with meta-analyses of observational studies. . *Journal of Sport and Health Science*. 2021;10:290-5.
42. Cuenca-Garcia M, Marin-Jimenez N, Perez-Bey A, Sánchez-Oliva D, Camiletti-Moiron D, Alvarez-Gallardo IC, Castro-Piñero J. Reliability of Field-Based Fitness Tests in Adults: A Systematic Review. *Sports Medicine*. 2022;1-20.
43. Cruz-León C, Marín-Jiménez N, Sánchez-Parente S, Borges Cosic M, Grao-Cruces A, Castro-Piñero J, Cuenca-Garcia M. Feasibility and Safety in health-related field-based physical fitness tests in adult population: The ADULT-FIT project. *Journal of Sports Sciences*. 2024;Under Rev.
44. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of clinical epidemiology*. 2007;34-42.
45. Ortega FB, Cadenas-Sánchez C, Sánchez-Delgado G, Mora-González J, Martínez-Téllez B, Artero EG, et al. Systematic review and proposal of a field-based physical fitness-test battery in preschool children: the PREFIT battery. *Sports medicine*. 2015;45:533-55.
46. Ruiz JR, Castro-Piñero J, España-Romero V, Artero EG, Ortega FB, Cuenca MM, et al. Field-based fitness assessment in young people: the ALPHA health-related fitness test battery for children and adolescents. *British journal of sports medicine*. 2011;45(6):518-24.
47. Rikli RE, Jones CJ. Development and validation of criterion-referenced clinically relevant fitness standards for maintaining physical independence in later years. *The Gerontologist*. 2013;53(2):255-67.
48. Suni JH, Oja P, Laukkanen RT, Mülunpalo SI, Pasanen ME, Vuori IM, et al. Health-related fitness test battery for adults: aspects of reliability. *Archives of physical medicine and rehabilitation*. 1996;77(4):399-405.
49. Physiology CSfE. The Canadian Physical Activity, Fitness & Lifestyle Approach: CSEP-Health-Related Appraisal & Counseling Strategy. Health Canada Ottawa; 2003.
50. Rodriguez FA, Gusi N, Valenzuela A, Nàcher S, Nogués J, Marina M. Valoració de la condició física saludable en els adults (I): Antecedents i protocols de la bateria AFISAL-INEFC. *Apunts Educació física i esports*. 1998;2(52):54-77.
51. Oja P, Tuxworth B. Eurofit for Adults: Assessment of Health-Related Fitness. Tampere, Finland: Council of Europe. Committee for the Development of Sport. UKK Institute for Health Promotion Research. 1995.
52. Castro-Piñero J, Marin-Jimenez, N., Fernandez-Santos, J. R., Martin-Acosta, F., Segura-Jimenez, V., Izquierdo-Gomez, R., ..., Cuenca-Garcia M. Criterion-Related Validity of Field-Based Fitness Tests in Adults: A Systematic Review. *Journal of clinical medicine*. 2021;10:3743.
53. Bowen DJ, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, Fernandez M. How we design feasibility studies. *American journal of preventive medicine*. 2009;36:452-7.

54. Suni JH, Miilunpalo SI, Asikainen TM, Laukkanen RT, Oja P, Pasanen ME, Vuori IM. Safety and feasibility of a health-related fitness test battery for adults. *Physical Therapy*. 1998;78:134-48.

HYPOTHESES

HYPOTHESES

The starting hypotheses of this Doctoral Thesis are listed below:

- I.** There is a lack of scientific evidence regarding the feasibility and safety in children, adolescents, adults, and older adults of each health-related field-based physical fitness tests, as well as the absence of a homogeneous methodology.
- II.** The health-related field-based physical fitness tests evaluated, based on scientific evidence, to assess the different components of physical fitness related to health are feasible and safe in adult population, regardless sex, age and physical activity level.
- III.** The health-related field-based physical fitness tests evaluated after a multicomponent exercise intervention of 12 weeks are responsive in adult population, regardless sex and age.

AIMS/OBJETIVOS

AIMS

The overall aim of the present International Doctoral Thesis was to propose a battery of field-based tests to assess health-related physical fitness in adults based on scientific evidence, feasible, safe and responsive according to sex, age and physical activity level.

The specific aims of the present International Doctoral Thesis are the following:

Study I. To comprehensively analyze the **feasibility** and **safety** of the existing health-related field-based physical fitness tests in adults, through a systematic review of the literature.

Study II. Develop methodological **feasibility** and **safety** studies of those field-based physical fitness tests identified in the reviews that do not present sufficient or strong scientific evidence on their feasibility and safety in adults, according to sex, age, and physical activity level.

Study III. To determine the **responsiveness** of the health-related field-based physical fitness tests used in adults.

OBJETIVOS

El objetivo general de la presente Tesis Doctoral Internacional ha sido: proponer una batería de test de campo para evaluar la condición física relacionada con la salud en adultos basada en la evidencia científica, viable, segura y que sea sensible al cambio de rendimiento, en función del sexo, la edad y el nivel de actividad física.

Los objetivos específicos de la presente Tesis Doctoral Internacional fueron:

Estudio I. Analizar exhaustivamente la viabilidad y seguridad de los test de campo para evaluar la condición física relacionada con la salud existentes en adultos, mediante una revisión sistemática de la literatura.

Estudio II. Desarrollar estudios metodológicos de viabilidad y seguridad de aquellos test de campo para evaluar la condición física relacionada con la salud identificados en la revisión que no presentan evidencia científica suficiente o sólida sobre su viabilidad y seguridad en adultos, en función del sexo, la edad y el nivel de actividad física.

Estudio III. Determinar la sensibilidad al cambio de rendimiento de los test de campo para evaluar la condición física relacionada con la salud utilizados en adultos.

MATERIAL
AND
METHODS

MATERIAL AND METHODS

The present Doctoral Thesis is composed of three studies, one of them designed as a systematic review (study I), the second as a methodological study (study II), and the third as an intervention study (study III) derived from the ADULT-FIT study.

Participants

Study I included healthy people of all ages (i.e., children and adolescents, adults and older). The sample size in the feasibility studies was 4,186 participants, of which 2,124 were children and adolescents, 1,857 were adults and 205 were older adults. The sample size in the safety studies was 2,665 participants, of which 808 were children and adolescents and 1,857 were adults

Study II included participants from Cádiz (Spain), homogeneously distributed by sex (49.5% females and 50.5% males), age (18-34 years, 35-49 years, and 50-64 years), and physical activity level (non-active and active). The final sample was 390 participants.

Study III included a total of 62 non-active participants aged 18 to 64 years were randomized into the intervention ($n=31$) or control group ($n=31$).

The inclusion and exclusion criteria of the three studies are shown in **Table 1**, and the flow chart of the study I is shown in **Figure 1**; and the studies II and III is shown in **Figure 2**.

Table 1. Inclusion and exclusion criteria of the studies of the present Doctoral Thesis.

Study	Inclusion	Exclusion
I. Feasibility and Safety of Field-Based Physical Fitness Test: A Systematic Review (<i>study I</i>)	i) All ages; ii) participants: the study population was based on a generally healthy population, who did not present any injury, physical and/or mental disabilities, irrespective of body mass index (BMI), diabetes or other cardiovascular risks (i.e., hypertension, hypercholesterolemia, lipid profiles, glucose levels, insulin sensitivity). In the course of this review, we found that some studies sampled healthy subjects and those with pathologies together. In those cases, we looked whether these studies performed stratified analyses by groups, separating the population with pathologies from the rest, if so, we included in the study; iii) study design: original studies or systematic reviews/meta-analyses; iv) language: articles were only published in English or Spanish and v) topic criterion: original studies analysing feasibility and safety in field-based physical fitness tests.	Studies that examined the feasibility and safety of field-based physical fitness tests designed for use in exclusive clinical settings.
II. Feasibility and Safety in health-related field-based physical fitness test in adult population: The ADULT-FIT project (<i>study II</i>)	i) Adults (i.e., 18 to 64 years old), ii) not suffer from any physical or mental illness that impairs their ability to participate in physical activity, iii) willingness to perform all the tests comprising the study and iv) be able to read and understand the informed consent form and the purpose of the study.	i) severe or end-stage illness, ii) myocardial infarction three months prior to study beginning, iii) unsteady cardiovascular condition, iv) medical indication which prevents the tests from being performed and v) any injury or circumstance that makes it incompatible with the correct performance of the tests.
III. Responsiveness of Health-Related Field-Based Physical Fitness Tests in Adults: The ADULT-FIT project (<i>study III</i>)	i) Being an adult (i.e., 18 to 64 years old) and non-active, ii) not having physical or mental illness that prevents from doing physical activity, iii) have intention to carry out all the tests included in the study and iv) being able to read and understand the informed consent as well as the aim of the study.	i) having acute or terminal illness, ii) myocardial infarction three months before starting the study, iii) unstable cardiovascular disease, iv) medical prescription that prevents the performance of the tests and v) injury or circumstance that makes it impossible to carry out the tests correctly.

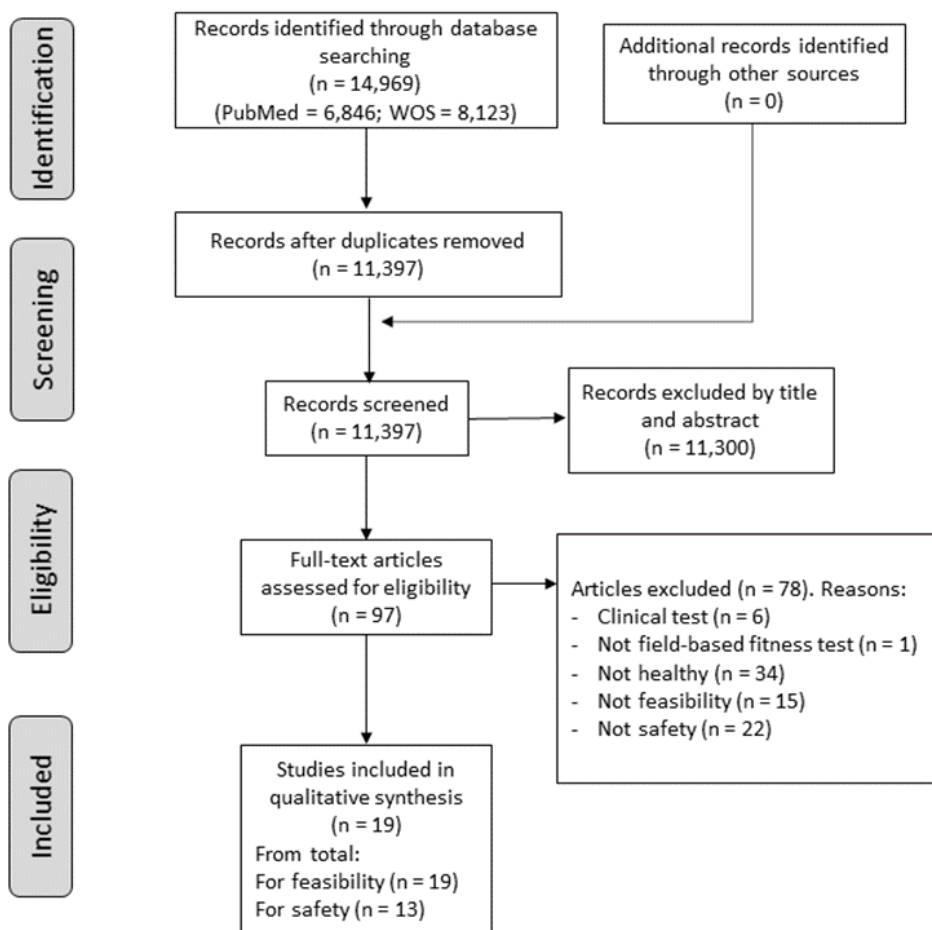


Figure 1. PRISMA flow-chart showing the selection of original studies (*study I*).

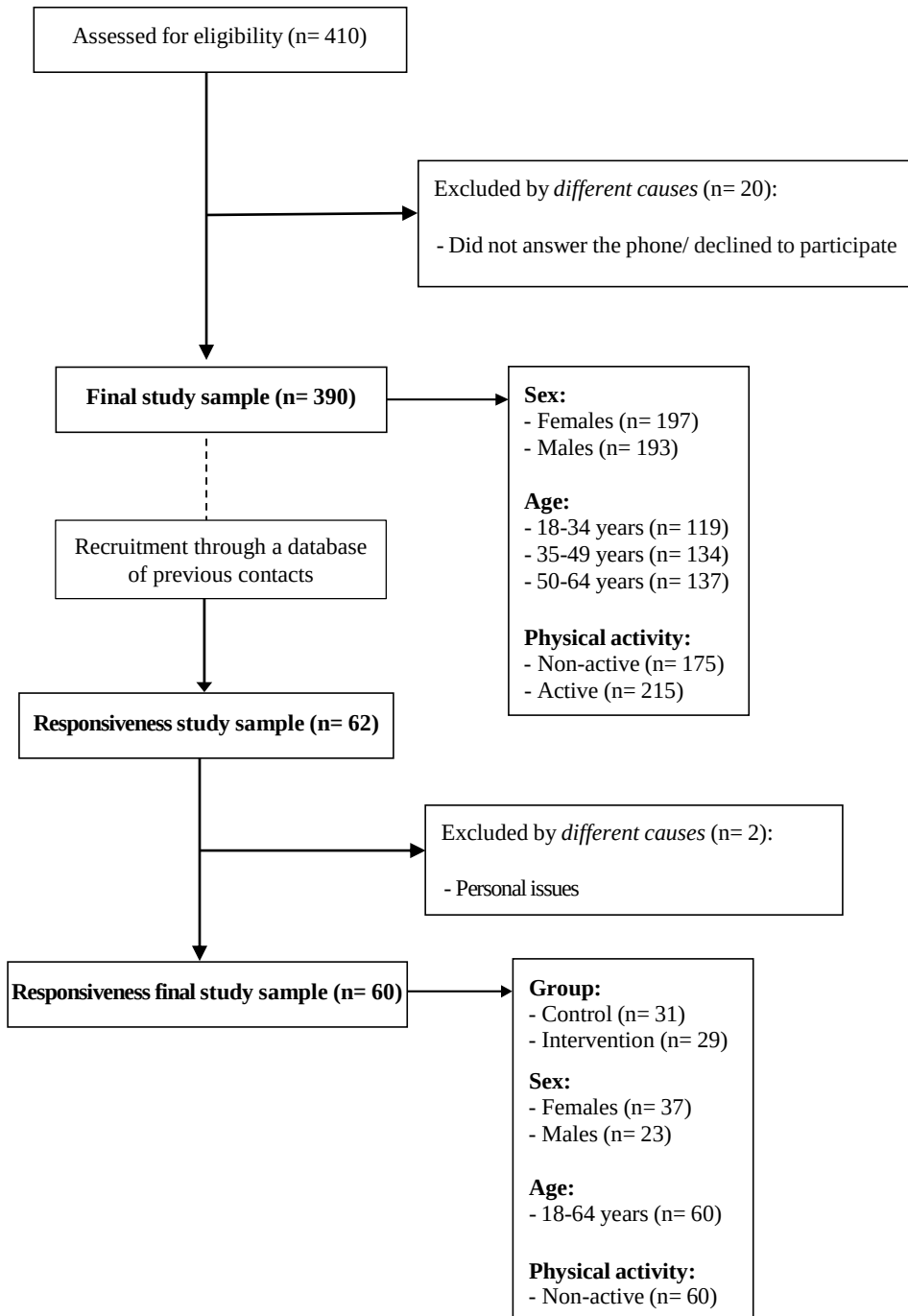


Figure 2. Flow-chart of the studies II and III derived from ADULT-FIT project.

Procedures

Study I. Feasibility and Safety of Field-Based Physical Fitness Test: A Systematic Review.

Protocol and Registration

This systematic review was developed according to Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines¹. The review protocol was registered in the International Prospective Register of Systematic Reviews (PROSPERO reference number, CRD42022298276)².

Data Sources and Search Strategy

A systematic search was performed of the electronic databases MEDLINE (via PubMed) and Web of Science (all databases) from database inception to 31 January 2023. The search terms used were those related to the following topics: (i) feasibility: feasibility, viability; (ii) safety: safety; and (iii) physical fitness components: physical fitness, muscular strength, range of motion, articular, postural balance, physical endurance, cardiorespiratory fitness, cardiovascular fitness, aerobic fitness, aerobic capacity, maximal oxygen consumption, VO2max, motor fitness, running speed and agility. The three search topics were combined with the Boolean operators ‘AND’ (inter topics) and ‘OR’ (intra topics). When using PubMed, we included Medical Subject Heading (MeSH) terms to enhance the power of the search. The same search strategy and combination of terms was repeated in Web of Science but without using MeSH terms or their equivalent as a similar option does not exist for this database.

Study selection

The study selection process was carried out in three stages. Firstly, records were identified via the PubMed and Web of Science databases and imported into the EndNote software (Desktop version 20.1). Duplicate files were then removed, initially automatically by the software and secondary by visual checking. Secondly, the titles and abstracts of the search results were checked for eligibility. Finally, those full-text reports that were apparently eligible were read for final inclusion or exclusion in the review.

The results of all stages were compared between two researchers. If no consensus was reached between the two investigators (< 5%), a third investigator made the final decision. Furthermore, the literature search was complemented by manual review of the reference lists obtained from the selected studies. The authors were not blinded to the selected articles, as the reviewers who performed the quality assessment were familiar with the literature.

Data Extraction

Two researchers independently extracted the following information from each eligible original study according to the standardized form: (i) the author's name, (ii) participants (sex and number), (iii) age of participants, (iv) field-based physical fitness test, (v) statistical methods (vi) main outcome and (vii) conclusions. Any disagreement in the extracted data was discussed among the studies until a consensus was reached.

Risk of Bias Assessment

Risk of bias was assessed by two independent investigators. They obtained 94% agreement on the quality assessment before consensus, and 100% agreement after resolving discrepancies in a consensus meeting, with a third investigator. Risks of bias of the original studies were determined, modifying the original quality assessment checklist developed by Castro-Piñero et al.,³, based on three criteria: (i) appropriate number of participants, (ii) appropriate description of the study population and (iii) appropriate number of item reported, to assess feasibility/safety (see **Table 2 and 3**). Each criterion was scored from 0 to 2 (where 2 is the best score). The scores for each criterion were summarised (0-6) for all studies. Studies were classified as very low quality (score ≤ 2), low quality (score 3-4) or high quality (score >5).

Table 2. Quality assessment list for field-based fitness test feasibility studies.

Grading system parameter	Grade	Criterion
Number of study subjects	0	$n \leq 10$
	1	$n = 11-50$
	2	$n \geq 51$
Description of the study population with respect to age, sex, health status, fitness level, ethnicity, physical activity patterns, body composition, etc.	0	Less items than required for grade 1
	1	At least age, sex, health status, and fitness level
	2	More items than required for grade 1
Number of items reported: time spent preparing and developing the test, percentage of participants completing the test, and self-reporter feasibility questions to participants/evaluators	0	Less items than required for grade 1
	1	At least time spent preparing and developing the test, AND percentage of participants completing the test OR self-reporter questions to participants/evaluators
	2	Time spent preparing and developing the test, percentage of participants completing the test and self-reporter questions to participants/evaluators
Rating for total score: High quality = 5–6; Low quality = 3–4; Very low quality = 0–2		

Table 3. Quality assessment list for field-based fitness test safety studies.

Grading system parameter	Grade	Criterion
Number of study subjects	0	$n \leq 10$
	1	$n = 11-50$
	2	$n \geq 51$
Description of the study population with respect to age, sex, health status, fitness level, pubertal status, ethnicity, physical activity patterns, body composition, etc.	0	Less items than required for grade 1
	1	At least age, sex, health status, and fitness level
	2	More items than required for grade 1
Number of questions/ items reported: adverse events (i.e., instrument allergy, sick feeling, pain, musculoskeletal injuries, falls...), delayed-onset muscle soreness, heart rate, Borg rating of shortness of breath, etc.	0	Less than required for grade 1 or 2
	1	At least adverse events
	2	At least adverse events, AND delayed-onset muscle soreness OR heart rate OR Borg rating of shortness of breath
Rating for total score: High quality = 5-6; Low quality = 3-4; Very low quality = 0– 2		

Level of Evidence

To formulate a final conclusion on the feasibility and safety of the field-based physical fitness tests, we determined three levels of evidence³: (i) strong evidence: consistent findings in three or more high-quality studies; (ii) moderate evidence: consistent findings in two high quality studies; and (iii) limited or inconclusive evidence: consistent findings in multiple low-quality studies, inconsistent results found in multiple high-quality studies, or results based on one single study.

Study II. Feasibility and Safety in health-related field-based physical fitness test in adult population: The ADULT-FIT project.

Participants were recruited through leaflet, local newspapers and social media from Cadiz. A total of 390 (50.5% women) participants were included in the study. The total sample was homogeneously distributed by sex, age (18–34, 35–49, and 50–64 years), and physical activity level (non-active and active). All participants provide voluntarily written informed consent to be part of the present study. The study was approved by the Committee for Research of Cadiz, Spain.

After providing written informed consent and being informed of the protocol to be carried out, they signed the “Physical Activity Readiness Questionnaire”⁴. Participants were initially classified as active/non-active according to the World Health Organization’s recommendations for adults⁵, once they answered the question: *how many days (in a typical week) do you practice physical activity/exercise or some sport, of at least moderate intensity, lasting at least 60 minutes per day?*

Health-related field-based physical fitness tests assessment

The health-related physical fitness components measured and the tests related to them were: body composition (i.e., weight, height, triceps and subscapular skinfolds, and neck, waist and hip circumferences), cardiorespiratory fitness (i.e., the 6-min walk, 2-km walk, and 20-m shuttle run tests), muscular fitness (i.e., the handgrip, standing long jump, 30-s sit-to-stand and front plank tests) and motor fitness (i.e., the 6-m gait speed, 2.45-m time up & go, 4x10-m shuttle run and single-leg stand test (eyes open and closed)).

The tests assessments were distributed in the following order: day 1) Body composition measurements, the 4x10-m shuttle run, handgrip strength, and 2-km walk tests; day 2) the single-leg stand, standing long jump, front plank, and 6-min walk tests; day 3) the 2.45-m time up & go, 6-m gait speed, 30-s sit-to-stand, and 20-m shuttle run tests.

Prior to the field-based physical fitness testing sessions, all participants underwent a standardized 10-minute warm-up. All participants received detailed instructions for the tests and were actively encouraged to do their best in each test. In addition, participants were instructed to rest for 24 hours before the assessments, to maintain their nutritional and hydration habits and wear appropriate sport clothes (i.e., sports clothing and sports shoes that allow correct movement and sliding).

The tests were carried out in three non-consecutive days, with at least 72 hours apart, in an indoor/outdoor facility under appropriate conditions (i.e., free space with good lighting and ventilation; no objects restricting the performance of the tests; flat, clean, undamaged, non-slip and non-abrasive surface and; specific test equipment in perfect and non-injurious condition) and were assessed and supervised by researchers with experience in field-based physical fitness testing. The procedure/protocol of field-based physical fitness tests were:

Body composition

The body composition measurements included weight, height, neck, waist and hip circumferences, subscapular and triceps skinfolds, following the protocol described by the International Society for the Advancement of Kynanthropometry⁶. The protocols established by the Centers for Disease Control and Prevention were followed to assess neck circumference⁷. All measurements were taken by the same trained evaluator (to avoid intra-evaluator variability) twice and the mean between measurements was calculated. The evaluator was of the same sex as the participant.

Motor fitness

6-m gait speed test

The participants had to cover a distance of 6 metres walking as fast as possible without stopping. This test has a total distance of 10 metres. In the first 2 metres the participant will accelerate his/her walking to the maximum, and for the next 6 metres, the participant will walk as fast as possible (recorded distance). Afterwards, the participant will have two metres to decelerate. The test was evaluated and supervised by two evaluators, one of them located at the starting line of the course (point 1) and the second one, lateral to the 6 metres line (point 2), measuring the time of execution⁸. One attempt was performed.

2.45-m time up & go test

The test assesses mobility skills, gait speed and functional ability. Participants started the test sitting on a chair with both arms crossed over the chest. The participant had to get up from the chair, walk 2.45-m (8 feet) as fast as possible, turn, and return to seated position. The test was evaluated and supervised by two evaluators; one of them gave the start signal and controlled the total time required for the participant to complete the out and back. The second one held the chair of the participant at the time of return to prevent possible incidents. One attempt was performed.

4x10-m shuttle run test

The aim of this test was to assess the speed of movement, coordination and agility. It is a run and turn test at maximum speed (4x10-m) between two parallel lines on the floor 10 metres apart. On the start line there was a sponge (B) and on the opposite line there were two sponges (A, C). At the signal, the participant ran without a sponge as fast as possible to the other line crossing the line with both feet, picking up the sponge A and returning to the starting line. Quickly changing sponge A for sponge B and running back as fast as possible to the opposite line and changing sponge B for sponge C, returning to the starting line and finishing the test. The test was evaluated and supervised by one evaluator located at the start/end line, who recorded the time spent by the participant⁹. Participants had to perform two non-consecutive attempts and the best attempt was recorded.

Single-leg stand test (eyes open and closed)

The aim of the test was to assess the postural balance over a reduced support area.

The test was performed by supporting on one foot and raising the opposite leg in a tandem position with their arms at their sides parallel to their trunk and with their eyes open during 30 seconds; and subsequently with the other leg. Participants also performed the test with eyes closed for 30 seconds per leg. The test was evaluated and supervised by two evaluators. Both recorded every time the participant lost balance and put his/her foot on the floor, or if the participant performed any support on the leg that was steady. When the test was performed with eyes closed, one of the evaluators recorded each time the participant opened his eyes and the second recorded the losses in balance. It was the duty of the evaluators (when required), to ask participants to maintain the balance position by not supporting the tandem leg on the floor or on the balance leg¹⁰. The test was performed once per leg, first with eyes open and then with eyes closed. Two non-consecutive attempts of each round were evaluated.

Muscular fitness

Handgrip strength test

The aim of the test was to assess the maximal isometric strength according to the protocol described by Ruiz et al.,¹¹, where in the case of women, the grip had to be adjusted according to the size of the hand. Participants had to squeeze the dynamometer continuously and gradually for at least 3 seconds. The elbow should be in full extension and slightly away from the body to avoid contact of the dynamometer with any other part of the body. The test was evaluated and supervised by one evaluator. Two non-consecutive attempts were made with each hand and the best result at each grip span was retained.

Standing long jump test

The aim of the test is to assess the lower body explosive strength through an energetic jump forward in an attempt to reach the greatest distance possible without losing balance. Participants had to jump with both feet from a marked line to the front, trying to reach as far as possible with both feet together. The test was evaluated and supervised by one evaluator. Two non-consecutive attempts were performed and the best attempt was recorded¹².

30-s sit-to-stand test

The aim of this test was to assess the lower body endurance strength to sit down and stand up for 30 seconds as fast as possible while maintaining a correct execution

throughout the test. Participants were seated in the centre of a chair, back straight, with their arms crossed on their chests, legs approximately shoulder-width apart, feet placed on the floor at a 90-degree angle to the knees and slightly in front of the chair legs to avoid possible impact when rising up. The participants had to stand up and sit down as many times as possible for 30 seconds. The test was evaluated and supervised by two evaluators, one of them holding the chair and registering the rate of perceived exertion using a Borg scale (1-10 points)¹⁴ at the end of the test, and the second one checking the time and number of repetitions performed by the participant¹⁵. One attempt was performed.

Front plank test

The aim of the test was to assess the trunk endurance strength by performing a frontal plank for as long as possible. The starting position of the test was with the knees, toes and elbows resting on the floor, the palms of the hands in contact with the floor, maintaining a neutral position of the cervical curvature and forming a straight line between the lateral area of the acromion, greater trochanter and malleolus. The forearm was at a 90° angle to the upper arm and the hands were placed shoulder width apart. Prior to the test, an elastic cord was attached horizontally to a pair of vertical sticks, which was moved to a height that was at the same level as the participant's iliac crest, and then adjusted 3 centimetres below. This adjustment served as a reference for objective control of hip displacement during the test. Each time the hip touches the reference line, a verbal warning from the assessors would be given so that the participant recovers the correct position. The test would be finished when the hip was not maintained at the required level after receiving two consecutive warnings or when the participant reached the maximum effort. At the end of the test, total time and the RPE was registered. The test was evaluated and supervised by two evaluators, who were positioned one on each side of the participant to ensure that the participant maintained the correct position and tracked the time performed with a hand-held stopwatch (Casio HS-EV-1RET Digital, Casio, Tokyo, Japan)¹⁶. One attempt was performed.

Cardiorespiratory fitness

The cardiorespiratory field-based physical fitness tests assessed were: i) 6-min walk¹⁷, ii) 2-km walk¹⁸ and iii) 20-m shuttle run test¹⁹. All the tests were performed in groups of 10 participants approximately.

6-min walk test

The aim of the 6-min walk test was to walk and reach the longest distance possible in 6 minutes without running and keeping a steady pace. The participants started the test simultaneously, walking in a space of 20 x 30 metres. Two evaluators monitored the time and subsequently recorded the total distance covered, and the heart rate (HR) of the participants at the end of the test. One attempt was performed.

2-km walk test

For the 2-km walk, the aim of the test was to complete the 2 kilometres distance walking in the shortest time possible without running and keeping a steady pace. The participants started the test simultaneously, walking in a space of 20 x 30 metres. Two evaluators recorded the number of laps (participants had to complete 20 laps) during the 2-km walk test, total time, HR and tibial pain through a numeric rating scale (0-10 points) at the end of the test. One attempt was performed.

20-m shuttle run test

For the 20-m shuttle run test, participants had to run between 2 lines 20 metres apart following the rhythm of an audible signal. The initial speed of the signal was 8.5 km/h, which was increased by 0.5 km/h/min (1 minute equaled 1 stage), so the participants had to rise their running pace to reach the corresponding line. The test was finished when participants reached their maximum effort and abandoned the test or failed to reach the end lines at the same time as the audio signals on two consecutive occasions. Two evaluators registered the stages performed by each participant and the HR at the end of the test. One attempt was performed. The HR was recorded using the activity bracelet Xiaomi Mi Band 4 (Xiaomi Inc, Beijing, China) validated in adults²⁰.

Feasibility assessment

The feasibility items proposed, based on scientific evidence²¹, were: i) “appropriate facilities” (i.e., the test facility is in optimal conditions in terms of lighting and ventilation and enough space), ii) “appropriate equipment” (i.e., the audio instruments and speakers for the assessment of the tests are working properly), iii) “appropriate sports clothes” (i.e., correct sport clothes and shoes of the participants), iv)

“test’s instruction understood” (i.e., the participant understand the test’s instructions), v) “easy to administer” (i.e., participants perform the test without difficulty and correctly, related to: technical performance, modifications to the original protocol or familiarisation session with technical performance; and evaluator-administered tests, recalled data without difficulty and without the need for profound learning about the items to be assessed), vi) “participants evaluated/not evaluated” (i.e., participants who perform the test and participants who did not perform), vii) “ratio participants/evaluators” (i.e., number of evaluators per test to evaluate a certain number of participants), viii) “time of preparation” (i.e., required time to prepare the test), and ix) “time of performance” (i.e., required time for the participant to complete de test).

Safety assessment

The safety items proposed, based on scientific evidence²¹, were: i) “instrument allergy” (i.e., manual dynamometer); ii) “adverse events” (i.e., falls, injury and/or sick feeling) that may occur during the performance of the test; iii) “heart rate” (registered at the end of the 6-min walk, 2-km walk, and 20-m shuttle run tests); iv) “rate of perceived exertion” (assessed in the 30-s sit-to-stand and front plank tests) using the Borg scale of 1 to 10 for self-perception of general exertion¹⁴, where 1 means “very light”, 2-3 “light”, 4-6 “moderate”, 7-8 “hard”, 9 “very hard” and 10 “maximum effort”; v) “tibial pain” (registered at the end of the 2-km walk test) in a scale of 0 to 10 for subjective self-perception²² (0 means “no pain”, 1-3 “low”, 4-6 “moderate” and 7-10 “severe” tibial pain); and vi) “delayed-onset muscle soreness” (DOMS)²³, through a self-reported questionnaire 48 hours post assessment of the field-based physical fitness tests. The items recorded in the DOMS questionnaire were: a) “did you experienced DOMS posterior to assessment of the field-based fitness test?” (yes/no answers); b) if the answer was yes, “how much pain do you feel?” (options were: “very low”, “low”, “moderate”, “much” and “severe”); c) “where you had the muscle soreness?” (options were: “left arm”, “right arm”, “both arms”, “lower zone back”, “upper zone back”, “lower and upper zones back”, “left leg”, “right leg”, “both legs”, “both arms and legs”, “lower zone back and both legs”, “upper zone back and both legs” and “other”); d) “what test do you think caused your muscle soreness?” (options were: “the muscular fitness tests” (i.e., the handgrip strength, standing long jump, 30-s sit-to-stand and front plank tests) or “the cardiorespiratory fitness tests” (i.e., the 6-min walk, 2-km walk, and 20-m shuttle run tests)); e) “you have difficulty performing activities of daily living due to muscle

soreness?” (options were: “none”, “when picking up a heavy object at the level of your abdomen/trunk”, “when picking up objects from a piece of furniture above my head”, “when descending stairs/climbing stairs”, “when bending down to pick up an object” and “other”); and f) “if in the previous question you have indicated other, please specify which on” (open answer).

Study III. Responsiveness of health-related field-based physical fitness test in adult population: The ADULT-FIT project.

Sample size was calculated according to the T-test for differences between differences two independent mean (two groups) with a one-sided, a level of significance of 5%, an effect size of 0.80, a power of 90%, and considering a dropout rate of 10% (G*power 3.1.9.6, Düsseldorf, Germany). Participants were recruited through leaflets, local newspapers, and social media. A total of 62 non-active participants aged 18 to 64 years were recruited and randomized into the intervention group (n=31) or the control group (n=31). All interested volunteers provided written informed consent to participate in the present study. The study was approved by the Committee for Research of Cádiz.

Participants were initially classified as active/non-active when following/not following World Health Organization recommendations for adults⁵. The following self-reported question was asked: *how many days (in a typical week) do you practice physical activity/exercise or some sport, of at least moderate intensity, lasting at least 50 min per day?*

Protocol Design

Physical fitness was assessed using the field-based physical fitness tests at baseline (pre-test) and after (post-test) 12-weeks in all participants. The tests were carried out in three non-consecutive days, with at least 72 hours apart, in an indoor/outdoor facility under appropriate conditions (i.e.: space, surface, equipment and sports clothing), and were assessed and supervised by researchers with experience in field-based physical fitness testing. The tests assessments were distributed in the following order: i) *1st day*: body weight, height, waist circumference, 4x10-m shuttle run, handgrip strength and 2-km walk tests; ii) *2nd day*: standing long jump, front plank and 6-min walk tests and; iii) *3rd day*: 2.45-m time up & go, 6-m gait speed, 30-s sit-to-stand

and 20-m shuttle run tests. The procedure/protocol of the health-related field-based physical fitness tests used for this study are fully described in **study II**.

Post-test evaluation was conducted 72 hours after the conclusion of the last season of the intervention period under the same conditions, with researchers following the same protocol as in the pre-test evaluation.

Participants in the intervention group engaged in a 12-week multicomponent exercise; while those in the control group maintained their usual routine. Possible changes were then registered.

Training Program

The exercise program included 3 supervised sessions per week of multicomponent exercise (i.e., including aerobic, muscle strengthening and motor fitness activities)²⁴⁻²⁶ for 12 weeks, based on the guidelines proposed by the American College of Sports Medicine for the prescription of physical exercise in healthy adults²⁷. Each session lasted 60 minutes, 10 minutes of warm-up, 45-minutes of multicomponent exercise and 5 minutes of cool-down.

Exercise intensity was checked individually by determination of the rating of perceived exertion on the Borg Scale (1-10 points)¹⁴ and by heart rate using the activity bracelet (Xiaomi Mi Band 4) validated in adults (Xiaomi Inc, Beijing, China)²⁰. The target was set between 5 and 8 (perceived “somewhat hard” or “hard” exertion) and between 60% and 91% of the reserve heart rate.

On one hand, the aim of the muscle strengthening and motor fitness exercises was to develop strength, stability and coordination and speed of movement. It was based on basic strength work aimed at fundamental motor patterns using free weights, elastic bands and self-loading. The exercises were: pulls and pushes (both horizontal and vertical) as well as hip/knee dominant exercises and lumbo-pelvic stabilization work. In terms of external load, the exercises progressed throughout the sessions in qualitative terms. They became increasingly complex and demanding from a motor point of view, by progressively including unilateral work, cross-chain and homolateral execution, combinations of patterns (e.g., hip dominance plus push-off) and more complex combinations (e.g., hip dominance plus push-off and hip dominant plus vertical thrust). The exercises also involved increasingly demanding bases of support from a stabilization

point of view (e.g., transitioning from doing an exercise while standing to performing it in a split stance or on one leg). From a quantitative point of view (i.e., kilograms to be displaced), the participants were urged to mobilise increasingly heavier implements, thicker rubber bands and to adopt more demanding angles in the exercises with self-loads.

On the other hand, the aim of aerobic exercises was to develop cardiorespiratory fitness; for this purpose, intermittent work dynamics such as High Intensity Interval Training (HIIT), As Many Repetitions as Possible (AMRAP) and Fractional Running methods were established. HIIT and AMRAP consisted of alternating periods of high intensity (relative to the capacity of each participant) based on different types of movements (frontal running, lateral movements, jumps, etc.) and strength exercises with very light loads. The evolution of the external load involved accumulating progressively more effective time of advanced cardiovascular work, increasing from 12-15 to 30-35 min (modifying the total duration of the intensity periods, as along with their numbers). To enhance motivation and adherence, variations were introduced in both the exercises and training methods.

Statistical analyses

Study II: Feasibility and Safety in health-related field-based physical fitness test in adult population: The ADULT-FIT project.

An acceptable level of feasibility occurs when the proposed items (i.e., “appropriate facilities”, “appropriate equipment”, “appropriate sport cloths”, “test’s instruction understood”, and “easy to administer”) shows $\geq 95\%$ positively answers from participants²³, presented as frequency and percentage. The items “participants evaluated/not evaluated” and “ratio participant/evaluator” are expressed as frequency. The “time of preparation” is presented as mean and standard deviation (SD; seconds) and the “time of performance” as minimum and maximum time ranges (minutes or seconds) for each field-based physical fitness test, with the exception of those tests whose time of performance is established by protocol.

An acceptable level of safety occurs when the proposed items (i.e., “instrument allergy” and “adverse events”) shows $\geq 99\%$ positively answers from participants²³, presented as frequency and percentage. The “HR” is presented in mean (SD) of beats per minute (bpm). Moreover, the “85% HRmax” is presented as the frequency and percentage of participants who reached or exceed their 85% age-predicted maximum HR. The maximum HR was calculated according to Tanaka et al.,²⁸ (i.e., $208 - 0.7 \times \text{age}$). The “RPE”, “tibial pain” and “DOMS” items are presented as frequency and percentage. All the statistical procedures were conducted using SPSS v. 25.0 software for Windows.

Study III: Responsiveness of health-related field-based physical fitness tests in adults: The ADULT-FIT project.

The distribution of each variable was examined with Kolmogorov-Smirnov normality test. All data were normally distributed. To compare the baseline variables between groups, the one-way ANOVA test was used.

To compare pre-post differences, we used pairwise comparison ANOVA for each group separately. Percentage of change (% Change) for each variable was calculated as:

$$\% \text{Change} = (\text{mean (Post-test value)} - \text{mean (Pre-test value)} / \text{mean (Pre-test value)}) * 100.$$

Cohen's d was computed to quantify the magnitude of the effect size as standardized mean difference. The criteria to interpret the magnitude of the effect size were as follows: <0.2 trivial, 0.2-0.5 small, >0.5-0.8 moderate and >0.8 large²⁹. Statistical analyses were performed using the Statistical Package for Social Sciences (IBM SPSS Statistics for Mac Os, version 29.0; Armonk, NY). Differences were considered significant when $P < 0.05$.

Responder analysis was performed using a two-step approach: 1) pre-post change and an 80% confidence interval (CI) were calculated for each participant (80% CI classifies individuals with 90% certainty and it had been proposed as a default value³⁰). CIs calculations were adjusted for sample size by using the following formula:

$$CI = TE \times \sqrt{2} \times qt(1-(\%CI / 100), dfs)$$

Where TE represents the typical error (i.e., standard deviation of the pre-post change divided by $\sqrt{2}$), qt() is the quantile function of the central t-distribution, %CI the desired width of the CI (80 in our analysis) and dfs the degrees of freedom (i.e., sample size - 1).

Then, the smallest worthwhile change (SWC) was calculated for every variable as $0.2 \times$ the baseline between individual standard deviation. Finally, participants were

classified as *likely responder* if the CI lies completely over or below the SWC depending of the interpretation of positive effect for each of the variables³¹. 2) We assume that each studied variable follows a normal distribution in the population with the mean observed score change and the standard deviation of individual responses (σ_{IR}) as mean and standard deviation, respectively. σ_{IR} was calculated with the following formula:

$$\sigma_{IR} = \sqrt{SD_{INT}^2 - SD_{CON}^2}$$

where SD_{INT}^2 is the square of the calculated standard deviation of the observed score change from the intervention group, and SD_{CON}^2 is the square of the calculated standard deviation of the observed change scores from the control group³². Next, the percentage of responders was calculated as the area of the normal distribution that lies beyond the SWC³³. This the analyses were performed with the R software environment for statistical computing³⁴.

REFERENCES

1. Page MJ, McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ..., Moher D. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *International journal of surgery*. 2021;88:105906.
2. Booth A, Clarke M, Dooley G, Ghera D, Moher D, Petticrew M, Stewart L. The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. *Systematic reviews*. 2012;1:1-9.
3. Castro-Pinero J, Artero EG, Espana-Romero V, Ortega FB, Sjostrom M, Suni J, Ruiz JR. Criterion-related validity of field-based fitness tests in youth: a systematic review. *British Journal of Sports Medicine*. 2009;44(13):934-43.
4. Gledhill N. Update of the PAR-Q and PARmed-X. Ottawa, Ontario Canadian Society for Exercise Physiology. 2002.
5. Bull FC, Al-Ansari SS, Biddle S, Borodulin K, Buman MP, Cardon G, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *British journal of sports medicine*. 2020;54(24):1451-62.
6. Marfell-Jones M, Olds T, Stewart A, Carter L. ISAK accreditation handbook 2006.
7. Control CfD, Prevention. National Health and Nutrition Examination Survey Anthropometry Procedures Manual. . 2017.
8. Dargent-Molina P, Schott A, Hans D, Favier F, Grandjean H, Baudoin C, et al. Separate and combined value of bone mass and gait speed measurements in screening for hip fracture risk: results from the EPIDOS study. *Osteoporosis International*. 1999;9:188-92.
9. Ruiz J, Artero E, Ortega F, Chillón P, MJ GR, Mora J, et al. ALPHA-fitness test battery: health-related field-based fitness tests assessment in children and adolescents. *Nutrición hospitalaria*. 2011;26(6):1210-4.
10. Wolinsky FD, Miller DK, Andresen EM, Malmstrom TK, Miller JP. Reproducibility of physical performance and physiologic assessments. . *Journal of Aging and Health*. 2005;17:111-24.
11. Ruiz-Ruiz J, Mesa JL, Gutiérrez A, Castillo MJ. Hand size influences optimal grip span in women but not in men. *The Journal of hand surgery*. 2002;27(5):897-901.
12. Castro-Piñero J, Ortega FB, Artero EG, Girela-Rejón MJ, Mora J, Sjöström M, Ruiz JR. Assessing muscular strength in youth: usefulness of standing long jump as a general index of muscular fitness. *Journal of Strength & Conditioning Research*. 2010;24:1810-7.
13. Bowen DJ, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, Fernandez M. How we design feasibility studies. *American journal of preventive medicine*. 2009;36:452-7.
14. Borg GA. Psychophysical bases of perceived exertion. *Medicine and science in sports and exercise*. 1982;14(5):377-81.
15. Jones CJ, Rikli RE, Beam WC. A 30-s chair-stand test as a measure of lower body strength in community-residing older adults. *Research quarterly for exercise and sport*. 1999;70(2):113-9.
16. De Blaiser C, De Ridder R, Willems T, Danneels L, Roosen P. Reliability and validity of trunk flexor and trunk extensor strength measurements using handheld dynamometry in a healthy athletic population. *Physical Therapy in Sport*. 2018;34:180-6.
17. Mänttari A, Suni J, Sievänen H, Hsu P, Vähä-Ypyä H, Valkeinen H, et al. Six-minute walk test: a tool for predicting maximal aerobic power (VO 2 max) in healthy adults. *Clinical physiology and functional imaging*. 2018;38(6):1038-45.
18. Oja P, Laukkanen R, Pasanen M, Tyry T, Vuori I. A 2-km walking test for assessing the cardiorespiratory fitness of healthy adults. 1991;12:356-62.

19. Leger LA, Mercier D, Gadoury C, Lambert J. The multistage 20 metre shuttle run test for aerobic fitness. *Journal of sports sciences*. 1988;6:93-101.
20. El-Amrawy F, Nounou MI. Are currently available wearable devices for activity tracking and heart rate monitoring accurate, precise, and medically beneficial? *Healthcare informatics research*. 2015;21(4):315-20.
21. Cruz-Leon C, Exposito-Carrillo P, Sanchez-Parente S, Jimenez-Iglesias J, Borges-Cosic M, Cuenca-Garcia M, Castro-Piñero J. Feasibility and Safety of Field-Based Physical Fitness Test: A Systematic Review. *Sports Medicine - Open*. 2024;Under Review.
22. Karcioglu O, Topacoglu, H., Dikme, O., & Dikme, O. A systematic review of the pain scales in adults: which to use?. *The American journal of emergency medicine*. 2018;36:707-14.
23. Espana-Romero V, Artero EG, Jimenez-Pavon D, Cuenca-Garcia M, Ortega FB, Castro-Pinero J, et al. Assessing health-related fitness tests in the school setting: reliability, feasibility and safety; the ALPHA Study. *Int J Sports Med*. 2010;31(7):490-7.
24. Edwards JJ, Griffiths M, Deenmamode AH, O'Driscoll JM. High-Intensity Interval Training and Cardiometabolic Health in the General Population: A Systematic Review and Meta-Analysis of Randomised Controlled Trials. *Sports Med*. 2023;1-11.
25. Lopez P, Radaelli R, Taaffe DR, Newton RU, Galvão DA, Trajano GS, Pinto RS. Resistance training load effects on muscle hypertrophy and strength gain: Systematic review and network meta-analysis. *Medicine and science in sports and exercise*. 2021;53:1206.
26. Markov AHL, Chaabene H. Effects of Concurrent Strength and Endurance Training on Measures of Physical Fitness in Healthy Middle-Aged and Older Adults: A Systematic Review with Meta-Analysis. *Sports Med*. 2023;53:437-55.
27. Liguori G, American College of Sports M. ACSM's guidelines for exercise testing and prescription. . USA2021.
28. Tanaka H, Monahan, K. D., & Seals, D. R. Age-predicted maximal heart rate revisited. *Journal of the american college of cardiology*. 2001;37:153-6.
29. Cohen J. Statistical power analysis for the behavioral sciences. . Revised edition ed: Academic press.; 2013.
30. Hopkins WG. A Spreadsheet for Monitoring an Individual's Changes and Trend. *Sportscience*. 2017;21:5-9.
31. Bonafiglia JT, Nelms MW, Preobrazenski N, LeBlanc C, Robins L, Lu S, et al. Moving beyond threshold-based dichotomous classification to improve the accuracy in classifying non-responders. *Physiological reports*. 2018;6:e13928.
32. Hopkins WG. Individual responses made easy. *Journal of applied physiology*. 2015;118:1444-6.
33. Swinton PA, Hemingway BS, Saunders B, Gualano B, Dolan E. A statistical framework to interpret individual response to intervention: paving the way for personalized nutrition and exercise prescription. *Frontiers in Nutrition*. 2018;5:41.
34. RCore TEAM. R: A language and environment for statistical computing. 2016.

RESULTS

RESULTS

The results section of the present Doctoral Thesis is shown as a compilation of three scientific studies.

Study I. Feasibility and Safety of Field-Based Physical Fitness Test: A Systematic Review.

Figure 1 presents the flow-chart of retrieved and selected studies and systematic reviews/meta-analysis. The literature search reported 14,969 studies, from PubMed and Web of Science databases. After the removal of duplicate references (3,572 studies), and the screening of titles and abstracts (11,300 studies), we excluded 14,872 studies. A total of 97 full-text studies were assessed for eligibility. Thereafter, 78 were excluded due to different reasons. Finally, a total of 19 original studies addressed the feasibility and 13 of them also studied the safety of field-based physical fitness tests.

One study was classified as "very low quality" (score <2) for feasibility¹. Four studies were classified as "low quality" (score 3-4) for feasibility²⁻⁵ and 2 for safety^{5, 6}. Fourteen studies were classified as "high quality" (score >5) for feasibility and 11 for safety (see **Table 4 and 5**).

Table 4. Quality assessment of field-based fitness test feasibility studies.

Study	Fitness Component	Field-based fitness test	Number of study subjects	Description of the study population	Statistical analysis	Total score
Borel et al., 2010	Cardiorespiratory fitness	6-min step	1	2	0	3
Bruggeman et al., 2020	Cardiorespiratory fitness	45-s squat	1	2	1	4
Bruggeman et al., 2020	Cardiorespiratory fitness	3-min step	1	2	1	4
Anderson & Dal Corso, 2016	Cardiorespiratory fitness	Chester step	1	2	1	4
Anderson & Dal Corso, 2016	Cardiorespiratory fitness	Modified incremental step	1	2	1	4
Anderson & Dal Corso, 2016	Cardiorespiratory fitness	6-min walk	1	2	1	4
Lamonedá et al., 2020	Cardiorespiratory fitness	20-m shuttle run music	2	1	1	4
Lamonedá et al., 2020	Cardiorespiratory fitness	20-m shuttle run	2	1	1	4
Oja et al., 1991	Cardiorespiratory fitness	2-km walk	2	2	1	5
Amado-Pacheco et al., 2019	Cardiorespiratory fitness	20-m shuttle run	2	2	1	5
Cadenas-Sanchez et al., 2014	Cardiorespiratory fitness	20-m shuttle run	2	2	1	5
Cadenas-Sánchez et al., 2016	Cardiorespiratory fitness	20-m shuttle run	2	2	1	5
Fjortoft et al., 2011	Cardiorespiratory fitness	Reduced Cooper	2	2	1	5
Aadahl et al., 2012	Cardiorespiratory fitness	Danish step	2	2	2	6
Langhammer & Stanghelle, 2018	Cardiorespiratory fitness	2-min step	2	2	2	6
Langhammer & Stanghelle, 2018	Cardiorespiratory fitness	6-min walk	2	2	2	6
Laukkanen et al., 1992	Cardiorespiratory fitness	2-km walk	2	2	2	6
Suni et al., 1998	Cardiorespiratory fitness	2-km walk	2	2	2	6
España-Romero et al., 2010	Cardiorespiratory fitness	20-m shuttle run	2	2	2	6

McAllister & Palombaro, 2019	Musculoskeletal fitness	Modified 30-s sit-to-stand	0	1	1	2
Hébert et al., 2011	Musculoskeletal fitness	Handgrip	2	2	1	5
Amado-Pacheco et al., 2019	Musculoskeletal fitness	Handgrip	2	2	1	5
Cadenas-Sánchez et al., 2016	Musculoskeletal fitness	Handgrip	2	1	2	5
Fjortoft et al., 2011	Musculoskeletal fitness	Standing long jump	2	2	1	5
Amado-Pacheco et al., 2019	Musculoskeletal fitness	Standing long jump	2	2	1	5
Cadenas-Sánchez et al., 2016	Musculoskeletal fitness	Standing long jump	2	1	2	5
Smits-Engelsman et al., 2020	Musculoskeletal fitness	Standing long jump	2	1	2	5
Fjortoft et al., 2011	Musculoskeletal fitness	Jumping a distance of 7 m on 1 feet	2	2	1	5
Fjortoft et al., 2011	Musculoskeletal fitness	Jumping a distance of 7 m on 2 feet	2	2	1	5
Fjortoft et al., 2011	Musculoskeletal fitness	Pushing a medicine ball	2	2	1	5
Boyer et al., 2013	Musculoskeletal fitness	Partial curl-ups, 60-s, 90-s and unlimited plank	2	2	1	5
Ito et al., 1996	Musculoskeletal fitness	Trunk flexor endurance	2	2	1	5
Ito et al., 1996	Musculoskeletal fitness	Isometric back endurance	2	2	1	5
Amado-Pacheco et al., 2019	Musculoskeletal fitness	Sit and reach	2	2	1	5
España-Romero et al., 2010	Musculoskeletal fitness	Handgrip	2	2	2	6
Suni et al., 1998	Musculoskeletal fitness	Handgrip	2	2	2	6
España-Romero et al., 2010	Musculoskeletal fitness	Standing long jump	2	2	2	6
Suni et al., 1998	Musculoskeletal fitness	Vertical jump	2	2	2	6

Langhammer & Stanghelle, 2018	Musculoskeletal fitness	30-s sit-to-stand	2	2	2	6
Langhammer & Stanghelle, 2018	Musculoskeletal fitness	Arm curl	2	2	2	6
Suni et al., 1998	Musculoskeletal fitness	Isometric back endurance	2	2	2	6
Suni et al., 1998	Musculoskeletal fitness	Modified push-ups	2	2	2	6
Langhammer & Stanghelle, 2018	Musculoskeletal fitness	Chair sit and reach	2	2	2	6
Langhammer & Stanghelle, 2018	Musculoskeletal fitness	Back scratch	2	2	2	6
Fjortoft et al., 2011	Motor fitness	Running 20-m	2	2	1	5
Amado-Pacheco et al., 2019	Motor fitness	4 × 10-m shuttle run	2	2	1	5
Cadenas-Sánchez et al., 2016	Motor fitness	4 × 10-m shuttle run	2	1	2	5
Fjortoft et al., 2011	Motor fitness	5 × 10-m shuttle run	2	2	1	5
Cadenas-Sánchez et al., 2016	Motor fitness	Single-leg stand	2	1	2	5
Smits-Engelsman et al., 2020	Motor fitness	Single-leg stand	2	1	2	5
Smits-Engelsman et al., 2020	Motor fitness	Dynamic balance	2	1	2	5
Langhammer & Stanghelle, 2018	Motor fitness	2.45-m time up & go	2	2	2	6
Suni et al., 1998	Motor fitness	Single-leg stand	2	2	2	6

Total score indicates high quality = 5-6; low quality = 3-4; very low quality = 0-2.

Table 5. Quality assessment of field-based fitness test safety studies.

Study	Fitness Component	Field-based fitness test	Number of study subjects	Description of the study population	Statistical analysis	Total score
Borel et al., 2010	Cardiorespiratory fitness	6-min step	1	2	1	4
Amado-Pacheco et al., 2019	Cardiorespiratory fitness	20-m shuttle run	2	2	0	4
Bruggeman et al., 2020	Cardiorespiratory fitness	45-s squat	1	2	2	5
Bruggeman et al., 2020	Cardiorespiratory fitness	3-min step	1	2	2	5
Anderson & Dal Corso, 2016	Cardiorespiratory fitness	Chester step	1	2	2	5
Anderson & Dal Corso, 2016	Cardiorespiratory fitness	Modified incremental step	1	2	2	5
Anderson & Dal Corso, 2016	Cardiorespiratory fitness	6-min walk	1	2	2	5
Lamonedá et al., 2020	Cardiorespiratory fitness	20-m shuttle run music	2	1	1	5
Lamonedá et al., 2020	Cardiorespiratory fitness	20-m shuttle run	2	1	1	5
Aadahl et al., 2012	Cardiorespiratory fitness	Danish step	2	2	2	6
Oja et al., 1991	Cardiorespiratory fitness	2-km walk	2	2	2	6
Laukkanen et al., 1992	Cardiorespiratory fitness	2-km walk	2	2	2	6
Suni et al., 1998	Cardiorespiratory fitness	2-km walk	2	2	2	6
España-Romero et al., 2010	Cardiorespiratory fitness	20-m shuttle run	2	2	2	6
Amado-Pacheco et al., 2019	Musculoskeletal fitness	Handgrip	2	2	0	4
Amado-Pacheco et al., 2019	Musculoskeletal fitness	Standing long jump	2	2	0	4
Hébert et al., 2011	Musculoskeletal fitness	Handgrip	2	2	1	5
Ito et al., 1996	Musculoskeletal fitness	Trunk flexor endurance	2	2	1	5
Ito et al., 1996	Musculoskeletal fitness	Isometric back endurance	2	2	1	5

Amado-Pacheco et al., 2019	Musculoskeletal fitness	Sit and reach	2	2	0	4
Smits-Engelsman et al., 2020	Musculoskeletal fitness	Standing long jump	2	1	2	5
España-Romero et al., 2010	Musculoskeletal fitness	Handgrip	2	2	2	6
Suni et al., 1998	Musculoskeletal fitness	Handgrip	2	2	2	6
España-Romero et al., 2010	Musculoskeletal fitness	Standing long jump	2	2	2	6
Suni et al., 1998	Musculoskeletal fitness	Vertical jump	2	2	2	6
Suni et al., 1998	Musculoskeletal fitness	Isometric back endurance	2	2	2	6
Suni et al., 1998	Musculoskeletal fitness	Modified push-ups	2	2	2	6
Amado-Pacheco et al., 2019	Motor fitness	4 x 10-m shuttle run	2	2	0	4
Smits-Engelsman et al., 2020	Motor fitness	Dynamic balance	2	1	2	5
Smits-Engelsman et al., 2020	Motor fitness	Single-leg stand	2	1	2	5
Suni et al., 1998	Motor fitness	Single-leg stand	2	2	2	6

Total score indicates high quality = 5-6; low quality = 3-4; very low quality = 0-2.

Characteristic of included studies

The sample size in the feasibility studies was 4,186 participants, of which 2,124 were children and adolescents, 1,857 were adults and 205 were older adults (see **Table 6**). From the 19 original studies, 14 specified the gender with a ratio of women of 26% (n = 991) for children and adolescents, and 27.4% (n = 1,037) for adults. None of the studies specified gender in older adults.

The sample size in the safety studies was 2,665 participants, of which 808 were children and adolescents and 1,857 were adults (see **Table 7**). From the 13 original studies, 10 specified the gender with a ratio of women of 13.4% (n = 334) for children and adolescents, and 41.6% (n = 1,037) for adults. None of the studies specified gender in older adults.

Feasibility and safety of the field-based physical fitness tests were measured through questionnaires. Fourteen items were identified to assess feasibility (reported by evaluators and participants) and 11 items were observed to assess safety of the field-based physical fitness tests (reported by evaluators and by participants). Most of the studies used the frequency in number or percentages to determine feasibility and safety. Heart rate (HR) and scales presented the data in mean and standard deviation.

Feasibility

Cardiorespiratory fitness

Distance and Time-Based Run/Walk Tests

Five high-quality studies⁷⁻¹¹ have analysed the feasibility of the distance run/walk test (see **Table 6**). Three studies⁹⁻¹¹ showed that the 2-km walk test was feasible (~20% of the participants rejected to perform the test, the time to perform the test were ~15 to 18 minutes, two-three evaluators assessed the test and the 81% of the participants found the test “easy”) to assess cardiorespiratory fitness in adults. One study⁷ showed that the 6-min walk test was feasible (100% of the participants performed the test) to assess cardiorespiratory fitness in older adults. One study⁸ reported that the Reduced Cooper test was feasible (100% of the participants performed the test, 100% of the participants understood test instructions and the test was considered easy to administer) to assess

cardiorespiratory fitness in children.

Only one low-quality study³ found that the 6-min walk test was feasible (100% of the participants perform the test and the time to perform the test was 6 minutes) to assess cardiorespiratory fitness in adults.

Twenty-meters Shuttle Run

Four high-quality studies^{6, 12-14} analysed the feasibility of the 20-m shuttle run test in children (see **Table 6**). All of them showed that this test was feasible to assess cardiorespiratory fitness (8 to 10 minutes to perform the test, a minimum of 2 evaluators assessed the test^{12, 13}; groups of 4 to 8 participants, 4 to 5 participants if they are 3 years old¹²; groups of 4 to 8 participants of the same age, a 100% of the participants understood test instructions and the test was easy to administer¹³; 96% of the participants performed the test, and the time to perform the test were 6.7 ~3 minutes for males and 5.3 ~2 minutes for females⁶ and; 5 minutes to prepare and 10 minutes to perform the test approximately, 20 participants per group, 1 physical education teacher as evaluator, 99.2% of participants wore appropriate sport clothes, 100% of participants understood test instructions, 0.8% rejected to perform the test by lack of motivation, 100% of facilities were appropriate and the test was easy to administer¹⁴).

Only one low-quality study⁴ considered feasible the 20-m shuttle run music and the 20-m shuttle run tests in adolescents to assess cardiorespiratory fitness. The study compared both tests (20-m shuttle run music vs. 20-m shuttle run) through self-reported questionnaire, finding a little more feasible to the 20-m shuttle run music test (preference: 89.6% vs. 10.4%; enjoyment: 94% vs. 6%; maximum effort: 72.4% in both tests; and less fatigue perception: 56.4% in 20-m shuttle run music).

Step Tests

Two high-quality studies^{7, 15} analysed the feasibility of the step tests (see **Table 6**). The Danish step test¹⁵ and the 2-min step test⁷ were feasible (99% and 100% of participants performed the test and the time to perform was 6 and 2 minutes, respectively) to assess cardiorespiratory fitness in adults and older adults, respectively.

Three low-quality studies^{2, 3, 5} analysed the feasibility of the step tests. The 45-s squat and the 3-min step tests² were feasible (45 seconds and 3 minutes to perform the tests, respectively, 100% of the participants found the tests as “low effort” and the tests were considered easy to administer) to assess cardiorespiratory fitness in children and adolescents. The Chester step and modified incremental step tests³ were considered feasible (10 and 12 minutes to perform the tests, respectively, and the 100% of the participants performed the tests) to assess cardiorespiratory fitness in adults. Moreover, the 6-min step test⁵ showed to be feasible (6 minutes to perform the test and 100% of the participants performed the test) to assess cardiorespiratory fitness in adults.

Musculoskeletal fitness

Maximal Isometric Fitness

Five high-quality studies^{6, 9, 13, 14, 16} analysed the feasibility of hand maximal isometric strength, using the handgrip strength test (see **Table 6**). Four studies^{6, 13, 14, 16} showed that the handgrip strength test was feasible (100% of acceptance for the test^{6, 16}; 100% of the participants performed the test⁶; 8 to 10 minutes to perform test, 8 participants per group, 100% of the participants understood the test instructions correctly and the test was considered easy to administer¹³ and; ~3 minutes of preparation, ~90 seconds to perform the test by participant, 20 participants per group, 1 physical education teacher as evaluator, 99.2% of participants wore appropriate sport clothes, 98.5% of participants understood test instructions, there was not rejected to perform the test, 100% of facilities were appropriate and the test was considered easy to administer¹⁴) to assess maximal isometric muscular strength in children and adolescents. One study⁹ reported that the handgrip strength test was feasible (time to prepare and perform the test trough 5 points scale: mean score = 4.3 - 4.7 (excellent) and 95% of the participants performed the test) to assess maximal isometric muscular strength in adults.

Explosive Fitness

Six high-quality studies^{6, 8, 9, 13, 14, 17} assessed the feasibility of lower body explosive strength tests (see **Table 6**). These studies reported that the standing long jump test was feasible (100% of the participants performed the test^{6, 8, 17}; 100% of the participants understood test instructions and the test was considered easy to administer⁸; 8 to 10 minutes to perform the test, maximum 8 participants per group, 100% of the

participants showed positive answer about “understanding the test instructions” and the test was considered difficult to administer (several trials were required)¹³; ~3 minutes of preparation, 50 seconds to perform the test by participant, 20 participants per group, 1 physical education teacher as evaluator, 99% of participants wore appropriate sport clothes and understood the test instructions, 0.8% of participants rejected to perform the test by lack of motivation, 100% of facilities were appropriate and the test was considered easy to administer¹⁴; high acceptance of the test for participants (that is, 100% of the participants performed the test and understood the test instructions, and 76% of participants wore appropriate sport clothes) and evaluators (that is, the test was considered easy to administer and the space to administer the test was appropriate (5-6 metres long by 3-4 metres wide)¹⁷) to assess lower body explosive strength in children and adolescents. Moreover, one study⁸ showed that the jumping a distance of 7-m on 1 feet and jumping a distance of 7-m on 2 feet tests, were feasible (100% of participants performed the tests, 100% of participants understood the tests instructions and the tests were considered easy to administer) to assess lower body explosive strength in children. Finally, one study⁹ showed that the vertical jump test was considered feasible (time to prepare and perform the test trough 5 points scale: mean score = 5.0 (excellent) and 95% of the participants performed the test) to assess lower body explosive strength in adults.

One high-quality study⁸ assessed the feasibility of upper body explosive strength test (see **Table 6**). This study reported that pushing a medicine ball test was feasible in children (100% of the participants performed the test, 100% of participants understood the test instructions and the test were considered easy to administer).

Endurance Fitness

Three high-quality studies^{9, 18, 19} assessed the feasibility of trunk endurance strength tests (see **Table 6**). One study¹⁹ analyzed the feasibility of the partial curl-ups, 60-s, 90-s and unlimited plank tests (time to perform the test for: partial curl-ups (Fitnessgram) = 29 s ± 28, partial curl-ups (Canadian Health Measures Surve (CHMS)) = 36 s ± 21, 60-s plank = 41 s ± 16, the 90-s plank = 52 s ± 25 and the unlimited plank = 56 s ± 40; zero, interpretable and maximum scores for: partial curl-ups(Fitnessgram) = 12% of children could not perform the test, partial curl-ups (CHMS) = 5% of children could not perform the test (both tests only differentiates 70% of children to assess upper body muscular strength), 60-s plank = only differentiates 55% of children, 90-s plank = only differentiates 86% of children, and for unlimited plank = 133 seconds was required

to provide a nonzero/nonmaximal score for 95% of the participants).

One study¹⁸ reported that the trunk flexor endurance and isometric back endurance were feasible (100% of the participants performed the test, 95.5% of the participants performed the trunk flexor endurance test within the given time (≤ 5 minutes) and, 89% of the participants performed the isometric back endurance test within the given time (≤ 5 minutes) to assess trunk endurance muscular strength in children. Finally, one study⁹ reported that isometric back endurance test was feasible (time to prepare and perform the test through 5 points scale: mean score = 4.3–4.7 (excellent) and 95% of the participants performed the test) to assess trunk endurance muscular strength in adults.

One high-quality study⁷ assessed the feasibility of lower body endurance strength (see **Table 6**). This study found that the 30-s sit-to-stand test was feasible in older adults (100% of the participants performed the test and the time to perform the test was 30 seconds).

Two high-quality studies^{7,9} analysed the feasibility of upper body endurance strength tests (see **Table 6**). One study⁹ reported that the modified push-ups test was feasible (time to prepare and perform the test through 5 points scale: mean score = 4.3–4.7 (excellent) and 95% of the participants performed the test) to assess upper body muscular strength in adults. Also, one study showed that the arm curl test⁷ was feasible (100% of the participants performed the test) to assess upper body endurance muscular strength in older adults.

Finally, one very low-quality study¹ found feasible the modified 30-s sit-to-stand test (82% of the participants performed at least 1 repetition of the test) to assess lower body endurance muscular strength in older adults.

Flexibility

Two high-quality studies^{6,7} assessed the feasibility of flexibility tests (see **Table 6**). One study⁶ reported that the sit and reach test was feasible (100% of the participants performed the test) to assess lower body flexibility in children. One study⁷ showed that the chair sit and reach and the back scratch tests were feasible (100% of the participants performed the test) to assess lower body flexibility and upper body flexibility in older adults, respectively.

Motor Fitness

Speed

Only one high-quality study⁸ evaluated the feasibility of the running 20-m test, concluding that this test was feasible (100% of the participants performed the test, 100% understood the test instruction and it was considered easy to administer) to assess speed in children (see **Table 6**).

Agility

Four high-quality studies^{6-8, 13} evaluated the feasibility of agility tests (see **Table 6**). Two studies^{6, 13} reported that the 4 x 10-m shuttle run test was feasible (100% of the participants performed the test, time to perform the test was for males = 17.3 ± 2.5 and for females = 18.7 ± 3.1 ⁶; 8 to 10 minutes to perform the test, maximum 8 participants per group, 2 evaluators assessed the performance of the test, 100% of the participants understood the test instructions and the test was considered easy to administer¹³) to assess agility and speed in children. One study⁸ observed that the 5 x 10-m shuttle run test was feasible (100% of the participants performed the test, and understood the test instructions, and the test was considered easy to administer) to assess agility and speed in children. One high-quality study⁷ found that the 2.45-m time up & go test (also known as 8-ft up and go test) was feasible (100% of the participants performed the test) to assess agility, balance and speed of walking in older adults.

Balance

Three high-quality studies^{9, 13, 17} evaluated the feasibility of balance tests (see **Table 6**). Two studies^{13, 17} reported that the single-leg stand test was feasible (8 to 10 minutes to perform the test, maximum 8 participants per group, 100% of the participants understood the test instructions and the test was considered easy to administer¹³; 100% of the participants performed the test and understood the test instructions, 76% of participants wore appropriate sport clothes, the test was considered easy to administer and the space to perform it was appropriate¹⁷) to assess balance in children. Another study⁹ observed that the single-leg stand test was feasible (time to prepare and perform the test through 5 points scale: mean score = 4.3–4.7 (excellent), and 95% of the participants performed the test) to assess balance in adults. Finally, one study¹⁷ showed that the dynamic balance test was feasible (100% of the participants performed the test and understood the test instructions, 76% of participants wore appropriate sport clothes,

the test was considered easy to administer and the space to administer the test was appropriate) to assess balance in children.

Table 6. Overview of field-based fitness test feasibility studies by study quality.

Author	Participants	Age, range/ SD	Field-based fitness test	Feasibility items	Results	Conclusions
Very low quality studies						
<i>Musculoskeletal fitness (endurance)</i>						
McAllister & Palombaro, 2019	33 elderly Healthy	85.2 ± 8.0	Modified 30-s sit-to-stand	- Participants performing at least 1 repetition of the test (%) - Time to perform the test (minutes)	82% of the participants perform at least 1 repetition of the test. 30 seconds.	Modified 30-s sit-to-stand was considered feasible to assess lower body endurance muscular strength.
Low quality studies						
<i>Cardiorespiratory fitness</i>						
Bruggeman et al., 2020	40 Children Males (n = 23) Females (n = 17) Healthy	12.8 ± 1.9 13.3 ± 1.7 12.1 ± 2.0	45-s squat	- Time to perform the test (minutes) - Self-perception effort (%) - Easy to administer	45 seconds. 100% of the participants found the test as “low effort”. Test was considered easy to administer.	45-s squat was considered feasible to assess lower body muscular strength.
			3-min step		3 minutes. 100% of the participants found the test as “low effort”. Test was considered easy to administer.	3-min step was considered feasible to assess cardiorespiratory fitness.
Anderson & Dal Corso, 2016	20 adults Healthy	52 41 - 66	Chester step	- Participants performing the test	100% of the participants perform the	Chester step was considered feasible

				(%) -Time to perform the test (minutes)	test. 10 minutes.	for assess cardiorespiratory fitness.
			Modified incremental step		100% of the participants perform the test. 12 minutes.	Modified incremental step was considered feasible for assess cardiorespiratory fitness.
Borel et al., 2010	15 adults Healthy	29.0 ± 6.0	6-min step	- Participants performing the test (%) - Time to perform the test (minutes)	100% of the participants perform the test. 6 minutes.	6-min step was considered feasible for assess cardiorespiratory fitness.
Anderson & Dal Corso, 2016	20 adults Healthy	52 41 - 66	6-min walk	- Participants performing the test (%) - Time to perform the test (minutes)	100% of the participants perform the test. 6 minutes.	6-min walk seems to be feasible for assess cardiorespiratory fitness.
Lamonedá et al., 2020	386 adolescents Males (n = 189) Females (n = 197) Healthy	14.5 ± 1.6 14.4 ± 1.6 14.5 ± 1.6	20-m shuttle run music vs 20-m shuttle run	-Self-reported questionnaire items by participants: “preference”, “enjoyment”, “maximum effort” and “perception of fatigue” (%)	Preference = 89.6% vs. 10.4%. Enjoyment = 94% vs. 6%. Maximum effort = 72.4 both tests. Less fatigue perception: 56.4% for 20-m shuttle run music.	20-m shuttle run music and 20-m shuttle run seem to be feasible for assess cardiorespiratory fitness. 20-m shuttle run music seems to be more feasible.

High quality studies

Cardiorespiratory fitness

Aadahl et al., 2012	749 adults Males (n=346) Females (n=449) Healthy	46.8 ± 8.2 46.7 ± 8.4 47.0 ± 8.0	Danish step	- Participants performing the test (%) - Time to perform the test (minutes)	98.8% of the participants perform the test. 6 minutes.	Danish step was considered feasible for assess cardiorespiratory fitness.
Langhammer & Stanghelle, 2018	172 elderly Healthy	73 ± 5.8	2-min step	- Participants performing the test (%) - Time to perform the test (minutes)	100% of the participants perform the test. 2 minutes.	2-min step was considered feasible to assess cardiorespiratory fitness.
			6-min walk		100% of the participants perform the test. 6 minutes.	6-min walk seems to be feasible to assess cardiorespiratory fitness.
Oja et al., 1991	64 adults Males (n=35) Females (n=29) Healthy	41 ± 13.7 42.9 ± 14.0 39.1 ± 13.4	2-km walk	- Participants performing the test (%) - Time to perform the test (minutes, mean (SD))	100% of the participants perform the test. Males = 15.2 (1.8) minutes Females = 16.9 (1.5) minutes.	2-km walk seems to be feasible to assess cardiorespiratory fitness.
Laukkanen et al., 1992	273 adults Males (n = 121) Females (n = 152) Healthy	44.5 ± 10.3 44.5 ± 10.2 44.5 ± 10.5	2-km walk	- Participants performing the test (%) - Time to perform	96.4% of the participants perform the test.	2-km walk seems to be feasible to assess cardiorespiratory fitness.

				the test (minutes) - Reject to perform the test (%) - Number of evaluators (n) - Self-perception effort (%)	15-18 minutes. 21% of the participants rejected to perform the test. 2 evaluators. 81% of the participants found the test “easy”, 17% “somewhat difficult” and 2% “difficult”.	
Suni et al., 1998	600 adults Males (n = 246) Females (n = 354) Healthy	37 - 57 47.0 ± 7.9 47.0 ± 8.6	2-km walk	- Participants performing the test (%) - Time to perform the test (minutes) - Number of evaluators (n) - Easy to administer	95% of the participants performed the walk test. 20-25 minutes. 3 evaluators. Test was considered easy to administer.	2-km walk seems to be feasible to assess cardiorespiratory fitness.
Amado-Pacheco et al., 2019	90 children Males (n = 48) Females (n = 42) Healthy	4.04 ± 0.82 3.95 ± 0.82	20-m shuttle run	- Participants performing the test (%) - Time to perform the test (minutes, mean (SD))	96% of the participants perform the test. Time males = 6.7 (3). Time females = 5.3 (2).	20-m shuttle run seems to be feasible to assess cardiorespiratory fitness.
Cadenas-Sanchez et al., 2014	137 children Males (n = 78) Females (n = 59) Healthy	4.91 ± 0.89 4.9 ± 0.8 4.8 ± 0.8	20-m shuttle run	- Time to perform the test (minutes) - Time to prepare the test (minutes) - Participants per	8-10 minutes (including explanations and demonstrations). 3 minutes.	20-m shuttle run seems to be feasible to assess cardiorespiratory fitness.

				group (n) - Number of evaluators (n)	Groups of 4-8 participants, 4-5 participants if they are 3-year-old. A minimum of 2 evaluators running with children.	
Cadenas-Sánchez et al., 2016	161 children Males (n = 92) Females (n = 69) Healthy	4.91 ± 0.83 4.94 ± 0.86 4.87 ± 0.79	20-m shuttle run	- Time to perform the test (minutes) - Participants per group (n) - Number of evaluators (n) - Understanding the test instruction (%) - Easy to administer	8-10 minutes (including explanations and demonstrations). Groups of 4-8 participants of the same age. A minimum of 2 evaluators running with children. 100% of the participants understood the test. Test was considered easy to administer.	20-m shuttle run seems to be feasible to assess cardiorespiratory fitness.
España-Romero et al., 2010	58 children 80 adolescents Healthy	6 - 11.9 12 - 18	20-m shuttle run	- Time to prepare the test (minutes) - Time to perform the test (minutes) - Participants per group (n) - Number of evaluators (n) - Questionnaire with items of dichotomous answer (%)	~5 minutes of preparation. ~10 minutes of perform. 20 participants. 1 evaluator (Physical Education teachers at school). 99.2% of participants wore appropriate sport	20-m shuttle run seems to be feasible to assess cardiorespiratory fitness.

Fjortoft et al., 2011	195 children Males (n = 94) Females (n = 101) Healthy	5-12	Reduced Cooper	- Easy to administer - Participants performing the test (%) - Understanding the test instruction (%) - Easy to administer	clothes, 100% understood test instructions and 0.8% rejected to perform the test by lack of motivation. 100% facilities were appropriate. Test was considered easy to administer. 100% of the participants perform the test. 100% of the participants understood the test. Test was considered easy to administer.	Reduced Cooper was considered feasible to assess cardiorespiratory fitness.
<i>Musculoskeletal fitness (isometric)</i>						
Hébert et al., 2011	33 children 41 adolescents Males (n = 37) Females (n = 37) Healthy	4 - 9.9 10 - 17.5	Handgrip	- Understanding the test instruction (%) - Easy to administer	100% of the participants understood the test. Test was considered easy to administer.	Handgrip was considered feasible to assess isometric muscular strength.
Amado-Pacheco et al., 2019	90 children Males (n = 48) Females (n = 42) Healthy	4.04 ± 0.82 3.95 ± 0.82	Handgrip	- Participants performing the test (%)	100% of the participants performed the test.	Handgrip was considered feasible to assess isometric muscular strength.
Cadenas-Sánchez et al., 2016	161 children Males (n = 92) Females (n = 69)	4.91 ± 0.83 4.94 ± 0.86 4.87 ± 0.79	Handgrip	- Time to perform the test (minutes) - Participants per	8-10 minutes (including explanations and demonstrations)	Handgrip was considered feasible to assess isometric

	Healthy			group (n) - Understanding the test instruction (%) - Easy to administer	Groups of maximum 8 participants. 100% of the participants understood the test. Test was considered easy to administer.	muscular strength.
España-Romero et al., 2010	58 children 80 adolescents Healthy	6 - 11.9 12 - 18	Handgrip	- Time to prepare the test (minutes) - Time to perform the test (minutes) - Participants per group (n) - Number of evaluators (n) - Questionnaire with items of dichotomous answer (%) - Easy to administer	~3 minutes of preparation. ~90 seconds of perform by participants. Groups of 20 participants. 1 evaluator (Physical Education teachers at school). 99.2% of participants wore appropriate sport clothes, 98.5% understood test instructions and 0% rejected to perform the test by lack of motivation. 100% facilities were appropriate. Test was considered easy to administer.	Handgrip was considered feasible to assess isometric muscular strength.
Suni et al., 1998	600 adults Males (n = 246)	37 - 57 47.0 ± 7.9	Handgrip	- Time to prepare and perform the test	Mean score = 4.3-4.7 (excellent).	Handgrip was considered feasible to

Females (n = 354) Healthy	47.0 ± 8.6		(5 points scale (mean)) - Participants performing the test (%) - Easy to administer	95% of the participants perform the test. Test was considered easy to administer.	asses isometric muscular strength.
------------------------------	------------	--	--	--	---------------------------------------

Musculoskeletal fitness (explosive)

Fjortoft et al., 2011	195 children Males (n = 94) Females (n = 101) Healthy	5-12	Standing long jump	- Participants performing the test (%) - Acceptance of the test (%) - Easy to administer	100% of the participants perform the test. 100% of the participants shows acceptance for the test. Test was considered easy to administer.	Standing long jump was considered feasible to assess lower body explosive strength.
Amado-Pacheco et al., 2019	90 children Males (n = 48) Females (n = 42) Healthy	4.04 ± 0.82 3.95 ± 0.82	Standing long jump	- Participants performing the test (%)	100% of the participants performed the test.	Standing long jump was considered feasible to assess lower body explosive strength.
Cadenas-Sánchez et al., 2016	161 children Males (n = 92) Females (n = 69) Healthy	4.91 ± 0.83 4.94 ± 0.86 4.87 ± 0.79	Standing long jump	- Time to perform the test (minutes) - Participants per group (n) - Understanding the test instruction (%) - Easy to administer	8-10 minutes (including explanations and demonstrations) Groups of maximum 8 participants. 100% of the participants understood the test. Test was considered difficult to administer (several more trials are	Standing long jump was considered feasible to assess lower body explosive strength; however, it was needed more trials to perform it correctly (because it requires plenty of coordination).

Smits-Engelsman et al., 2020	80 children Males (n = 39) Females (n = 41) Healthy	9.2 ± 1.1	Standing long jump	- Participants performing the test (%) - Understanding the test instruction (%) - Appropriate sport clothes (%) - Easy to administer - Space (m)	required).	Standing long jump seems to be feasible to asses lower body explosive strength.
					100% of the participants performed the test. 100% of the participants understood the test. 76% appropriate sport clothes. Test was considered easy to administer. 5–6m long by 3–4m wide.	
España-Romero et al., 2010	58 children 80 adolescents Healthy	6 - 11.9	Standing long jump	- Time to prepare the test (minutes) - Time to perform the test (minutes) - Participants per group (n) - Number of evaluators (n) - Questionnaire with items of dichotomous answer (%) - Easy to administer	~3 minutes of preparation.	Standing long jump seems to be feasible to asses lower body explosive strength.
		12 - 18			~50 seconds of perform by participants. Groups of 20 participants. 1 evaluator (Physical Education teachers at school). 99.2% of participants wore appropriate sport clothes, 99.2% understood test instructions and 0.8% rejected to perform the test by lack of	

					motivation. 100% facilities were appropriate. Test was considered easy to administer.	
Suni et al., 1998	600 adults Males (n = 246) Females (n = 354) Healthy	37 - 57 47.0 ± 7.9	Vertical jump	- Time to prepare and perform the test (5 points scale (mean score)) - Participants performing the test (%) - Easy to administer	Mean score = 5.0 (excellent). 95% of the participants perform the test. Test was considered easy to administer.	Standing long jump was considered feasible to assess lower body explosive strength.
Fjortoft et al., 2011	195 children Males (n = 94) Females (n = 101) Healthy	5-12	Jumping a distance of 7-m on 1 feet	- Performance of the test (%) - Understanding the test instruction (%) - Easy to administer	100% of the participants perform the test. 100% of the participants understood the test. Test was considered easy to administer.	Jumping a distance of 7-m on 1 feet, seems to be feasible to assess lower body explosive strength.
Fjortoft et al., 2011	195 children Males (n = 94) Females (n = 101) Healthy	5-12	Jumping a distance of 7-m on 2 feet	- Performance of the test (%) - Understanding the test instruction (%) - Easy to administer	100% of the participants perform the test. 100% of the participants understood the test. Test was considered easy to administer.	Jumping a distance of 7-m on 2 feet, seems to be feasible to assess lower body explosive strength.

Fjortoft et al., 2011	195 children Males (n = 94) Females (n = 101) Healthy	5-12	Pushing a medicine ball (1 kg)	- Participants performing the test (%) - Understanding the test instruction (%) - Easy to administer	100% of the participants perform the test. 100% of the participants understood the test. Test was considered easy to administer.	Pushing a medicine ball (1 kg) to be feasible to assess upper body explosive strength.
<i>Musculoskeletal fitness (endurance)</i>						
Langhammer & Stanghelle, 2018	172 elderly Healthy	73 ± 5.8	30-s sit-to-stand	- Participants performing the test (%) - Time to perform the test (minutes)	100% of the participants perform the test. 30 seconds.	30-s sit-to-stand was considered feasible to assess lower body endurance muscular strength.
			Arm curl		100% of the participants perform the test.	Arm curl was considered feasible to assess upper body endurance muscular strength.
Boyer et al., 2013	823 children Males (n = 395) Females (n = 428) Healthy	8-12	Fitnessgram partial curl-ups CHMS (Canadian Health Measures Survey) partial curl-ups 60-s plank 90-s plank Unlimited time plank	- Participants performing the test (%): zero (floor effect), interpretable (scores between nonzero and a nonmaximum), and maximum scores (ceiling effect).	12%, 5% and 0.8% of participants cannot perform (zero score) the partial curl-up (Fitnessgram and CHMS) and plank protocols tests, respectively. CHMS partial curl-ups: Only differentiates 70% of participants.	The unlimited plank was considered feasible to evaluate trunk endurance muscular strength (it provides the best differentiation; it takes significantly more time to administer in comparison with the

					-Time to perform the tests (seconds, mean (SD))	60-s plank: only differentiates 55% of participants. 90-s plank: only differentiates 86% of participants. Unlimited time plank:133s was required to provide a nonzero/ nonmaximal score for 95% of the participants. Fitnessgram partial curl-ups (29 s (28)), CHMS partial curl-ups (36 s (21)), 60-s plank (41 s (16)), 90-s plank (52 s (25)) and unlimited time (56 s (40)).	Fitnessgram partial curl-ups protocol).
Ito et al., 1996	90 adults Males (n = 37) Females (n = 53) Healthy	44.3 46.8	Trunk flexor endurance	Participants performing the test (%) -Time to perform the test (minutes, mean (SD)) -Easy to administer	100 % of the participants perform the test. Male = 182.6 (69.3) minutes. Female = 208.2 (66.2) minutes. 4.4% participants were able to achieve more than 5 minutes. Test was considered easy to administer.	Trunk flexor and isometric back endurance tests were considered feasible to asses trunk endurance muscular strength.	
			Isometric back endurance			100 % of the	

					participants perform the test. Male = 85.1 (44.8) minutes. Female = 128.4 (53.0) minutes. 11.1% participants were able to achieve more than 5 minutes. Test was considered easy to administer.	
Suni et al., 1998	600 adults Males (n = 246) Females (n = 354) Healthy	37 - 57 47.0 ± 7.9	Isometric back endurance	- Participants performing the test (%) - Time to prepare and perform the test (5 points scale (mean score)) - Easy to administer	95% of the participants perform the test. Mean score = 4.3-4.7 (excellent). Test was considered easy to administer.	Isometric back endurance was considered feasible to asses trunk endurance muscular strength.
			Modified push-ups		95% of the participants perform the test. Mean score = 4.3-4.7 (excellent). Test was considered easy to administer.	Modified push-ups was considered feasible to asses upper body endurance muscular strength.
<i>Musculoskeletal fitness (flexibility)</i>						
Amado-Pacheco et al., 2019	90 children Males (n = 48) Females (n = 42)	4.04 ± 0.82 3.95 ± 0.82	Sit and reach	- Participants performing the test (%)	100% of the participants performed the test.	Sit and reach and was considered feasible to asses lower body

	Healthy					flexibility.
Langhammer & Stanghelle, 2018	172 elderly Healthy	73 ± 5.8	Chair sit and reach	- Participants performing the test (%)	100% of the participants perform the test.	Chair sit and reach was considered feasible to asses lower body flexibility.
			Back scratch		100% of the participants perform the test.	Back scratch was considered feasible to asses upper body flexibility.
<i>Motor fitness (speed)</i>						
Fjortoft et al., 2011	195 children Males (n = 94) Females (n = 101) Healthy	5-12	Running 20-m	- Participants performing the test (%) - Understanding the test instruction (%) - Easy to administer	100% of the participants perform the test. 100% of the participants understood the test. Test was considered easy to administer.	Running 20-m seems to be feasible to evaluate speed.
<i>Motor fitness (agility)</i>						
Amado-Pacheco et al., 2019	90 children Males (n = 48) Females (n = 42) Healthy	4.04 ± 0.82 3.95 ± 0.82	4 × 10-m shuttle run	- Participants performing the test (%) - Time to perform the test (seconds, mean (SD))	100% of the participants performed the test. Time males = 17.3 (2.5) Time females = 18.7 (3.1).	4 × 10-m shuttle run was considered feasible to evaluate agility and speed.
Cadenas-Sánchez et al., 2016	161 children Males (n = 92) Females (n = 69) Healthy	4.91 ± 0.83 4.94 ± 0.86 4.87 ± 0.79	4 × 10-m shuttle run	- Time to perform the test (minutes) - Participants per group (n)	8-10 minutes (including explanations and demonstrations) Groups of maximum 8	4 × 10-m shuttle run was considered feasible to evaluate agility and speed.

				- Number of evaluators (n) - Understanding the test instruction (%) - Easy to administer	participants. 2 evaluators were positioned in both extremes and participants had to touch the evaluator hand. 100% of the participants understood the test. Test was considered easy to administer.	
Fjortoft et al., 2011	195 children Males (n = 94) Females (n = 101) Healthy	5-12	5 × 10-m shuttle run	- Participants performing the test (%) -Understanding the test instruction (%) - Easy to administer	100% of the participants perform the test. 100% of the participants understood the test. Test was considered easy to administer.	5 × 10-m shuttle run was considered feasible to evaluate agility and speed.
Langhammer & Stanghelle, 2018	172 elderly Healthy	73 ± 5.8	2.45-m time up & go	- Participants performing the test (%)	100% of the participants perform the test.	2.45-m up and go was considered feasible to evaluate agility, balance and speed of walking.
<i>Motor fitness (balance)</i>						
Cadenas-Sánchez et al., 2016	161 children Males (n = 92) Females (n = 69) Healthy	4.91 ± 0.83 4.94 ± 0.86 4.87 ± 0.79	Single-leg stand	- Time to perform the test (minutes) -Subjects per group (n) - Understanding the	8-10 minutes (including explanations and demonstrations) Groups of maximum 8 participants.	Single-leg stand was considered feasible to evaluate balance and lower body strength

				test instruction (%)	100% of the participants understood the test.	
				- Easy to administer	Test was considered easy to administer.	
Smits-Engelsman et al., 2020	80 children Males (n = 39) Females (n = 41) Healthy	9.2 ± 1.1	Singe-leg stand	- Participants performing the test (%) - Understanding the test instruction (%) - Appropriate sport clothes (%) - Easy to administer	100% of the participants performed the test. 100% of the participants understood the test. 76% of participants wore appropriate sport clothes.	Singe-leg stand and dynamic balance were considered feasible to evaluate balance and lower body strength.
			Dynamic balance	- Space (m)	Test was considered easy to administer. 5–6m long by 3–4m wide.	
					100% of the participants performed the test. 100% of the participants understood the test. 76% of the participants wore appropriate sport clothes. Test was considered easy to administer. 5–6m long by 3–4m wide.	

Suni et al., 1998	600 adults Males (n = 246) Females (n = 354) Healthy	37 - 57 47.0 ± 7.9	Single-leg stand	- Time to prepare and perform the test (5 points scale (mean score)) - Participants performing the test (%) - Easy to administer	Mean score = 4.3-4.7 (excellent). 95% of the participants perform the test. Test was considered easy to administer.	Single-leg stand was considered feasible to evaluate balance and lower body strength.
-------------------	---	-----------------------	------------------	--	---	---

Safety

Cardiorespiratory fitness

Distance and Time-Based Run/Walk Tests

Four high-quality studies^{3, 9-11} analysed the safety of the distance and time-based run/walk test (see **Table 7**). Three studies⁹⁻¹¹ showed that the 2-km walk test was safe to assess cardiorespiratory fitness in adults (delayed onset muscle soreness (DOMS) males = 60% and females = 78%, HR bpm males = 153 ± 17.4 and females = 151 ± 18.3 , exertion 85%HRmax males = 43% and females = 37%, and $\geq 95\%$ performed the test without adverse events⁹; rating of perceived exertion (RPE) males = 3 ± 1.8 and females = 2.9 ± 1.5 (exertion perceived as “moderate”), HR bpm males = 153.1 ± 22.2 and females = 153.8 ± 16 and, $\geq 70\%$ performed the test without adverse events¹⁰ and; RPE = 3.1 ± 2.4 (exertion perceived as “moderate”), HR bpm males = 139.1 ± 19.1 and females = 143.5 ± 17.1 ¹¹). One study³ reported that the 6-min walk test was safe (HR bpm = 128 ± 25 , Borg scale dyspnea = 2 (1–3) and Borg scale leg fatigue = 2 (1–3) and 100% of participants perform the test without adverse events) to assess cardiorespiratory fitness in adults.

Twenty-meters Shuttle Run Test

Two high-quality studies^{4, 14} analysed the safety of the 20-m shuttle run test, only in children and adolescents (see **Table 7**). One study¹⁴ observed that the 20-m shuttle run test was safe ($\geq 99\%$ participants performed the test without adverse events and 71.2% experienced some degree of DOMS, of them, the 29.3% assumed that the 20-m shuttle run test could be the cause) to assess cardiorespiratory fitness in children and adolescents. Another study⁴ compared the 20-m shuttle run test with music vs. the original version; both tests were considered safe (RPE = 8.3 ± 2 vs. 7.9 ± 2.2 and HR bpm = 177.1 ± 30.6 vs. 177.7 ± 29.2) to assess cardiorespiratory fitness in adolescents.

Only one low-quality study⁶ reported that the 20-m shuttle run test was safe (100% of participants performed the test without adverse events) to assess cardiorespiratory fitness in children.

Step and squat tests

Three high-quality studies^{2, 3, 15} analysed the safety of the step and squats tests (see **Table 7**). One study² observed that the 45-s squat ($RPE = 3.4 \pm 1.7$ and, 100% of participants perform the test without adverse event) and the 3-min step tests ($RPE = 2.7 \pm 1.4$ and, 100% of participants perform the test without adverse event) were safe to assess cardiorespiratory fitness in children. Two studies^{3, 15}, showed that the Chester step (Borg scale dyspnea = 2 (1–3), Borg scale leg fatigue = 2 (1–3), HR bpm = 143 ± 28 and, 100% of participants perform the test without adverse events); modified incremental step (Borg scale dyspnea = 2 (1–3), Borg scale leg fatigue = 2 (1–4), HR bpm = 142 ± 27 and, 100% of participants perform the test without adverse event³) and Danish step tests (0.1% of participants experienced adverse events (1 participant had strained calf after 30 s)) were safe to assess cardiorespiratory fitness in adults¹⁵.

Only one low-quality study⁵ reported that the 6-min step test was safe (Borg scale dyspnea = 2.5 ± 1.5 , Borg scale leg fatigue = 3.4 ± 2.5 and HR bpm = 118.2 ± 18.7) to assess cardiorespiratory fitness in adults.

Muscular Fitness

Maximal Isometric Fitness

Three high-quality studies^{9, 14, 16} analysed the safety of hand maximal isometric strength, using the handgrip strength test (see **Table 7**). Two studies^{14, 16} concluded that the test was safe (100% of participants performed the test without adverse events (i.e., pain or discomfort)¹⁶; $\geq 99\%$ participants performed the test without adverse events (i.e., “sick feeling” and “pain in hand or forearm”) and instrument allergy, and 0% experienced DOMS¹⁴) to assess maximal isometric muscular strength in children and adolescents. One study⁹ reported that handgrip strength test was safe (DOMS males and females = 0%, HR bpm males = 95 ± 15.2 and females = 92 ± 15.3 , exertion $85\%HR_{max} = 0\%$ (participants exceeded the maximum percentage) and $\geq 95\%$ of participants performed the test without adverse events) to assess isometric muscular strength and in adults.

Only one low-quality study⁶ found that the handgrip strength test was safe (100% of participants perform the test without adverse events) to assess isometric muscular strength in children.

Explosive Fitness

Three high-quality studies^{9, 14, 17} assessed the safety of lower body explosive strength tests (see **Table 7**). Two studies^{14, 17} showed that the standing long jump test was safe (100% of the participants performed the test without adverse events or injuries¹⁷; $\geq 99\%$ participants performed the test without adverse events (i.e., “sick feeling” and “injury”) and 71.2% experienced some degree of DOMS between differences that could be the cause (i.e., 20-m shuttle run or standing long jump tests)¹⁴) to assess lower body explosive strength in children and adolescents. One study⁹ reported that the vertical jump test was safe (HR bpm males = 110 ± 16.3 and females = 112 ± 17.4 , exertion 85%HRmax = 0% (participants exceeded the maximum percentage) and, $\geq 95\%$ performed the test without adverse events) to assess lower body explosive strength in adults.

Only one low-quality study⁶ observed that the standing long jump test was safe (100% of the participants perform the test without adverse events) to assess lower body explosive strength.

Endurance Fitness

Two high-quality studies^{2, 9} analysed the safety of trunk endurance strength tests (see **Table 7**). The trunk flexor endurance and isometric back endurance tests were safe (100% of participants performed the test without adverse events^{9, 18}; HR bpm males = 112 ± 18.3 and females = 121 ± 19.8 , exertion 85%HRmax = 1.28% (participants exceeded the maximum) and, $\geq 90\%$ performed the test without adverse events⁹) to assess trunk endurance muscular strength in adults.

One high-quality study⁹ analysed the safety of upper body endurance strength tests (see **Table 7**). The modified push-ups test was safe (HR bpm males = 140 ± 18 and females = 144 ± 17.5 , exertion 85%HRmax = 9.2% (participants exceeded the maximum) and, $\geq 90\%$ performed the test without adverse events) to assess upper body muscular strength in adults.

Motor Fitness

Balance

Two high-quality studies^{9, 17} evaluated the safety of the balance tests (see **Table 7**). One study¹⁷ observed that the dynamic balance and the single-leg stand tests were safe (100% of participants performed the test without adverse events) to assess balance in children. Another study⁹ found that the single-leg stand test was safe (HR bpm males = 100 ± 16.1 and females = 94 ± 16 , exertion 85%HRmax = 0% (participants exceeded the maximum) and, $\geq 90\%$ performed the test without adverse events) to assess balance in adults.

Table 7. Overview of field-based fitness test safety studies by study quality.

Author	Participants	Age, range/ SD	Field-based fitness test	Safety items	Results	Conclusions
Low quality studies						
<i>Cardiorespiratory fitness</i>						
Borel et al., 2010	15 adults Healthy	29.0 ± 6.0	6-min step	-Heart rate ((HR), bpm) (mean (SD)) -Dyspnea (Borg scale 0-10) (mean (SD)) -Leg fatigue (Borg scale 0-10) (mean (SD))	HR = 118.2 (18.7). Borg scale dyspnea = 2.5 (1.5). Borg scale leg fatigue = 3.4 (2.5).	6-min step was considered safe to assess cardiorespiratory fitness.
Amado-Pacheco et al., 2019	90 children Males (n = 48) Females (n = 42) Healthy	4.04 ± 0.82 3.95 ± 0.82	20-m shuttle run	-Adverse events (%)	100% of participants perform the test without adverse events.	20-m shuttle run seems to be safe to assess cardiorespiratory fitness.
<i>Musculoskeletal fitness (isometric)</i>						
Amado-Pacheco et al., 2019	90 children Males (n = 48) Females (n = 42) Healthy	4.04 ± 0.82 3.95 ± 0.82	Handgrip	-Adverse events (%)	100% of participants perform the test without adverse events.	Handgrip was considered safe to assess isometric muscular strength.
<i>Musculoskeletal fitness (explosive)</i>						
Amado-Pacheco et al., 2019	90 children Males (n = 48) Females (n = 42)	4.04 ± 0.82 3.95 ± 0.82	Standing long jump	-Adverse events (%)	100% of participants perform the test without adverse events.	Standing long jump seems to be safe to asses lower body

Healthy

explosive strength.

*Musculoskeletal fitness
(flexibility)*

Amado-Pacheco et al., 2019	90 children Males (n = 48) Females (n = 42) Healthy	4.04 ± 0.82 3.95 ± 0.82	Sit and reach	-Adverse events (%)	100% of participants perform the test without adverse events.	Sit and reach was considered safe to assess lower body flexibility.
----------------------------	--	----------------------------	---------------	---------------------	---	---

*Motor fitness
(agility)*

Amado-Pacheco et al., 2019	90 children Males (n = 48) Females (n = 42) Healthy	4.04 ± 0.82 3.95 ± 0.82	4 × 10-m shuttle run	-Adverse events (%)	100% of participants perform the test without adverse events.	4 × 10-m shuttle run was considered safe to evaluate agility and speed.
----------------------------	--	----------------------------	----------------------	---------------------	---	---

High quality studies
Cardiorespiratory fitness

Bruggeman et al., 2020	40 Children Males (n = 23) Females (n = 17) Healthy	12.8 ± 1.9 13.3 ± 1.7 12.1 ± 2.0	45-s squat	-Rating of perceived exertion (RPE) (Borg scale 0-10) (mean (SD)) -HR (bpm) (mean (SD)) -Adverse events (%)	RPE = 3.4 (1.7). HR not available data. 100% of participants perform the test without adverse events.	45-s squat was considered safe to assess cardiorespiratory fitness.
			3-min step		RPE = 2.7 (1.4). HR not available data. 100% of participants perform the test without adverse events.	3-min step was considered safe to assess cardiorespiratory fitness.
Anderson & Dal Corso, 2016	20 adults Healthy	52 41 - 66	Chester step	-HR (bpm) (mean (SD)) -Dyspnea (Borg scale 0-10) (mean (range))	HR = 143 (28.0). Borg scale dyspnea = 2 (1-3).	Chester step was considered safe to assess

			Modified incremental step	-Leg fatigue (Borg scale 0-10) (mean (range)) -Adverse events (%)	Borg scale leg fatigue = 2 (1-3). 100% of participants perform the test without adverse events. HR = 142 (27.0). Borg scale dyspnea = 2 (1- 3). Borg scale leg fatigue = 2 (1-4). 100% of participants perform the test without adverse events.	cardiorespiratory fitness. Modified incremental step was considered safe to assess cardiorespiratory fitness.
Aadahl et al., 2012	795 adults Males (n=346) Females (n=449) Healthy	46.8 ± 8.2 46.7 ± 8.4 47.0 ± 8.0	Danish step	-Adverse events (n)	1 participant had a strained calf muscle after 30 sec, the test was not finished.	Danish step was considered safe to assess cardiorespiratory fitness.
Suni et al., 1998	600 adults Males (n = 246) Females (n = 354) Healthy	37 - 57 47.0 ± 7.9 47.0 ± 8.6	2-km walk	-Delayed-onset muscle soreness (DOMS) questionnaire (%) -HR (bpm) (mean (SD)) -Exertion 85%HRmax (%) -Adverse events (%)	DOMS male = 60%. DOMS female = 78%. HR males = 153 (17.4). HR females = 151 (18.3). Exertion 85%HRmax males = 43% Exertion 85%HRmax females = 37% exceeded the maximum. ≥95% of participants perform the test without adverse events.	2-km walk was considered safe to assess cardiorespiratory fitness.

Oja et al., 1991	159 adults Males (n=79) Females (n=80) Healthy	41 ± 13.7 42.9 ± 14.0 39.1 ± 13.4	2-km walk	-RPE: 0-10 points (mean (SD)) -HR (bpm) (mean (SD)) -Exertion perceived -Adverse events (%)	RPE males = 3.0 (1.8). RPE females = 2.9 (1.5). HR males = 153.1 (22.2). HR females = 153.8 (16.0). Exertion perceived as “moderate”. ≥70% of participants perform the test without adverse events.	2-km walk was considered safe to assess cardiorespiratory fitness.
Laukkanen et al., 1992	277 adults Males (n = 124) Females (n = 153) Healthy	44.5 ± 10.3 44.5 ± 10.2 44.5 ± 10.5	2-km walk	-RPE: Borg scale 0-10 (mean (SD)) -HR (bpm) (mean (SD))	RPE = 3.1 (2.4). Exertion perceived as “moderate”. HR males = 139.1 (19.1) HR females = 143.5 (17.1). HR = 128 (25.0). Borg scale dyspnea = 2 (1-3).	2-km walk was considered safe to assess cardiorespiratory fitness.
Anderson & Dal Corso, 2016	20 adults Healthy	52 41 - 66	6-min walk	-HR (bpm) (mean (SD)) -Dyspnea (Borg scale 0-10) (mean (range)) -Leg fatigue (Borg scale 0-10) (mean (range)) -Adverse events (%)	Borg scale dyspnea = 2 (1-3). Borg scale leg fatigue = 2 (1-3). 100% of participants perform the test without adverse events.	6-min walk seems to be safe to assess cardiorespiratory fitness.
España-Romero et al., 2010	58 children 80 adolescents Healthy	6 - 11.9 12 - 18	20-m shuttle run	-Questionnaire with items of dichotomous answer (%) -DOMS questionnaire (%)	Sick feeling = 0.8%. Injury = 0.8%. 71.2% experienced some degree of DOMS; of them, the 29.3% assumed that the 20 m shuttle run test could be the cause.	20-m shuttle run was considered safe to assess cardiorespiratory fitness.

Lamoneda et al., 2020	386 adolescents Males (n =189) Females (n = 197) Healthy	14.5 ± 1.6 14.4 ± 1.6 14.5 ± 1.6	20-m shuttle run music vs 20-m shuttle run	-RPE: Borg scale 0-10 (mean (SD)) -HR (bpm) (mean (SD))	RPE = 8.3 (2.0) vs. 7.9 (2.2). HR = 177.1 (30.6) vs. 177.7 (29.2).	20-m shuttle run music and 20-m shuttle run were considered safe to assess cardiorespiratory fitness.
<i>Musculoskeletal fitness (isometric)</i>						
Hébert et al., 2011	33 children 41 adolescents Males (n = 37) Females (n = 37) Healthy	4 - 9.9 10 - 17.5	Handgrip	-Adverse events (%)	100% of participants perform the test without adverse events (i.e., pain or discomfort).	Handgrip was considered safe to assess isometric muscular strength.
España-Romero et al., 2010	58 children 80 adolescents Healthy	6 - 11.9 12 - 18	Handgrip	-Questionnaire with items of dichotomous answer (%) -DOMS questionnaire (%)	Instrument allergy = 0%, Injury = 0%, sick feeling = 0%, pain in hand or forearm = 0.8%. DOMS= 0%.	Handgrip was considered safe to assess isometric muscular strength.
Suni et al., 1998	600 adults Males (n = 246) Females (n = 354) Healthy	37 - 57 47.0 ± 7.9	Handgrip	-DOMS questionnaire (%) -HR (bpm) (mean (SD)) - Exertion 85%HRmax (%) -Adverse events (%)	DOMS male = 0% DOMS female = 0% HR males = 95 (15.2). HR females = 92 (15.3). Exertion 85%HRmax = 0% participants exceeded de maximum percentage. ≥95% of participants perform the test without adverse events.	Handgrip was considered safe to assess isometric muscular strength.
<i>Musculoskeletal fitness (explosive)</i>						

Smits-Engelsman et al., 2020	80 children Males (n = 39) Females (n = 41) Healthy	9.2 ± 1.1	Standing long jump	-Adverse events (%)	100% of participants perform the test without adverse events or injuries.	Standing long jump seems to be safe to assess lower body explosive strength.
España-Romero et al., 2010	58 children 80 adolescents Healthy	6 - 11.9 12 - 18	Standing long jump	-Questionnaire with items of dichotomous answer (%) -DOMS questionnaire (%)	Sick feeling = 0. Injury = 0%. 71.2% experienced some degree of DOMS between differences tests that could be the cause (i.e. 20-m shuttle run or standing long jump).	Standing long jump seems to be safe to assess lower body explosive strength.
Suni et al., 1998	600 adults Males (n = 246) Females (n = 354) Healthy	37 - 57 47.0 ± 7.9	Vertical jump	-DOMS questionnaire (%) -HR (bpm) (mean (SD)) -Exertion 85%HRmax (%) -Adverse events (%)	DOMS male = 0%. DOMS female = 0%. HR males = 110 (16.3). HR females = 112 (17.4). Exertion 85%HRmax = 0% participants exceeded de maximum percentage. ≥95% of participants perform the test without adverse events.	Vertical jump seems to be safe to assess lower body explosive strength.
<i>Musculoskeletal fitness (endurance)</i>						
Ito et al., 1996	90 adults Males (n = 37) Females (n = 53) Healthy	44.3 46.8	Trunk flexor endurance	-Adverse events (%)	100% of participants perform the test without adverse events (pain, problems).	Trunk flexor endurance and isometric back endurance tests seem to be safe to assess trunk endurance muscular strength.
			Isometric back endurance		100% of participants perform the test without adverse events (pain, problems).	

Suni et al., 1998	600 adults Males (n = 246) Females (n = 354) Healthy	37 - 57 47.0 ± 7.9	Isometric back endurance	-DOMS questionnaire (%)	DOMS male = 0%. DOMS female = 0%. HR males = 112 (18.3). HR females = 121 (19.8). Exertion 85%HRmax = 1.28% participants exceed the maximum.	Isometric back endurance was considered safe to assess trunk endurance muscular strength.
			Modified push-ups	-HR (bpm) (mean (SD))	≥90% of participants perform the test without adverse events.	
				-Exertion 85%HRmax (%)		
				-Adverse events (%)		
<i>Motor fitness (balance)</i>						
Smits-Engelsman et al., 2020	80 children Males (n = 39) Females (n = 41) Healthy	9.2 ± 1.1	Single-leg stand	-Adverse events (%)	100% of participants perform the test without adverse events.	Single-leg stand and dynamic balance tests were considered safe to evaluate balance.
			Dynamic balance		100% of participants perform the test without adverse events.	
Suni et al., 1998	600 adults Males (n = 246) Females (n = 354) Healthy	37 - 57 47.0 ± 7.9	Single-leg stand	-DOMS questionnaire (%)	DOMS male = 0%. DOMS female = 0%. HR males = 100 (16.1). HR females = 94 (16.0).	Single-leg stand was considered safe to evaluate balance.
				-HR (bpm) (mean (SD))		

-Exertion 85%HRmax (%)	Exertion 85%HRmax = 0% participants exceeded de maximum. percentage. ≥90% of participants perform the test without adverse events.
-Adverse events (%)	

Adverse events refer to falls, sick feeling, pain and/or injuries.

Study II. Feasibility and Safety in health-related field-based physical fitness test in adult population: The ADULT-FIT project.

Feasibility

Body composition

The final sample was composed of 390 adults aged 18-64 years. The data collected from the feasibility items for the assessment of body composition (i.e., weight, height, neck, waist and hip circumferences, and triceps and subscapular skinfolds) are presented in **Table 8**. All the items assessed (i.e., “appropriate facilities”, “appropriate equipment”, “appropriate sport clothes”, “test’s instruction understood” and “easy to measure”) presented a successful answer of 100%. The “ratio participant/evaluator” was 1/1 for the assessment.

The “time of preparation” of the tests took 71.7 (11.8) seconds and for “time of performance” ranged from 3.5 to 5.5 minutes by participant.

Cardiorespiratory fitness tests

The data collected from the feasibility items for the assessment of cardiorespiratory fitness tests are presented in **Table 8**. All the items assessed (i.e., “appropriate facilities”, “appropriate equipment”, “appropriate sport clothes”, “test’s instruction understood” and “easy to measure”) presented a successful answer of 100%.

The “ratio participant/evaluator” was 10/2 for the correct assessment of the cardiorespiratory fitness tests (i.e., 2-km walk, 6-min walk and 20-m shuttle run tests). Referring to the item "time of preparation", the 2-km walk test took 75.3 (10.7) seconds, the 6-min walk test 75.3 (10.7) seconds, and the 20-m shuttle run test 114.01 (22.3) seconds.

The “time of performance” for the 2-km walk test ranged from 11 to 23 minutes, the 6-min walk test the time was set by the test criterion (i.e., 6 minutes). Finally, the time to perform the 20-m shuttle run test ranged from 1.5 to 13 minutes (which also correspond to number of stages).

Table 8. Feasibility of the measurement protocol of field-based body composition and cardiorespiratory fitness tests.

	Body composition ¹		6-min walk test		2-km walk test		20-m shuttle run test	
	Yes	No	Yes	No	Yes	No	Yes	No
Appropriate facilities (n/%)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)
Appropriate equipment (n/%)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)
Appropriate sport clothes (n/%)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)
Test's instruction understood (n/%)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)
Easy to administer (n/%)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)
Participants evaluated/not evaluated (n)	390/0		390/0		390/0		390/0	
Ratio participants/evaluators (n)	1/1		10/2		10/2		10/2	
Time of preparation (sec) [#]	71.7 (11.8)		75.3 (10.7)		75.3 (10.7)		114.01 (22.3)	
Time of performance (min)	3.5 - 5.5 ^{2§}		6.0*		11 - 23 [§]		1.5 - 13 [§]	

[#] Data are shown in mean and (standard deviation); [§] data are shown in minimum and maximum time values; *time established by test.

¹Body composition measures were: weight, height, neck, waist and hip circumferences and triceps and subscapular skinfolds.

²Time of performance is the sum of two non-consecutive measures.

Muscular fitness

The data collected from the feasibility items for the assessment of muscular fitness tests are presented in **Table 9**. All the items assessed (i.e., “appropriate facilities”, “appropriate equipment”, “appropriate sport clothes”, “test’s instruction understood” and “easy to measure”) presented a successful answer of 100%. The “ratio participant/evaluator” for the handgrip strength and the standing long jump tests was 1/1. Regarding the 30-s sit-to-stand and front plank tests the ratio was 1/2.

The “time of preparation” was 39.1 (9.3) seconds for the handgrip strength test, 69.1 (10.3) seconds for the standing long jump test, 60.8 (9.4) seconds for the 30-s sit-to-stand test and 77 (11.7) seconds for the front plank test.

Referring to “time of performance” item, the handgrip strength test ranged from 1.8 to 3 minutes (including two attempts and 1-minute rest between attempts), the standing long jump test from 1.3 to 2.3 minutes (also including two attempts and 1-minute rest between attempts) and the front plank test from 1.1 to 9.3 minutes. In the 30-s sit-to-stand test, the time to perform the test was set by the test criterion (i.e., 30 seconds).

Motor fitness

The data collected from the feasibility items for the assessment of motor fitness tests are presented in **Table 9**. All the items assessed (i.e., “appropriate facilities”, “appropriate equipment”, “appropriate sport clothes”, “test’s instruction understood” and “easy to measure”) presented a successful answer of 100%. For the proper assessment of the tests, the “ratio participant/evaluator” was 1/2 for the 6-m gait speed, 2.45-m time up & go and single-leg stand tests, and 1/1 for the 4x10-m shuttle run test.

Regarding "time of preparation" for the 6-m gait speed test was 55.8 (9.9) seconds, 52.3 (8.5) seconds for the 2.45-m time up & go test, 91.8 (10.3) seconds for the 4x10-m shuttle run test, and 55.3 (9.4) seconds for single-leg stand test.

Regarding to the “time of performance” item, the 6-m gait speed test ranged from 1.8 to 3.6 seconds, the 2.45-m time up & go test ranged from 4 to 7 seconds, the 4x10-m shuttle run test ranged from 1.15 to 1.28 minutes, and for the single-leg stand test the time was already established by the test criterion; however, the assessment took between 3 to 3.6 minutes per participant (including 30 seconds for each leg and 1-minute rest between eyes open and eyes closed modality).

Table 9. Feasibility of the measurement protocol of field-based muscular and motor fitness tests.

	Handgrip strength		Standing long jump		30-s sit-to-stand		Front plank		6-m gait speed		2.45-m time. up & go		4x10-m shuttle run		Single-leg stand	
	Yes	No	Yes	Yes	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Appropriate facilities (n/%)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)
Appropriate equipment (n/%)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)
Appropriate sport clothes (n/%)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)
Test's instruction understood (n/%)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)
Easy to measure (n/%)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)	390 (100)	0 (0)
Participants evaluated/not evaluated (n)	390/0		390/0		390/0		390/0		390/0		390/0		390/0		390/0	
Ratio participants/evaluators (n)	1/1		1/1		1/2		1/2		1/2		1/2		1/1		1/2	
Time of preparation (s) [#]	39.1 (9.3)		69.1 (10.3)		60.8 (9.4)		77 (11.7)		55.8 (9.9)		52.3 (8.5)		91.8 (10.3)		55.3 (9.4)	
Time of performance (min or <u>sec</u>)	1.8 - 3.0 ^{1 §*}		1.3 - 2.3 ^{1 §*}		<u>30</u> ^{**}		1.1 - 9.3 ^{§*}		<u>1.8 – 3.6</u> [§]		<u>4 – 7</u> [§]		1.15 - 1.28 ^{1 §*}		3.0 - 3.6 ^{1 §*}	

[#] Data are shown in mean and (standard deviation); [§] data are shown in minimum and maximum time values; *average minutes per participant; **time established by test.

¹ Time of performance is the sum of two attempts and 1-minute rest between attempts.

Safety

The data collected from the safety items for the assessment of the health-related field-based physical fitness tests answered are presented in **Table 10**. No complications occurred during the testing procedure referred to “instruments allergy” and/or “adverse events” (i.e., falls, injury or sick feeling), except one participant (1.4%) that felt sick at the end of the 20-m shuttle run test.

The mean (SD) value of “HR” in the 2-km walk test was 144 (7.1) bpm, 134.9 (9.3) bpm for 6-min walk test, and 169.2 (5.7) bpm for 20-m shuttle run test. None of the participants reached its maximum HR in the 2-km walk and 6-min walk tests; however, the 22.8% of participants reached their maximum HR in the 20-m shuttle run test.

Regarding the “85%HRmax” item, the 32% of the participants reached it in the 2-km walk test. From this percentage, the 59% were females, the 45% were physically active and belonged to two age groups (26% for 35-49 year and 74% for 50-64 years). For the 6-min walk test, only the 7% of the participants reached their “85%HRmax”, from this percentage the 79% were females, the 32% were physically active, and belonged to two age groups (7% for 35-49 year and 93% for 50-64 years). For the 20-m shuttle run test, the 97% of the participants reached their “85%HRmax”, this percentage was balanced by sex (50%) and physical activity level (50%), but not by age groups (28 % for 18-34 years, 36% for 35-49 years and 36% for 50-64 years).

For “RPE” item, the 39.2% of participants reported a “very light/light” effort, 51.3% a “moderate” effort and 9.5% a “hard” effort in the 30-s sit-to-stand test. The 11.8% of participants reported a “moderate” effort, 60.8% a “hard” effort, 24.9% a “very hard” effort and 2.5% a “maximum effort” in the front plank test.

The 31.8% of the participants expressed having felt pain associated with the anterior tibial muscles at the end of the 2-km walk test. Of them, the 37.2% reported “low”, 34.6% “moderate” and 28.2% “severe” tibial pain.

The “DOMS” questionnaire had a response rate of 381 participants (97.7%). The 44.6% of the participants expressed some degree of “DOMS”. Among the responders, the 58.8% experienced “very low/low” degree, 25.3% “moderate” degree, 12.9% “quite” degree and the 2.9% “severe” degree. Moreover, the 89.4% of the sample that answered

the questionnaire, reported pain in “both legs”, 6.5% in the “lower zone back and both legs” and the 4.1% reported pain in “both arms and legs”. The 95.3% of participants reported that cardiorespiratory fitness tests (i.e., the 2-km walk, the 6-min and the 20-m shuttle run tests) could be the cause of their “DOMS”; however, the 4.7% of participants reported that the standing long jump test could be the cause. And finally, regarding to the difficulty, they may have experienced in carrying out daily activities due to muscular soreness, the 18% reported difficulties associated to this. Of those who did experience difficulties, the 58% reported that felt them when they “descending stairs/climbing stairs”, the 28% reported “when bending down to pick up an object”, and 14% reported “other” (i.e., difficult to walk normally).

Feasibility and safety were not affected by sex, age and level of physical activity, except in the “85%HRmax” item, since as age increased, more participants reached the “85%HRmax”.

Table 10. Safety of the measurement protocol of field-based fitness test answered by evaluators and participants.

Field-based physical fitness test	Instrument allergy n (%)	Adverse events			HR [§]	85% HRmax	RPE [§]	Tibial pain [#]	DOMS*
		Falls n (%)	Injury n (%)	Sick feeling n (%)	Mean (SD)	n (%)	n (%)	n (%)	n (%)
Body composition	0 (0)	0 (0)	0 (0)	0 (0)	n/a	n/a	n/a	n/a	n/a
6-min walk	0 (0)	0 (0)	0 (0)	0 (0)	134.9 (9.3)	28 (7.2)	n/a	n/a	
2-km walk	0 (0)	0 (0)	0 (0)	0 (0)	144 (7.1)	125 (32.1)	n/a	No pain: 266 (68.2) Pain: 124 (31.8): low: 46 (11.8), moderate: 43 (11), severe: 35 (9)	162 (95.3) ¹
20-m shuttle run	0 (0)	0 (0)	0 (0)	1 (0.3)	169.2 (5.7)	378 (97.0)	n/a	n/a	
Handgrip strength	0 (0)	0 (0)	0 (0)	0 (0)	n/a	n/a	n/a	n/a	n/a
Standing long jump	0 (0)	0 (0)	0 (0)	0 (0)	n/a	n/a	n/a	n/a	8 (4.7)
30-s sit-to-stand	0 (0)	0 (0)	0 (0)	0 (0)	n/a	n/a	Very light/light: 153 (39.2) Moderate: 200 (51.3) Hard: 37 (9.5)	n/a	n/a
Front plank	0 (0)	0 (0)	0 (0)	0 (0)	n/a	n/a	Moderate: 46 (11.8) Hard: 237 (60.8) Very hard: 97 (24.9) Maximum effort: 10 (2.5)	n/a	n/a
6-m gait speed	0 (0)	0 (0)	0 (0)	0 (0)	n/a	n/a	n/a	n/a	n/a
2.45-m time up & go	0 (0)	0 (0)	0 (0)	0 (0)	n/a	n/a	n/a	n/a	n/a
4x10 shuttle run	0 (0)	0 (0)	0 (0)	0 (0)	n/a	n/a	n/a	n/a	n/a
Single-leg stand	0 (0)	0 (0)	0 (0)	0 (0)	n/a	n/a	n/a	n/a	n/a

DOMS: Delayed-onset muscle soreness; HR: heart rate; HR max.: maximum heart rate; RPE: Rate of perceived exertion; SD: standard deviation.

[§]RPE Borg scale (range 1-10): very light/light (1-3), moderate (4-6), hard (7-8), very hard (9), and maximum effort (10).

[#]Tibial pain self-perception (scale range 0-10): no pain (0), low (1-3), moderate (4-6), and severe (7-10) tibial pain.

*DOMS self-reported questionnaire answered by participants 48 hours post evaluation of the field-based physical fitness tests.

¹Participants reported that cardiorespiratory fitness tests (i.e., the 6-min walk, the 2-km walk, and the 20-m shuttle run tests) could be the cause of their DOMS.

Study III. Responsiveness of health-related field-based physical fitness tests in adults: The ADULT-FIT project.

Sixty participants finally completed the study (62% female); 31 and 29 in control and intervention group, respectively. The 94% of the participants ($n=29$) in the intervention group completed, at least, the 80% of the intervention program sessions. There was a missing value for 20-m shuttle run, 4x10-m shuttle run, handgrip strength, standing long jump and front plank, and 2 missing values for 6-m gait speed test, otherwise, all the participants completed both assessments.

The main baseline characteristics of participants are shown in **Table 11**. There were no baseline differences between the groups in any tests.

Table 11. Participants baseline characteristics.

	CONTROL GROUP		INTERVENTION GROUP		<i>p-values</i>
	(n=31)		(n=29)		
Sex (female %)	18 (58)		19 (66)		0.743
Age (years)	42.13	11.68	45.41	10.04	0.249
Height (cm)	167.19	10.17	165.44	7.74	0.458
Weight (kg)	74.36	15.42	74.12	19.36	0.958
Body mass index (kg/m ²)	26.47	4.56	26.85	5.77	0.782
Waist circumference (cm)	87.58	12.10	86.96	15.39	0.862
20-m shuttle run (stage)	3.05	2.09	2.84	1.92	0.691
2-km walk (s)	1107.45	102.43	1076.93	91.71	0.230
6-min walk (m)	678.51	58.55	680.56	54.64	0.889
6-m gait speed (s)	2.40	0.22	2.44	0.24	0.448
2.45-m time up & go (s)	5.74	0.63	5.67	0.62	0.664
4x10-m shuttle run (s)	14.08	1.68	14.39	1.55	0.473
Handgrip strength* (kg)	32.75	10.23	31.53	10.24	0.645
Standing long jump (cm)	134.61	33.67	130.17	29.13	0.588
30-s sit-to-stand (rep)	16.97	2.76	16.24	2.06	0.256
Front plank (s)	80.48	32.17	75.48	31.18	0.544

Results are presented as mean (SD) except for sex variable n (%). *Handgrip righth and left arm mean.

Responsiveness of health-related field-based physical fitness tests analysed are shown in **Table 12**. There were no differences between pre-post-tests in the control groups; whereas in the intervention group, differences pre-post tests were observed in all tests (all $P<0.01$), except in waist circumference ($P=0.091$). Additional analyses were performed adjusted for sex and age and the results remained unchanged (data not shown). The 6-min walk, 2-km walk, 30-s sit-to-stand, front plank, 6-m gait speed and 2.45-m time up & go tests showed a large effect size (*Cohen's* $d>0.8$). A moderate effect size was found in the 20-m shuttle run, 4x10-m shuttle run and standing long jump tests (*Cohen's* $d>0.5-0.8$). While the weight, BMI and handgrip tests showed a trivial effect size (*Cohen's* $d<0.20$). In absolute data, the difference in weight, BMI, waist circumference and handgrip test pre-post intervention were -0.94 kg, -0.5kg/m², -0.84 cm and 1.3 kg, respectively. Additional analyses were performed using the data of handgrip strength test relative to corporal weight and the results remained unchanged (data not shown).

Table 12. Pre-post differences for each groups and *percentage of the population that is expected to respond to the intervention.**Handgrip righth and left arm mean. Δ : difference on post to pre-test. Effect size values <0.2 were considered trivial, >0.2-0.5, small, >0.5-0.8, moderate and >0.8 large. %Change =

TESTS	CONTROL GROUP					INTERVENTION GROUP						
	(n=31)					(n=29)						
	Δ (95% IC)		<i>P</i> -values	Cohen's d	% Change	Δ (95% IC)		<i>P</i> -values	Cohen's d	% Change	σ_{IR}	% Responder
Weight (kg)	0.31	(-0.40; 1.02)	0.390	0.02	0.4	-0.94	(1.67; -0.20)	0.014	0.05	-1.3	2.31	13
Body mass index (kg/m ²)	0.09	(-0.37; 0.19)	0.543	-0.01	0.3	-0.45	(-0.16; 0.74)	0.003	0.03	-1.7	0.92	26
Waist circumference (cm)	-0.31	(-1.26; 0.63)	0.512	0.03	-0.4	-0.84	(-1.82; 0.14)	0.091	0.05	-1.0	2.58	23
20-m shuttle run (stage)	-0.15	(-0.41; 0.12)	0.276	0.07	-4.8	1.30	(1.03; 1.58)	<0.001	0.62	45.9	0.82	86
2-km walk (s)	1.29	(-16.45; 13.87)	0.865	0.01	-0.1	-82.62	(98.29; 66.95)	<0.001	0.93	-7.7	36.44	96
6-min walk (m)	-2.82	(-15.25; 9.61)	0.651	0.05	-0.4	62.50	(49.66; 75.35)	<0.001	1.04	9.2	40.11	90
6-m gait speed (s)	0.02	(-0.07; 0.09)	0.697	0.06	0.6	-0.33	(-0.41; -0.24)	<0.001	1.19	-13.4	0.26	89
2.45-m time up & go (s)	-0.13	(-0.29; 0.02)	0.096	0.19	-2.3	-0.71	(-0.87; -0.55)	<0.001	1.29	-12.5	0.30	99
4x10-m shuttle run (s)	0.05	(-0.14; 0.24)	0.994	0.03	0.4	-1.23	(-1.43; 1.02)	<0.001	0.80	-8.1	0.52	99
Handgrip strength* (kg)	-0.22	(-0.84; 0.40)	0.472	0.03	-1.0	1.35	(0.72; 1.98)	<0.001	0.13	4.3	1.66	33
Standing long jump (cm)	-1.68	(-5.18; 1.82)	0.341	0.05	-1.2	17.25	(13.57; 20.93)	<0.001	0.59	14.0	6.09	96
30-s sit-to-stand (rep)	-0.19	(-0.79; 0.40)	0.516	0.07	-1.1	3.62	(3.01; 4.23)	<0.001	1.50	22.3	1.98	94
Front plank (s)	-7.07	(-16.72; 2.59)	0.148	0.21	-8.8	44.11	(34.0; 54.27)	<0.001	1.06	59.7	28.69	91

(mean (Post-test value) / mean (Pre-test value) / mean (Pre-test value)) * 100. σ_{IR} : intervention response standard deviation; %Responders: percentage of the experimental group that respond to the intervention.

Figures 3 to 6 showed participants categorized into responders / no responders to intervention based on individual change (Δ), 80%CI and SWC in all tests. Responder analysis showed a higher proportion of participants categorized as responders in the intervention group compared to the control group in all the variables analyzed. Among all of tests evaluated in the intervention group, we observed the highest proportion of individual responders in the 4x10-m shuttle run, 30-s sit-to-stand, 6-m gait speed and 2.45-m time up & go tests (78%, 69%, 68 and 66% of responders, respectively) (**Figures 5-6**). While in the 6-min walk, front plank, 2-km walk, 20-m shuttle run and standing long jump tests the proportion of individual responders were 57%, 54%, 52%, 46% and 46%, respectively. The lower proportion of individual responders was observed in the weight, BMI, waist circumference and handgrip tests (<15%) in the intervention group. The proportion of responders in the control group ranged between 0 and 7% in all the tests studied, except in the 2.45-m time up & go test with 23%.

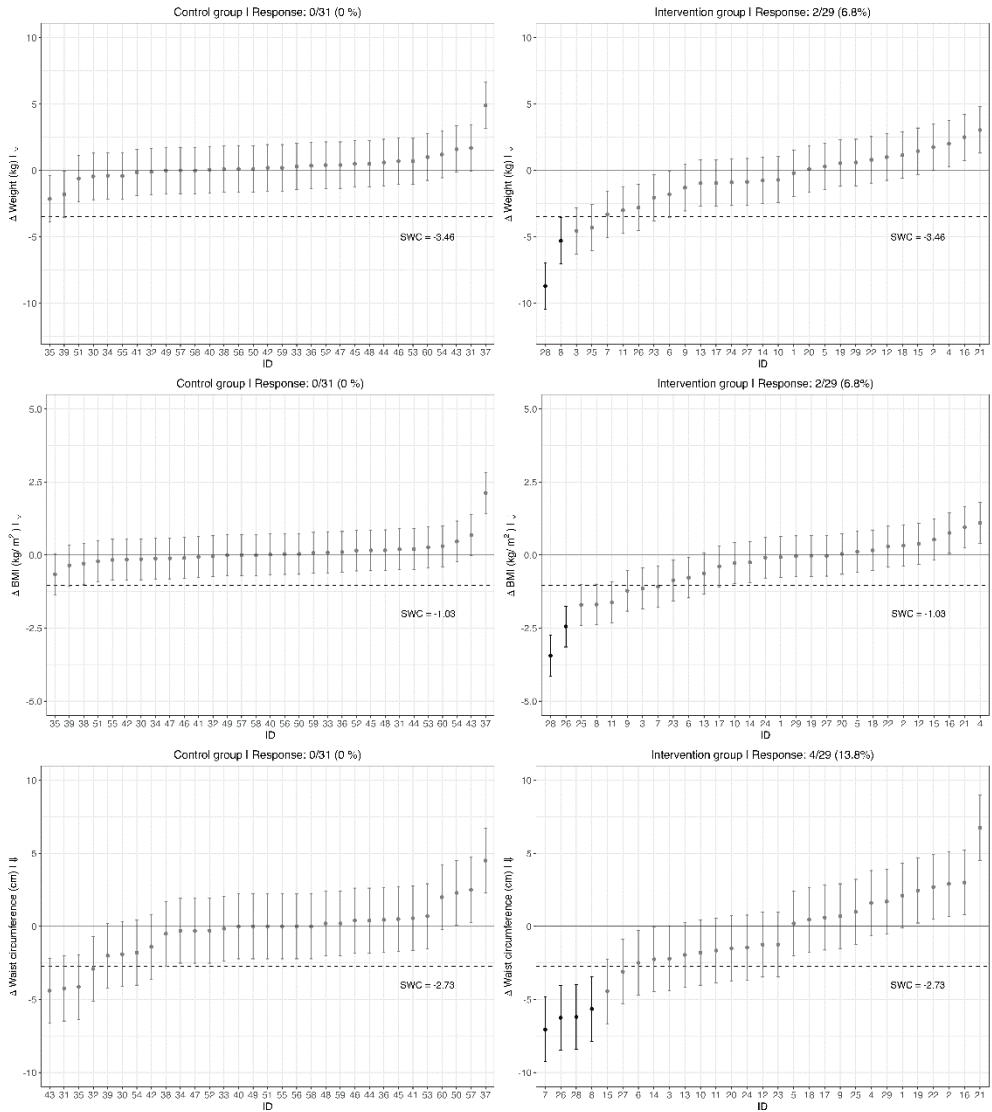


Figure 3. Participants categorized into responders / no responders to intervention in weight, BMI and waist circumference variables based on individual change (Δ), 80%CI and Smallest Worthwhile Change (SWC). Arrow next to y-axis title represents direction of the positive effect for each variable. Participants were classified as likely responder if the CI lies completely over or below the SWC (horizontal dotted line) depending of the interpretation of positive effect.

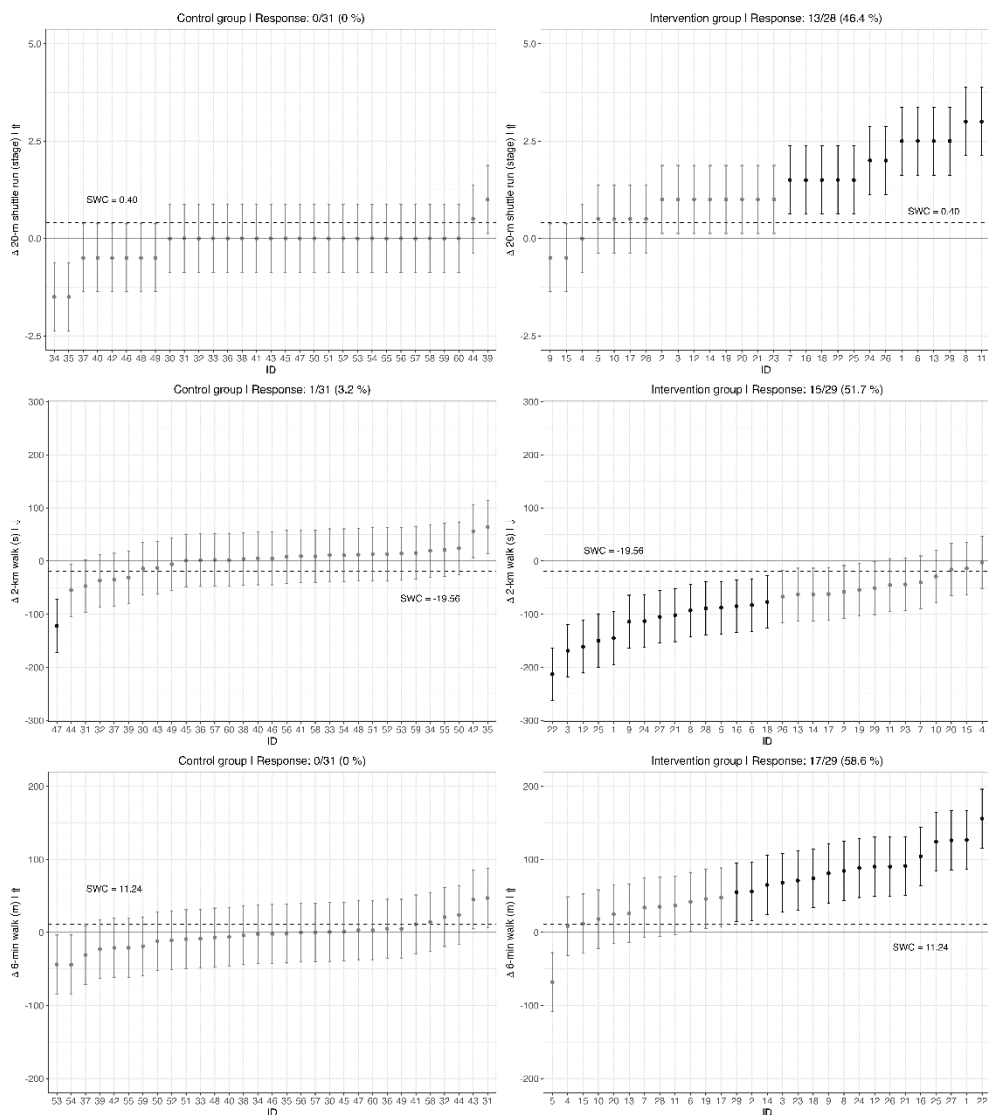


Figure 4. Participants categorized into responders / no responders to intervention in 20-m shuttle run, 2-km walk and 6-min walk tests based on individual change (Δ), 80%CI and Smallest Worthwhile Change (SWC). Arrow next to y-axis title represents direction of the positive effect for each variable. Participants were classified as likely responder if the CI lies completely over or below the SWC (horizontal dotted line) depending of the interpretation of positive effect.

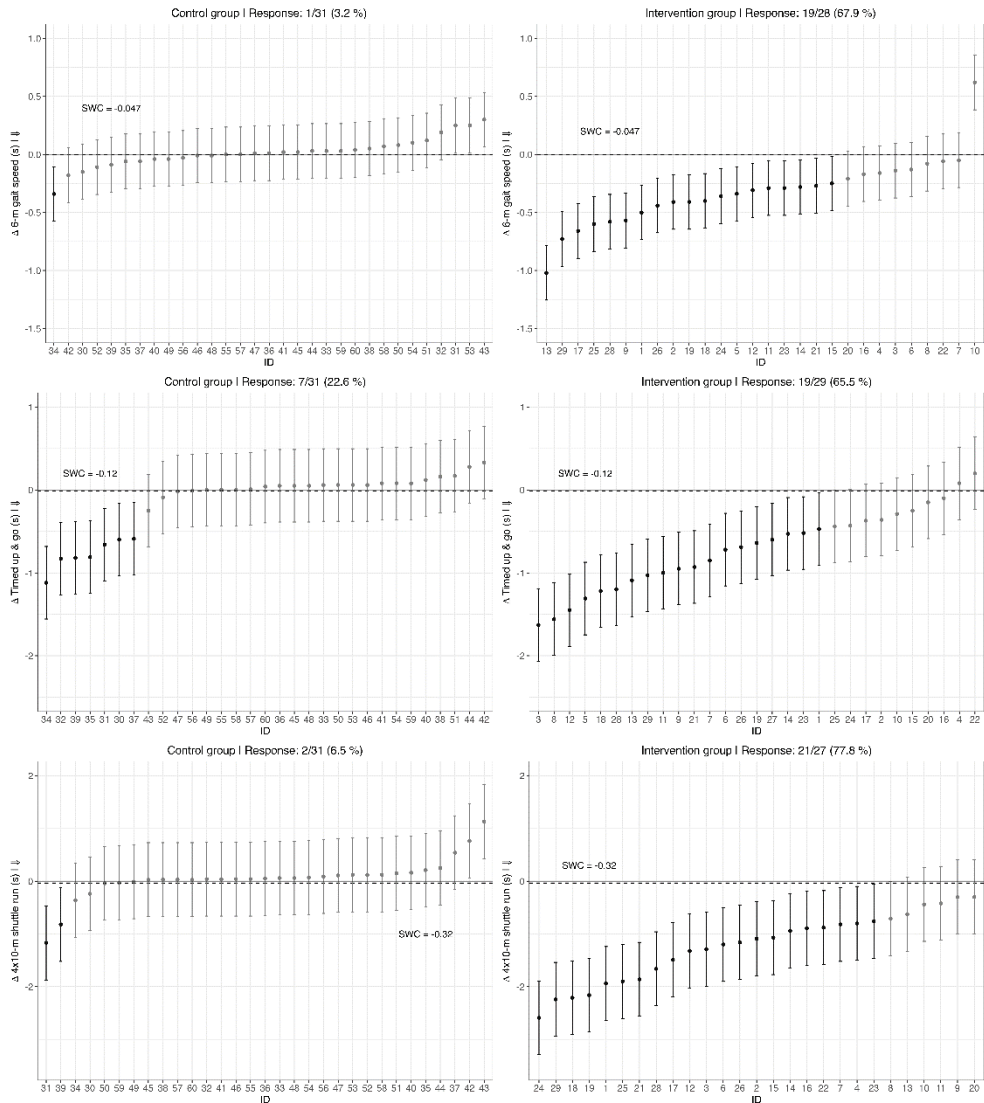


Figure 5. Participants categorized into responders / no responders to intervention in 6-m gait speed, 2.45-m time up & go and 4x10-m shuttle run tests based on individual change (Δ), 80%CI and Smallest Worthwhile Change (SWC). Arrow next to y-axis title represents direction of the positive effect for each variable. Participants were classified as likely responder if the CI lies completely over or below the SWC (horizontal dotted line) depending of the interpretation of positive effect.

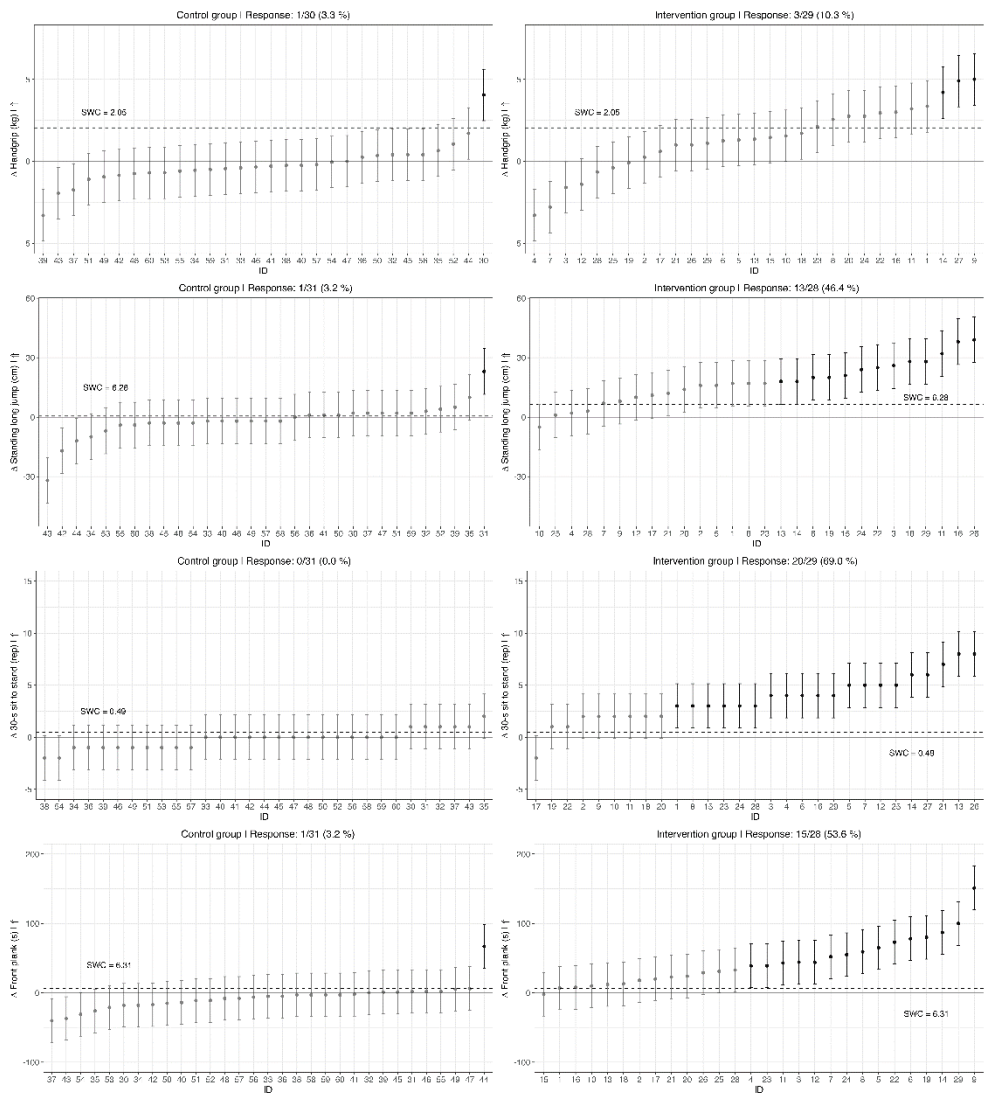


Figure 6. Participants categorized into responders / no responders to intervention in handgrip strength, standing long jump, 30-s sit-to-stand and front plank tests based on individual change (Δ), 80%CI and Smallest Worthwhile Change (SWC). Arrow next to y-axis title represents direction of the positive effect for each variable. Participants were classified as likely responder if the CI lies completely over or below the SWC (horizontal dotted line) depending of the interpretation of positive effect.

The percentage of the population that is expected to respond to the intervention are presented in **Table 12**. The expected proportion of responders based on σ_{IR} and SWC was higher to 85% for the 4x10-m shuttle run (99%), 2.45-m time up & go (99%), 2-km walk (96%), standing long jump (96%), 30-s sit-to-stand (94%), front plank (91%), 6-min walk (90%), 6-m gait speed (89%) and 20-m shuttle run (86%) tests (**Table 12**). While for the weight, BMI, waist circumference and handgrip tests were lower than 35%.

DISCUSSION

DISCUSSION

The specific discussions for each study are shown below.

Study I. Feasibility and Safety of Field-Based Physical Fitness Test: A Systematic Review.

The present systematic review comprehensively studied the feasibility and the safety of the existing field-based physical fitness tests used in people of all ages. The findings of this review provide an evidence-based proposal for most feasible and safe field-based physical fitness tests for population of all ages.

The validity^{1,2} and the reliability^{3,4} of field-based physical fitness tests have been widely studied in children, adolescents and adults. However, in this review we have found that few studies have addressed the feasibility and the safety in children and adolescents, adults or/and older adults.

Only the 2-km walk, 20-m shuttle run, handgrip strength and standing long jump tests showed strong evidence on feasibility (see **Table 13**). In neither case can this evidence be generalized for the three age groups (i.e., children and adolescents, adults or/and older adults), as there were no studies with strong evidence in older adults. Thus, there was no homogeneity in the study of the tests feasibility by age groups.

Regarding safety, only the 2-km walk test was shown to have strong evidence in adults. The evidence presented was only in youth and adults. The evidence was only analysed in 11 studies (see **Table 14**). Therefore, we found that most field-based physical fitness tests, which are often widely used in one or more age groups, were not shown to have strong or moderate evidence. Based on these findings, more studies are needed to determine the feasibility and the safety of field-based physical fitness test in people of all ages.

Levels of evidence

Feasibility

Cardiorespiratory fitness

Strong evidence indicated that the 2-km walk and 20-m shuttle run tests were feasible to assess cardiorespiratory fitness in adults and children and adolescents, respectively. Limited evidence showed that the Reduced Cooper test in children and

adolescents, the Danish step test in adults, and the 6-min walk and 2-min step tests in older population were feasible to assess cardiorespiratory fitness.

Musculoskeletal fitness

Strong evidence indicated that a) the handgrip strength test was feasible to assess isometric muscular strength in children and adolescents; and b) the standing long jump test was feasible to assess lower body explosive strength in children and adolescents. Moderate evidence indicates that the isometric back endurance test was feasible to assess trunk endurance strength in adults. Limited evidence indicated that a) the handgrip strength test was feasible to assess isometric muscular strength in adults; b) the vertical jump test in adults and, the jumping a distance of 7-m on 1 feet and jumping a distance of 7-m on 2 feet tests in children and adolescents were feasible to assess lower body explosive strength; c) the pushing a medicine ball test was feasible to assess upper body explosive strength in children and adolescents; d) the 60-s, 90-s and unlimited plank tests in children and adolescents; the trunk flexor endurance test in adults; and the partial curl-ups test in children and adolescents were feasible to assess trunk endurance strength; e) the 30-s sit-to-stand test was feasible to assess lower body endurance strength in older adults; f) the arm curl test in older adults; and the modified push-ups test in adults were feasible to assess upper body endurance strength.

Flexibility

Limited evidence indicated that the sit and reach test in children, the chair sit and reach in older adults and the back scratch test in older adults were feasible to assess flexibility.

Motor fitness

Moderate evidence indicated that a) the 4 x 10-m shuttle run test was feasible to assess agility in children and adolescents; and b) the single-leg stand test was feasible to assess balance and lower body strength in children and adolescents. Limited evidence indicated that a) the running 20-m test was feasible to assess speed in children and adolescents; b) the 5 x 10-m shuttle run test in children and adolescents and the 2.45-m time up & go test in older adults were feasible to assess agility; c) the dynamic balance

test in children and adolescents, and the single-leg stand test in adults were feasible to assess balance and lower body strength.

Table 13. Levels of evidence of feasibility in field-based fitness tests by age group.

Field-based fitness tests	Children/Adolescents	Adults	Older adults
<i>Cardiorespiratory fitness</i>			
Distance and time-based walk			
2-km walk		●	
6-min walk			○
Reduced Cooper	○		
Shuttle run			
20-m shuttle run	●		
Step			
Danish step		○	
2-min step			○
<i>Musculoskeletal fitness</i>			
Maximal isometric fitness			
Handgrip strength	●	○	
Lower body explosive fitness			
Standing long jump	●		
Vertical jump		○	
Jumping a distance of 7-m on 1 feet	○		
Jumping a distance of 7-m on 2 feet	○		
Upper body explosive fitness			
Pushing medicine ball	○		
Trunk endurance fitness			
60-s, 90-s and unlimited plank	○		
Trunk flexor endurance		○	
Isometric back endurance		◐	
Partial curl-ups	○		
Lower body endurance fitness			
30-s sit-to-stand			○
Upper body endurance fitness			
Arm curl			○
Modified push-ups		○	
Flexibility			
Sit and reach	○		
Chair sit and reach			○
Back scratch			○
<i>Motor fitness</i>			
Speed			
Running 20-m	○		
Agility			
4 x 10-m shuttle run	◐		
5 x 10-m shuttle run	○		
2.45-m time up & go			○
Balance			
Dynamic balance	○		
Single-leg stand	◐	○	
<i>Indicates strong evidence:</i>			
<i>Indicates moderate evidence:</i>			
<i>Indicates limited evidence:</i>			

Safety

Cardiorespiratory fitness

Strong evidence indicated that only the 2-km walk test was safe to assess cardiorespiratory fitness in adults. Moderate evidence indicated that 20-m shuttle run test was safe to assess cardiorespiratory fitness in children and adolescents. Limited evidence indicated that the 6-min walk test in children and adolescents, the Danish step test in adults, the 3-min step test in children and adolescents, the Chester step and modified incremental step tests in adults, and the 45-s squat test in children and adolescents were safe to assess cardiorespiratory fitness.

Musculoskeletal fitness

Moderate evidence indicated that a) the handgrip strength test was safety to assess maximal isometric strength in children and adolescents; b) the standing long jump test was safe to assess lower body explosive strength in children and adolescents; and c) the isometric back endurance test was safe to assess trunk endurance strength in adults. Limited evidence indicated that a) the handgrip strength test was safe to assess maximal isometric strength in adults; b) the vertical jump test was safe to assess lower body explosive strength in adults; c) the trunk flexor endurance test was safe to assess trunk endurance strength in adults; d) the modified push-ups test was safe to assess upper body endurance strength in adults.

Flexibility

No studies examined the safety of the flexibility field-based physical fitness tests; therefore, null evidence was found.

Motor fitness

Limited evidence indicated that a) the dynamic balance test was safe to assess balance and lower body strength in children and adolescents; and b) the single-leg stand test was safe to assess balance in children and adolescents and adults.

No evidence was found in safety studies for the older adult population.

Table 14. Levels of evidence of safety in field-based fitness tests by age group.

Field-based fitness tests	Children/Adolescents	Adults	Older adults
<i>Cardiorespiratory fitness</i>			
Distance and time-based walk			
2-km walk		●	
6-min walk	○		
Shuttle run			
20-m shuttle run	●		
Step and squat			
Danish step		○	
3-min step	○		
Chester step		○	
Modified incremental step		○	
45-s squat	○		
<i>Musculoskeletal fitness</i>			
Maximal isometric strength			
Handgrip strength	●	○	
Lower body explosive strength			
Standing long jump	●		
Vertical jump		○	
Trunk endurance strength			
Trunk flexor endurance		○	
Isometric back endurance		●	
Upper body endurance strength			
Modified push-ups		○	
<i>Motor fitness</i>			
Balance			
Dynamic balance	○		
Single-leg stand	○	○	
<i>Indicates strong evidence:</i>			
<i>Indicates moderate evidence:</i>			
<i>Indicates limited evidence:</i>			

Criteria to assess feasibility and safety

Homogeneous criteria were not observed when assessing the field-based physical fitness test feasibility and safety, what could influence in the final results.

There was no consensus on which items were appropriate to assess feasibility, and whether they should have been reported by the participants and/or evaluators (see **Table 15**). Three studies included items reported by participants⁵⁻⁷, all studies, except Lamonedá et al.,⁶, included items reported by evaluators, and two studies used a mixed evaluation^{5, 7}, predominating items based on the perception of the evaluators and/or participants and few objective items.

Among the items reported to assess feasibility by participants, the “self-perception effort”⁵⁻⁷ were the most used. “Self-perception fatigue”, “self-perception preference test” and “self-perception enjoyment” were the other items reported by the participants, which were used only once in a single study⁶. It is important to highlight that the “self-perception effort” and “self-perception fatigue” items are more related to the safety assessment. In fact, they have been used in most of the safety studies included in this systematic review, which corroborates that there is no unanimity of criteria. Likewise, Hébert et al.,⁸, include the item “adverse event” to assess feasibility, when it is an item used by most authors to assess safety.

The items most frequently recorded by evaluators were “time to perform the test” used in 16 studies^{5, 7, 9-22}, “participants performing the test” in 13 studies^{7-10, 12-15, 19-23}, “easy to administer” in seven studies^{5, 8, 17, 18, 20, 22, 23}, “number of evaluators” in five studies^{7, 9, 16-18}, and “understand test instructions” in four studies^{8, 18, 20, 23}.

It can be observed that these items reported by evaluators, were mostly focused on the time necessary for the preparation, development and implementation of the test, which varied depending on the appropriate facilities and equipment, test procedures, number of tests assessed, number of testers and number of participants that can be tested at the same time. Only two studies^{16, 18} provided information about the time for preparation, number of evaluators, number of participants assessed by test and time to perform the test/s. This information would be interesting when choosing a test according to its feasibility, and designing future field-based physical fitness test batteries.

Interestingly, only two studies have explicitly assessed whether sport clothes and facilities were appropriate or not^{18, 20}. This could be due to the other studies assumed that to administer a test to sportswear and facilities have to be appropriate.

Regarding the total number of feasibility items included in each study, only one study included nine items assessed by evaluators¹⁸, one study used six items²⁰, and five items⁷, three studies used four items^{6, 16, 17} and 13 studies (68.4%) only used between one and three items^{5, 8-15, 19, 21-23}.

Table 15. Items to establish feasibility reported by evaluators and participants[§].

Study	Participants performing the test (%)	Evaluators (n)	Participants measured (n)	Tests	Time to prepare test	Time to perform test	Easy to administer
McAllister & Palombaro, 2019	82%			Modified 30-s sit-to-stand		30 seconds	
Bruggeman et al., 2020				45-s squat and 3-min step		45 seconds / 3 minutes	Yes
Anderson & Dal Corso, 2016				Chester step, modified incremental step and 6-min walk		10 / 12 / 6 minutes	
Borel et al., 2010	100%			6-min step		6 minutes	
Lamonedá et al., 2020				20-m shuttle run music vs. 20-m shuttle run			
Aadahl et al., 2012	98.8%			Danish step		6 minutes	
Langhammer & Stanghelle, 2018	100%			2-min step and 6-min walk		2 / 2 minutes	
Oja et al., 1991	100%			2-km walk		15.2 /16.9 minutes (mean for males /females)	
Laukkanen et al., 1992	96.4%	2		2-km walk		15 to 18 minutes	
Suni et al., 1998	95%	3		2-km walk, handgrip, vertical jump,		1hr 10minutes for HRFI test-battery	

				single-leg stand, isometric back endurance and modified push-ups				
Amado-Pacheco et al., 2019	96%			20-m shuttle run		6.7/ 5.3 minutes (mean for males/females)		
Cadenas-Sánchez et al., 2014		2*	Groups of 4-8 (4-5 year-old) 4-5 (3-year-old)	20-m shuttle run	3 minutes	8 a 10 minutes		
Cadenas-Sánchez et al., 2016		2-5	Groups of 4-8 (same age)	20-m shuttle run; handgrip, standing long jump, 4×10-m shuttle run and single-leg stand		2.5hrs (20 participants /2 evaluators**) for PREFIT battery	Yes all, except standing long jump	
España-Romero et al., 2010		1	Group of 20	20-m shuttle run, handgrip and standing long jump	11minutes	2.5hrs for Evidence-based ALPHA-Fitness test battery	Yes	
Fjortoft et al., 2011	100%			Reduced Cooper, jumping a distance of 7-m on 1 feet/ 2 feet, pushing a medicine ball, running 20-m and 5×10-m shuttle run			Yes	

Hébert et al., 2011	100%	Handgrip				Yes
Smits-Engelsman et al., 2020	100%	Standing long jump, single-leg stand and dynamic balance			20-40 minutes per participants /30-50 minutes per 2 participants for PREFIT test battery	Yes
Boyer et al., 2013	88-99%	Partial curl-ups, 60-s, 90-s and unlimited plank test			Fitnessgram /CHMS partial curl-ups (29/36 s); 60-s /90-s /unlimited time planks (41/52/56s)	
Ito et al., 1996	100%	Trunk flexor endurance and isometric back endurance			95-89% within the time (≤5minutes)	Yes
Total studies by item	13	5	3	2	16	7

Table 15 (continued). Items to establish feasibility reported by evaluators and participants[§].

Study	Understand test instructions (% yes)	Appropriate sport clothes (% yes)	Appropriate facilities (% yes)	Reject (% yes)	Self- perception effort [§] (%)	Self- perception fatigue [§] (%)	Self- perception preference test [§] (%)	Self- perception enjoyment [§] (%)	Total items by study
McAllister & Palombaro, 2019									2
Bruggeman et al., 2020					100% low				3
Anderson & Dal Corso, 2016									1
Borel et al., 2010									2
Lamoneda et al., 2020					72.4%max. for both	56.4% less fatigue with music	89.6% vs. 10.4%	94% vs. 6%	4
Aadahl et al., 2012									2
Langhammer & Stanghelle, 2018									2
Oja et al., 1991									3
Laukkanen et al., 1992				21%	81% easy				5
Suni et al., 1998									3
Amado-Pacheco et al., 2019									2
Cadenas-Sánchez et al., 2014									4
Cadenas-Sánchez et al., 2016									4
España-Romero et al., 2010	≥ 95% [#]	≥ 95% [#]	≥ 95% [#]	2.3% [#]					9
Fjortoft et al., 2011	100%								3
Hébert et al., 2011	100%								3

Smits-Engelsman et al., 2020	100%	76%	100%							6
Boyer et al., 2013										2
Ito et al., 1996										3
Total studies by item	4	2	2	2	3	1	1	1		
\$ Questions answered by participants. #An acceptable level of feasibility was considered when the items were ‘positively’ answered in at least 95 % of the cases.										

Similarly, safety was assessed with items reported by evaluator and participants (see **Table 16**). Most of the studies used items reported by participants and evaluators^{5-7, 9, 11, 12, 15, 18}, except five studies that used only items reported by evaluators^{8, 13, 19, 20, 22}. Note that these five studies only used the “adverse events” item to assess safety.

We only found two objective items based on “HR” and “exertion 85% HRmax” which were included in seven studies^{5-7, 9, 11, 12, 15}. The “HR” item together with “adverse events”^{5, 8, 9, 11, 13, 15, 18-20, 22} and “self-perception general exertion” items^{5-7, 11, 15} were the most used to assess safety.

Most of the items were related to the perception of effort/fatigue or feeling sick (i.e., self-perception dyspnea and fatigue, general exertion, DOMS and sick feeling). Only one study¹⁸ included a questionnaire related to “instrument allergy”.

Heart rate assessed at the end of the test, was the most commonly used safety item in cardiorespiratory fitness tests (i.e., the 6-min walk, 2-km walk, 20-mt shuttle run, 45-s squat, 3-min step, Chester step, modified incremental step and 6-min step tests). Instead of using the absolute HR as a safety item, we suggest to use the relative value as a percentage of HRmax or similar to assess safety. Only one study⁹ employed this last item, expressing the exertion as the percentage of age-predicted HRmax, establishing 85% as cut-off point. On the other hand, “adverse events” was the most used item to evaluate cardiorespiratory fitness (i.e., 20-mt shuttle run, 45-s squat, 3-min step, Chester step, modified incremental step, Danish step, 2-km walk tests), musculoskeletal fitness tests (i.e., the handgrip strength, vertical jump, standing long jump, isometric back endurance, modified push-ups and trunk flexor endurance), and motor fitness tests (i.e., the dynamic balance and single-leg stand).

Of the 11 items proposed in the studies to assess safety, two studies used four items^{11, 18}, and 11 studies (85% of the studies) only utilized between one and three items^{5-8, 12, 13, 15, 18-20, 22}.

Table 16. Items to establish safety reported by evaluators and participants^{\$}.

Study	Test	Self-perception dyspnea ^{\$}	Self-perception leg fatigue ^{\$}	Self-perception general exertion ^{\$}	DOMS ^{\$}	Exertion 85% HRmax*	HR*	Adverse events	Instrument allergy	Sick feeling	Pain in hand or forearm	Injury	Total items by study
Borel et al., 2010	6-min step	2.5	3.4				118						3
Amado-Pacheco et al., 2019	20-m shuttle run							0 %					1
Bruggeman et al., 2020	45-s squat/ 3-min step			3.4/2.7			N/A	0%					3
Anderson & Dal Corso, 2016	Chester step/modified incremental step/ 6-min walk	2/2/2	2/2/2				143/142/128	0.12%					4
Aadahl et al., 2012	Danish step							0%					1
Suni et al., 1998	2-km walk/ handgrip/vertical jump/ single-leg stand/ isometric back endurance/ modified push-ups				♂60/0/0/	♂43/0/0/	♂153/95/110/						4
					0/0/0	0/1.3/9.2	00/112/140						
					♀78/0/0/	♀37/0/0/	♀151/92/112/	<5 %					
Oja et al., 1991	2-km walk												3
Laukkanen et al., 1992	2-km walk												2
España-Romero et al., 2010	20-m shuttle run/ handgrip /standing long jump				71				≥99%#	≥99%#	≥99%#	≥99%#	5
Lamoneda et al., 2020	20-m shuttle run music vs. 20-m shuttle run			8.3 vs. 7.9			177 vs.178						2
Hébert et al., 2011	Handgrip							0%					1

Smits-Engelsman et al., 2020	Standing long jump/ single-leg stand/ dynamic balance							0%				1
Ito et al., 1996	Trunk flexor endurance/isometric back endurance							0%				1
Total studies by item		2	2	4	2	1	7	9	1	1	1	1

DOMS: Delayed-onset soreness; HR: Heart rate. A/N: Not available data. \$ Questions answered by participants after the test thought Borg scale (1-10 points). \$\$Item answered by participants 48 hours post test thought questionnaire. *After the test. #An acceptable level of safety was considered when the items were ‘positively’ answered in at least 99 % of the cases.

There are no own analyses to assess the feasibility and safety (see **Tables 15 and 16**). All the studies used the frequency in number, mean or percentages to decide whether the items evaluated meet or not the feasibility and safety criteria. However, a value or cut-off point is not established to determine whether or not an item meets the feasibility and safety criteria. Only España-Romero et al.,¹⁸ considered that a field-based physical fitness test met the feasibility and safety criteria when the items were positively answered in at least 95% and 99%, respectively. Thus, there was no consensus on whether or not an item met the criteria for feasibility or safety. All these differences when assessing, analysing and concluding with the findings, could lead to biases when presenting the results and objectively establishing that a study is feasible and/or safe to measure physical fitness in people of all ages, especially if there are no valid tools designed to assess them.

Finally, based on our findings, there are items that should be considered as a high priority to assess the feasibility and safety of field-based physical fitness test in children, adolescents, adults or/and older adults. We consider that to assess feasibility we should use at least the following items: a) “time to perform the test”; b) “time to prepare the test”; c) “percentage of rejection”; d) “percentage of participants performed the test” e) “number of participants measured by test”; f) “number of evaluators”; g) “easy to administer”; and h) “percentage of participants understand test instructions”. Appropriate facilities, equipment and sportswear must be assumed as a condition, whenever we assess field-based physical fitness test. Regarding the assessment of safety, we highlight a) “exertion %HRmax”; b) “percentage of adverse events” (i.e., falls, sick feeling, pain and/or injury); c) “general exertion and/or fatigue”; d) “DOMS”; and e) “instrument allergy”, in case we use instruments that were in direct contact with the participant's skin (see **Figure 7**).

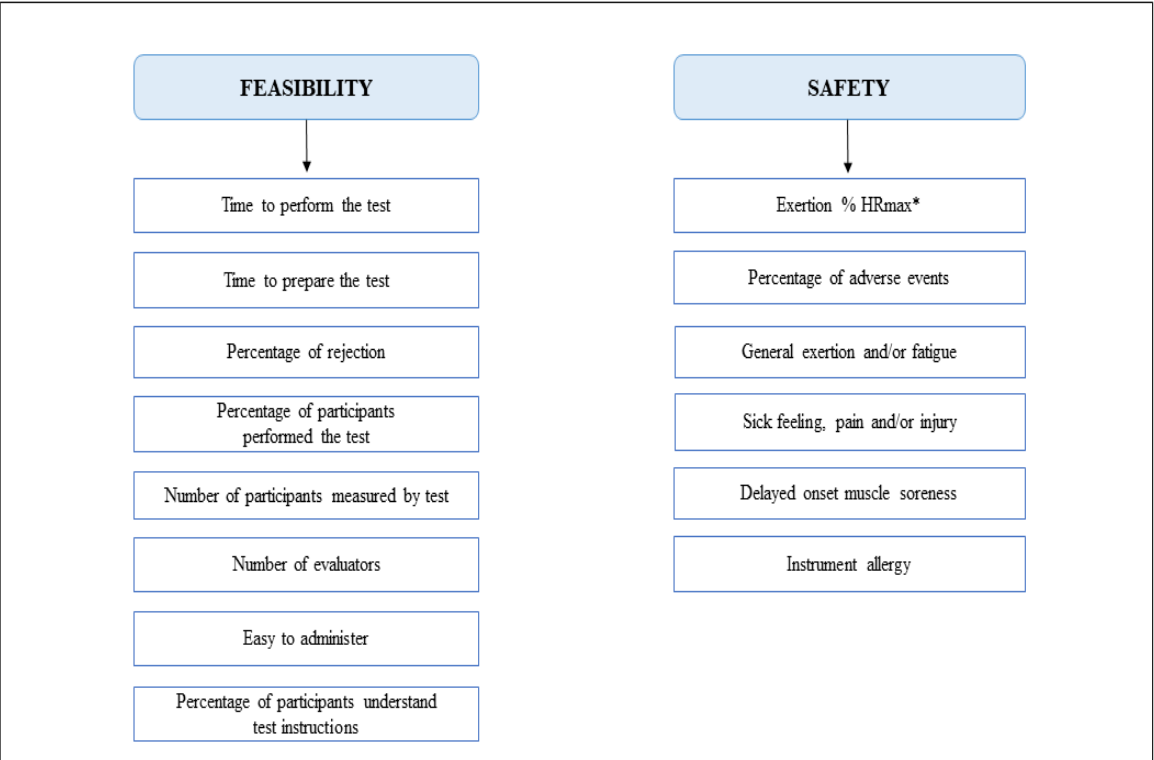


Figure 7. Proposed items for assessing the feasibility and the safety of field-based physical fitness tests in people of all ages.

**The exertion % HR max. will depend on the maximum and submaximal characteristics of the test.*

Study II. Feasibility and Safety in health-related field-based physical fitness test in adult population: The ADULT-FIT project.

The main findings of the present study suggest that the most frequently used health-related field-based physical fitness tests are feasible and safe in adult population regardless sex, age, and physical activity level.

Feasibility

All participants completed the proposed field-based physical fitness tests without any difficulties. All the proposed items presented a successful answer of 100% by the participants. The items “appropriate facilities”, “appropriate equipment” and “appropriate sports clothes” should be implicitly included in the assessment of health-related field-based physical fitness tests; however, we have included them based on a previous study in children and adolescents¹⁸, which results are in line with our findings.

Regarding to the items “test’s instruction understood” and “easy to administer”, several studies presented a similar successful answer in these items in children and adolescents^{8, 16, 20, 23, 24} and adults⁹. However, one study in children and adolescents¹⁸, where the assessments were conducted by Physical Education teachers, observed difficulties in the evaluation of the body composition, specifically with the skinfolds measurements due to unfamiliarity. On the contrary, in our study, the evaluators had previous experience. In another study in preschoolers¹⁷, difficulties in the standing long jump were observed, especially in the youngest (3-year-old group). This might have occurred because children do not have a developed jumping pattern, and in our case, we approached the standing long jump as a natural movement and not as a specific skill.

Regarding the “participants evaluated/not evaluated” item, in our study all participants performed the tests. Other studies also had a high performance rate, between 95-100% in children and adolescents^{8, 13, 20, 23}, adults^{7, 9, 15}, and older participants¹⁴.

Concerning to the “ratio participants/evaluators” item, in the cardiorespiratory fitness tests was 10 participants by 2 evaluators. Previous studies in preschoolers showed an optimal ratio of 4 to 8 participants by 2 evaluators (1 of them running with the children to set the pace) to assess the 20-m shuttle run test^{16, 17}; while in children and adolescents an optimal ratio was 20 participants by 1 evaluator¹⁸. The “ratio participants/evaluators” most used in muscular and motor fitness tests were 1/1 and 1/2, respectively. Other

studies also showed 1-2 evaluator to assess the handgrip strength^{17, 18}, the standing long jump^{17, 18} and the 4x10-m shuttle run tests¹⁷ in children and adolescents. Only one study⁹ mentioned that 1-3 evaluators were needed for the assessment of the muscular and motor fitness tests in adults. We believe that it is important to include 2 evaluators in the muscular fitness and motor fitness tests in order to have a correct measurement (i.e., registration of data, better control of correct execution those tests of technical complexity and more safety), especially when assessing adult population with low physical fitness.

The “time of preparation” was approximately 1 minute for the majority of the tests except, for the handgrip strength test, which showed the shortest time, which corresponded mainly to the time needed to measure the hand size in women. The 20-m shuttle run test took ~2 minutes of preparation, in concordance with Cadenas-Sánchez et al.,¹⁶ which took ~3 minutes in preschoolers, although España-Romero et al.,¹⁸ took ~5 minutes in children and adolescents. We could consider that all the tests were easy to prepare and the time that takes to prepare them is adequate.

For the “time of performance” item, we took between 3.5 to 5.5 minutes per participant in body composition assessments (i.e., weight, height, neck, waist, and hip circumferences and triceps and subscapular skinfolds); similar to the study by España-Romero et al.,¹⁸ in which 5 minutes were required per child or adolescent, although, less anthropometric measurements were assessed (i.e., height, weight, waist circumference, triceps and subscapular skinfolds).

For the 2-km walk test the “time of performance” ranged between 11 to 23 minutes with a mean of 15.6 minutes for males and 16.7 minutes for females; similar results were found in previous studies^{7, 15} in adults. In the 20-m shuttle run test, a range of 1.5 to 13 minutes with a mean of 6.3 minutes for males and 5.7 minutes for females. No studies have been found that record the “time of performance” in the adult population, but a previous study in preschoolers showed a range of 5.3 to 10 minutes^{13, 16, 17} and a mean of 5.1 for children and adolescents¹⁸. It is important to note that the time spent in this study will depend on the cardiorespiratory fitness level of the participants, where sex and age are also determinants.

In general, the muscular fitness tests did not exceed 2 minutes in “time of performance”, including those with two attempts and 1-minute rest between attempts. Similar results have been found in two studies^{17, 18}, in which the “time of performance” for the handgrip strength and standing long jump (including 2 or 3 jumps and 1-minute

rest between attempts) tests were ~2 minutes in children and adolescents.

Regarding to the motor fitness tests, the 6-m gait speed and 2.45-m time up & go tests took less than 7 seconds (one attempt per test), the 4x10-m shuttle run test (including two attempts and 1-minute rest between attempts) took 1.2 to 1.3 minutes, and the single-leg stand test (30 seconds per leg and 1-minute rest between eyes open and eyes closed modality) took 3.0 to 3.6 minutes. Two studies that analysed the “time of performance” of the 4x10-m shuttle run test in preschoolers found a range from 1.3 to 1.4 minutes¹⁷, and from 1.3 to 1.4¹³ (including two attempts and 1-minute rest between attempts). No other studies were found that examined this test in adults. Concerning the single-leg stand test, only one study⁹ measured the “time of performance” in adults, using a 5-point scale. It was concluded that this test showed an excellent “time of performance” (i.e., range from 4.3 to 4.7).

Safety

Referred to our safety assessments, >99% of the items presented a successful answer. No participant reported any “instrument allergy”, which is in line with a previous study in children and adolescents¹⁸.

Regarding to “adverse events” item, there were no falls, and no participants suffered any injuries during the assessments; however, only one participant (0.3%) felt sick at the end of the 20-m shuttle run test, who recovered without complications after resting and hydrating. Only one study¹¹ in adults analysed “adverse events” item in the 6-min walk test and the results were similar to our study. Previous studies in adults observed a 0.5%⁹ and <30%¹⁵ of adverse events assessing the 2-km walk test; both expressed that the discomfort was focused on the lower limb musculature. Also, one study¹³ reported no “adverse events”, and other study¹⁸ showed than one participant reported muscle cramps during the assessment of the 20-m shuttle run test in children and adolescents.

Regarding muscular fitness tests, 4 studies have included this item in the assessment of handgrip strength test. Three showed no adverse events in children^{8, 13} and adults⁹, and only one study¹⁸ reported that one participant had “pain in the hand or forearm” during the assessment in children. Moreover, no adverse events were observed in the assessment of the standing long jump test in children^{13, 20}.

Finally, one study⁹ in adults reported that three participants felt dizzy during the modified version of single-leg stand test (eyes-opened, 1 minute by legs) and one study¹³ reported no adverse events in the 4x10-m shuttle run test in children. As we mentioned above, no “adverse events” occurred in our study, only the case of one participant who felt sick at the end of the 20-m shuttle run test, but recovered without complications after resting and hydrating.

In general, it may be noted that the incidence of adverse events in the field-based physical fitness tests is low and mostly related to strain of the predominant musculature involved in the performance of the test, which is something normally observed in physical exercise. This allows us to conclude that the application of these field-based physical fitness tests is safe.

Regarding to “HR” item, the most relevant aspect of monitoring it is the opportunity to assess the intensity of exercise, which is closely related to VO₂ in submaximal (6-min walk and 2-km walk) and maximal (20-m shuttle run) field-based fitness tests, which allows to know accurately the intensity of these²⁵. The “85%HRmax” is considered a rough estimate of the anaerobic threshold²⁶. We observed that more participants reached “85% HRmax” as the test duration (i.e., the 6-min walk vs. the 2-km walk tests) and test intensity (i.e., 6-min walk and 2-km walk tests vs. the 20-m shuttle run test) increased.

Similarly, with increasing age, more participants reached the “85%HRmax”, however, the level of physical activity of the participants did not influence reaching this threshold. This item is also used by Suni et al.,⁹ to calculate cardiovascular exertion in cardiorespiratory fitness test (i.e., the 2-km walk) as well as in muscular and motor fitness tests in adults (i.e., 37-57 years). Regarding the 2-km walk test, they reported that 43% of males and 37% of females reached or exceed their “85% HRmax”. In our study, the 32% reached their “85%HRmax”, being all the participants from the older groups (i.e., 35-49 and 50-64 years). This could mean that younger participants from our study (i.e., 18-34 years) would have better adaptations to the cardiovascular effort of this submaximal test. Only one study⁷ reported the “HR” at the end of the 20-m shuttle run test in adolescents (177.7 (29.2) bpm), however, the “85% HRmax” of the participants was not reported. “HR” assessment has shown to be a useful measure of the cardiovascular stress level of participants²⁵. For safety reasons, “HR” monitoring could be included in endurance strength tests like 30-s sit-to-stand and front plank tests.

Concerning to the item “RPE”, the results on perceived exertion in the muscular fitness tests (i.e., 30-s sit-to-stand and front plank tests) showed that the most demanding test for the participants was the front plank test, where more than 60% perceived their general exertion as “hard”. No other studies that registered “RPE” item in muscular fitness tests were found. We did not include “RPE” registering in the cardiorespiratory and motor fitness tests, however, others studies in adults analysed this item in cardiorespiratory fitness tests and showed “moderate” effort in the 2-km walk test^{7, 15} and “very hard” effort in the 20-m shuttle run test⁶. We believe that the “RPE” can be also assessed in cardiorespiratory fitness tests, but it is essential to include an objective variable such as the “HR”.

In relation to “tibial pain” item, it was only registered after finishing the 2-km walk test, where the 32% of the total sample felt some degree of pain but any participant interrupted the test because of pain. There are no other safety studies that analyse this item. It is relevant to include the assessment of “tibial pain” as a method for quantifying muscle pain at the end of field-based walking tests (i.e., the 6-min walk and 2-km walk tests) in order to analyse the physiological response of the musculoskeletal system to the intensity and duration of walking tests. During the walking a higher pressure is generated on the tibialis anterior musculature, which is responsible for controlling the descent of the foot to floor contact and is involved in the dorsiflexion of the foot in the swing phase, so the higher intensity of walking may lead to the appearance of muscle pain^{27, 28}. This assessment could be a good complement to the “DOMS” perceived by the participants.

We consider that it is relevant to include the “HR” and “RPE” items in cardiorespiratory and endurance strength fitness tests (i.e., 30-s sit-to-stand and front plank tests); and “tibial pain” in walking/running cardiorespiratory fitness tests, as they have been less analysed in previous studies, but are highly related to the control of the muscular and cardiovascular exertion of the participants and helps to quantify the exercise intensities during the performance of the field-based physical fitness tests^{25, 29}. All this would provide further advances in the study of safety in health-related field-based physical fitness tests.

Referred to the “DOMS” item, in our study the 43.6% of the total sample (44.1% active) showed some degree of “DOMS”, and only the 2.9% reported a “severe” degree. From this sample, the 89.4% felt pain in lower extremities and additionally, the 95.3% of participants reported that cardiorespiratory fitness tests could be the cause of their

“DOMS”. Only one study⁹ assessed “DOMS” in adults indicating that 60% of males and 78% of females experienced some degree of “DOMS”, but only the 5% of males and 10% of females reported “severe” degree. Although this study also included the 2-km walk as a cardiorespiratory fitness test, 83% of participants reported that their “DOMS” were caused by the one-leg squat test. Another study¹⁸ reported that the 35.3% of children and adolescents experienced some degree of “DOMS”, where the 21.3% reported “severe” degree and the 83% reported that the “DOMS” was due to performing the 20-m shuttle run test, which is in line with our study. It is important to include “DOMS” as a tool to detect the degree of muscular overload (intensity and/or duration) resulting from the assessment of the field-based physical fitness tests and also, for identify the muscular fitness of individuals in order to establish more personalized training loads related to gender, age and level of physical activity³⁰. This tool, in addition to those mentioned above (HR, RPE and tibial pain), helps us to understand more precisely the physical demands of the most frequently used health-related field-based physical fitness tests in the adult population.

Study III. Responsiveness of health-related field-based physical fitness tests in adults: The Adult-Fit project.

The aim of the present study was to determine the responsiveness of the selected health-related field-based physical fitness tests in adults. Overall, all field-based tests were found to be responsive after the intervention regardless sex and age with an effects size moderate to large, except to the weight, BMI and handgrip tests that was trivial. The highest proportion of individual responders was observed in the 4x10-m shuttle run, 30-s sit-to-stand, 6-m gait speed and 2.45-m time up & go tests. The proportion of the population that is expected to respond to the intervention was higher 85% in most of the health-related field-based fitness tests evaluated, but $\leq 33\%$ in weight, BMI, waist circumference and handgrip strength tests. To the best of the authors' knowledge, no previous studies have addressed the responsiveness on the health-related field-based physical fitness tests used in healthy adult population after an exercise intervention. Most of them have been focused on physical performance in clinical practice, in populations with pathologies and older adults with or without disease.

Responsiveness is considered the more essential propriety of an evaluative instruments, that is, the ability of a test to detect changes over time³¹. Therefore, apart from evaluating validity^{2, 32-35}, reliability⁴, feasibility and safety³⁶, the responsiveness needs also to be examined.

Our results displayed modest magnitude of change in whole-body weight (-0.94 kg), BMI (-0.50 kg/m²) and waist circumference (-0.84 cm) after exercise training in adults with a mean age of 45 years and BMI of 26.9 kg/m². The effects of moderate-intensity continuous training (MICT) combined or not with resistance training, HIIT or, MICT vs. HIIT on body composition has been analysed in recent years. It has been widely proven that exercise has a positive effect on body-weight, fat mass or waist circumference in adults, but small effects of change³⁷⁻⁴¹. These studies conclude that the effect of exercise intervention might be influenced by age (18-65 years), complications (obesity), duration (>6 weeks), frequency, intensity, time-efficient and exercise modality. In addition, diet control is an important variable to the control of the body weight since a combination of diet with exercise being more effective than just exercise in reduction of body mass and fat mass⁴².

Regarding to field-based cardiorespiratory fitness tests, we observed moderate

and large magnitude of change after exercise training. At the individual level, ~50-55% of the participants were responders. The proportion of the population that is expected to respond to the intervention was $\geq 86\%$ in all tests. Responsiveness of 20-m shuttle run test has not been defined in healthy adult populations. A study developed in asthmatic adolescents showed an increase in the number of stages completed (7 ± 1.4 to 10.1 ± 1.3) in the 20-m shuttle run test after an individualized aerobic training program, and concluded that this test is sensible to change over the time⁴³. We observed a moderate magnitude of change after exercise training, with an improvement of 1.3 states in healthy no-active adults. A study conducted in healthy nonathletic middle-aged adults observed an decrease in the time (minutes) spent to complete the 2-km walk test (13.1 ± 1.0 to 15.1 ± 0.6 in males; 15.1 ± 1.0 to 16.3 ± 1.1 in females) after 15-week walking training⁴⁴. The authors concluded that the 2-km walk test can be used as a reasonably accurate field-based test to predict changes in the cardiorespiratory capacity due to aerobic training. These results are in line with our study in healthy no-active middle-aged adults, which improved 82 s in the 2-km walk test after exercise training program. Respect to the 6-min walk test, it is commonly employed to evaluate responsiveness to rehabilitation or exercise program in population with chronic respiratory diseases, cardiovascular disease and older adults⁴⁵⁻⁴⁷. Available evidence suggests a minimal important difference of 30 meters in 6-min walk test in adults with chronic respiratory disease. Responsiveness data of 6-min walk test in healthy adults are limited. A study in adults in military service considered an improvement in more than 81.3 meters (via minimal detectable change (MDC)) in 6-min walk test as a true change in performance⁴⁸; while our analyses showed an average increase of 62.5 meters after the intervention. Despite limited evidence in responsiveness of cardiorespiratory fitness tests in health adults, 6-min walk, the 2-km walk and the 20-m shuttle run tests, are responders to an exercise training program in our sample. The verbal encouragement and patient motivation are very important when evaluate a running or walking cardiorespiratory fitness tests influencing in the person's effort capacity, especially in longer distance and self-regulated pace tests.

In field-based motor fitness tests, we observed a large magnitude of change in 6-m gate speed (-0.3) and 2.45-m time up & go tests (-0.7 s), and moderate to the 4x10-m shuttle run test (-1.2 s) after exercise training; and ~66-78% of the participants were responders. The proportion of the population that is expected to respond to the intervention was $\geq 89\%$ in all tests. Responsiveness of these tests has not been defined in

healthy adult populations. A reduction of 1.0 s for the 2.45-m time up & go test has been determined to be the MDC in older adults with type 2 diabetes⁴⁹, 0.5 s in people with epilepsy⁵⁰. A study in older adults showed that the 4-m gait speed test is responsive with 0.05 m/s and 0.1 m/s indicating a substantial change (via small meaningful change)⁵¹. No responsiveness data found in the 4x10-m shuttle run test. The 6-m gait speed and 2.45-m time up & go tests are frequently used to evaluate motor fitness (i.e., speed of reaction, speed of movement or gait speed, agility, coordination) in older adults, while the 4x10-m shuttle run test is frequently used in the young population, making it difficult to find responsive data in adult population.

Regarding to responsiveness of muscular fitness tests, we observed significant but very small effects ($P < 0.001$; *Cohen's d* = 0.13) on handgrip strength test after exercise training, with a change of 1.3 kg. This is in line with previous studies in older adults, where change above 1.6 kg was considered a real change⁵². A recent metanalytical review examined the effects of different exercise training on handgrip strength in healthy community-dwelling older adults of 60 years or older. The effects change was small with a large effect in favor of program based on task-specific training and multimodal training modes to provide an appropriate stimulus to improve handgrip strength⁵³; which is an interesting contribution to take to be account when evaluating the efficacy of interventions with exercise training. Our intervention has incorporated different training focuses (cardiorespiratory, muscular and motor fitness (coordination, balance, agility and speed)), where grip strength was training thought grips and tractions with TRX (suspension training), elastic bands, dumbbell and kettlebell. However, initial level of handgrip strength of participants (where better levels could show smaller changes), intervention time (>3 moths), intensity or task-specific training should be also taken into consideration to achieve a greater change.

The standing long jump responded after exercise training, with a moderate magnitude of change of 17.3 cm; the 46% of the participants were responders and the proportion of the population that is expected to respond to the intervention was $\geq 96\%$. Responsiveness of standing long jump test has not been defined in healthy adult populations. The 30-s sit-to-stand stand and front plank tests showed a large magnitude of change after exercise training, the 69% and 54% of the participants were responders, respectively. The proportion of the population that is expected to respond to the intervention was $\geq 91\%$ in both tests. A MDC value of 3.24 in adults with low back pain

has been defined to the 30-s chair stand test; that is, an improvement in more than 3.24 repetitions has been considered a true change in performance⁵⁴. Other studies have observed similar values in older adults with type 2 diabetes or dementia^{49, 55}. Which is in line with our results, participants increase an average of 3.6 repetition more after the exercise training intervention. Our analyses showed an average increase of ~44 s in the front plank test after exercise training in healthy no-active adults. In a study developed in younger and older adults (20-35 and 60-79 years old, respectively), an improvement of 57 s in front plank test was considered a true change in performance (via MDC)⁵⁶; this improvements in the front plank test is relatively higher that the change observed in our study, authors affirm that the sample was probably more fit and active than the general population. Thus, more studies about responsiveness of standing long jump, 30-s sit-to-stand and front plank tests in healthy adults are required.

REFERENCES

1. Castro-Pinero J, Artero EG, Espana-Romero V, Ortega FB, Sjostrom M, Suni J, Ruiz JR. Criterion-related validity of field-based fitness tests in youth: a systematic review. *British Journal of Sports Medicine*. 2009;44(13):934-43.
2. Castro-Piñero J, Marin-Jimenez, N., Fernandez-Santos, J. R., Martin-Acosta, F., Segura-Jimenez, V., Izquierdo-Gomez, R., ..., Cuenca-García M. Criterion-Related Validity of Field-Based Fitness Tests in Adults: A Systematic Review. *Journal of clinical medicine*. 2021;10:3743.
3. Artero EG, Espana-Romero, V., Castro-Pinero, J., Ortega, F. B., Suni, J., Castillo-Garzon, M. J., Ruiz JR. Reliability of field-based fitness tests in youth. *International journal of sports medicine*. 2011;32:159-69.
4. Cuenca-García M, Marin-Jimenez N, Perez-Bey A, Sánchez-Oliva D, Camiletti-Moiron D, Alvarez-Gallardo IC, Castro-Piñero J. Reliability of Field-Based Fitness Tests in Adults: A Systematic Review. *Sports Medicine*. 2022:1-20.
5. Bruggeman BS, Vincent HK, Chi X, Filipp SL, Mercado R, Modave F, et al. Simple tests of cardiorespiratory fitness in a pediatric population. *Plos one*. 2020;15(9):e0238863.
6. Lamonedá J, Huertas-Delgado, F. J., Cadenas-Sanchez C. Feasibility and concurrent validity of a cardiorespiratory fitness test based on the adaptation of the original 20 m shuttle run: The 20 m shuttle run with music. *Journal of Sports Sciences*. 2021;39:57-63.
7. Laukkanen RM, Oja P, Ojala KH, Pasanen ME, Vuori IM. Feasibility of a 2-km walking test for fitness assessment in a population study. . *Scandinavian journal of social medicine*. 1992;20:119-26.
8. Hébert LJ, Maltais, D. B., Lepage, C., Saulnier, J., Crête, M., Perron M. Isometric muscle strength in youth assessed by hand-held dynamometry: A feasibility, reliability, and validity study: A feasibility, reliability, and validity study. *Pediatric Physical Therapy*. 2011;23:289-99.
9. Suni JH, Miilunpalo SI, Asikainen TM, Laukkanen RT, Oja P, Pasanen ME, Vuori IM. Safety and feasibility of a health-related fitness test battery for adults. *Physical Therapy*. 1998;78:134-48.
10. McAllister LS, Palombaro KM. Modified 30-second sit-to-stand test: reliability and validity in older adults unable to complete traditional sit-to-stand testing. *Journal of Geriatric Physical Therapy*. 2020;43:153-8.
11. José A, Dal Corso S. Step tests are safe for assessing functional capacity in patients hospitalized with acute lung diseases. *Journal of Cardiopulmonary Rehabilitation and Prevention*. 2016;36:56-61.
12. Borel B, Fabre, C., Saison, S., Bart, F., Grosbois JM. An original field evaluation test for chronic obstructive pulmonary disease population: the six-minute stepper test. *Clinical rehabilitation*. 2010;24:82-93.
13. Amado-Pacheco JC, Prieto-Benavides, D. H., Correa-Bautista, J. E., García-Hermoso, A., Agostinis-Sobrinho, C., María Alonso-Martínez, A., ..., Ramírez-Vélez R. Feasibility and reliability of physical fitness tests among colombian preschool children. *International journal of environmental research and public health*. 2019;16:3069.
14. Langhammer B, Stanghelle JK. Senior fitness test; a useful tool to measure physical fitness in persons with acquired brain injury. *Brain injury*. 2019;33:183-8.
15. Oja P, Laukkanen R, Pasanen M, Tyry T, Vuori I. A 2-km walking test for assessing the cardiorespiratory fitness of healthy adults. 1991;12:356-62.
16. Cadenas-Sanchez C, Alcántara-Moral F, Sanchez-Delgado G, Mora-Gonzalez J, Martinez-Tellez B, Herrador-Colmenero M, Ortega FB. Assessment of cardiorespiratory fitness in preschool children: adaptation of the 20 metres shuttle run test. . *Nutricion hospitalaria*. 2014;30:1333-43. .
17. Cadenas-Sanchez C, Martinez-Tellez B, Sanchez-Delgado G, Mora-Gonzalez J, Castro-Pinero J, Lof M, et al. Assessing physical fitness in preschool children: Feasibility, reliability and practical recommendations for the PREFIT battery. *J Sci Med Sport*. 2016;19(11):910-5.

18. Espana-Romero V, Artero EG, Jimenez-Pavon D, Cuenca-Garcia M, Ortega FB, Castro-Pinero J, et al. Assessing health-related fitness tests in the school setting: reliability, feasibility and safety; the ALPHA Study. *Int J Sports Med.* 2010;31(7):490-7.
19. Aadahl M, Zacho, M., Linneberg, A., Thuesen, B. H., Jørgensen T. Comparison of the Danish step test and the watt-max test for estimation of maximal oxygen uptake: the Health2008 study. *European journal of preventive cardiology.* 2013;20:1088-94.
20. Smits-Engelsman B, Bonney, E., Neto, J. L. C., Jelsma DL. Feasibility and content validity of the PERF-FIT test battery to assess movement skills, agility and power among children in low-resource settings. *BMC Public Health.* 2020;20:1-11.
21. Boyer C, Tremblay M, Saunders T, McFarlane A, Borghese M, Lloyd M, Longmuir P. Feasibility, Validity, and Reliability of the Plank Isometric Hold as a Field-Based Assessment of Torso Muscular Endurance for Children 8-12 Years of Age. *Pediatric Exercise Science.* 2013;25(3):407-22.
22. Ito T, Shirado, O., Suzuki, H., Takahashi, M., Kaneda, K., Strax TE. Lumbar trunk muscle endurance testing: an inexpensive alternative to a machine for evaluation. *Archives of physical medicine and rehabilitation.* 1996;77:75-9.
23. Fjørtoft IPA VSH, Vereijken B. Measuring physical fitness in children who are 5 to 12 years old with a test battery that is functional and easy to administer. *Physical therapy.* 2011;91:1087-95.
24. Cadenas-Sanchez C, Sanchez-Delgado G, Martinez-Tellez B, Mora-Gonzalez J, Löf M, Espana-Romero V, et al. Reliability and validity of different models of TKK hand dynamometers. *The American Journal of Occupational Therapy.* 2016;70(4):7004300010p1-p9.
25. Achten J, Jeukendrup AE. Heart rate monitoring: applications and limitations. *Sports Med.* 2003;33:517-38.
26. Liguori G, American College of Sports M. ACSM's guidelines for exercise testing and prescription . . USA2021.
27. Sanzen LFA, Westlin N. Anterior tibial compartment pressure during race walking. . *The American journal of sports medicine.* 1986;14:136-8.
28. Chleboun GS, Busic AB, Graham KK, Stuckey HA. Fascicle length change of the human tibialis anterior and vastus lateralis during walking. . *Journal of orthopaedic & sports physical therapy.* 2007;37:372-9.
29. Chen MJ, Fan, X., Moe ST. Criterion-related validity of the Borg ratings of perceived exertion scale in healthy individuals: a meta-analysis. *Journal of sports sciences , .* 2002;20: 873-99.
30. Cheung KHPA, Maxwell L. Delayed onset muscle soreness. . *Sports medicine.* 2003;33:145-64.
31. Hopkins WG. A Spreadsheet for Monitoring an Individual's Changes and Trend. *Sportscience.* 2017;21:5-9.
32. Garcia-Hermoso A, Cavero-Redondo I, Ramirez-Velez R, Ruiz JR, Ortega FB, Lee DC, Martinez-Vizcaino V. Muscular Strength as a Predictor of All-Cause Mortality in an Apparently Healthy Population: A Systematic Review and Meta-Analysis of Data From Approximately 2 Million Men and Women. *Arch Phys Med Rehabil.* 2018;99(10):2100-13 e5.
33. Harber M, Kaminsky L, Arena R, Blair S, Franklin B, Myers J, al. e. Impact of cardiorespiratory fitness on all-cause and disease-specific mortality: advances since 2009. . *Progress in cardiovascular diseases.* 2017;60:11-20.
34. Marin-Jimenez N, Cruz-Leon C, Sanchez-Oliva D, Jimenez-Iglesias J, Caraballo I, Padilla-Moledo C, Castro-Piñero J. Criterion-Related Validity of Field-Based Methods and Equations for Body Composition Estimation in Adults: A Systematic Review. . *Current Obesity Reports.* 2022:1-14.
35. Marín-Jiménez N, Cruz-León C, Perez-Bey A, Conde-Caveda J, Grao-Cruces A, Aparicio VA, ..., Cuenca-García M. Predictive Validity of Motor Fitness and Flexibility Tests in Adults and Older Adults: A Systematic Review. *Journal of Clinical Medicine.* 2022;11:328.

36. Cruz-León C, Marín-Jiménez N, Sánchez-Parente S, Borges Cosic M, Grao-Cruces A, Castro-Piñero J, Cuenca-García M. Feasibility and Safety in health-related field-based physical fitness tests in adult population: The ADULT-FIT project. *Journal of Sports Sciences*. 2024;Under Rev.
37. Guo L, Chen J, Yuan W. The effect of HIIT on body composition, cardiovascular fitness, psychological well-being, and executive function of overweight/obese female young adults. *Frontiers in Psychology*. 2023;13:1095328.
38. Keating SE, Johnson NA, Mielke GI, Coombes JS. A systematic review and meta-analysis of interval training versus moderate-intensity continuous training on body adiposity. *Obesity reviews*. 2017;18:943-64.
39. Wewege M, Van Den Berg R, Ward RE, Keech A. The effects of high-intensity interval training vs. moderate-intensity continuous training on body composition in overweight and obese adults: a systematic review and meta-analysis. *Obesity Reviews* 2017;18:635-46.
40. Wilson JM, Marin PJ, Rhea MR, Wilson SM, Loenneke JP, Anderson JC. Concurrent training: a meta-analysis examining interference of aerobic and resistance exercises. *The Journal of Strength & Conditioning Research*. 2012;26:2293-307.
41. Sultana RN, Sabag A, Keating SE, Johnson NA. The effect of low-volume high-intensity interval training on body composition and cardiorespiratory fitness: a systematic review and meta-analysis. *Sports Medicine*. 2019;49:1687-721.
42. Clark JE. Diet, exercise or diet with exercise: comparing the effectiveness of treatment options for weight-loss and changes in fitness for adults (18–65 years old) who are overfat, or obese; systematic review and meta-analysis. *Journal of Diabetes & Metabolic Disorders*. 2015;14:1-28.
43. Ahmaidi SB, Varray AL, Savy-Pacaux AM, Pnetaut CG. Cardiorespiratory fitness evaluation by the shuttle test in asthmatic subjects during aerobic training. *Chest*. 1993;103:1135-41.
44. Laukkanen RMT, Kukkonen-Harjula TK, Oja P, Pasanen ME, Vuori IM. Prediction of change in maximal aerobic power by the 2-km walk test after walking training in middle-aged adults. *International journal of sports medicine*. 2000;21:113-6.
45. Holland AE, Spruit MA, Troosters T, Puhan MA, Pepin V, Saey D, ... & Singh SJ. An official European Respiratory Society/American Thoracic Society technical standard: field walking tests in chronic respiratory disease. *European Respiratory Journal*. 2014;44:1428-46.
46. O'Keefe S, Lye M, Donnellan C, Carmichael D. Reproducibility and responsiveness of quality of life assessment and six minute walk test in elderly heart failure patients. *Heart*. 1998;80:337.
47. Bellet RN, Adams L, Morris NR. The 6-minute walk test in outpatient cardiac rehabilitation: validity, reliability and responsiveness: a systematic review. *Physiotherapy*. 2012;98:277-86.
48. Wilken JM, Darter BJ, Goffar SL, Ellwein JC, Snell RM, Tomalis EA, Shaffer SW. Physical performance assessment in military service members. *JAAOS-Journal of the American Academy of Orthopaedic Surgeons*. 2012;20:S42-S7.
49. Alfonso-Rosa RM, del Pozo-Cruz B, del Pozo-Cruz J, Sanudo B, & Rogers ME. Test-retest reliability and minimal detectable change scores for fitness assessment in older adults with type 2 diabetes. *Rehabilitation Nursing Journal*. 2014;39:260-8.
50. Aktar B, Balci B, Oztura I, & Baklan B. The test-retest reliability and minimal detectable change of the six-minute walk test, timed up and go test, and 30-second chair stand test in people with epilepsy. *Physiotherapy theory and practice*. 2023;1-10.
51. Perera S, Mody SH, Woodman RC, & Studenski SA. Meaningful change and responsiveness in common physical performance measures in older adults. *Journal of the American Geriatrics Society*. 2006;54:743-9.
52. Bohannon RW. Test-Retest Reliability of Measurements of Hand-Grip Strength Obtained by Dynamometry from Older Adults: A Systematic Review of Research in the PubMed Database. *The Journal of frailty & aging*. 2017;6:83-7.

53. Labott B, H. B, Morat M, Morat T, Donath L. Effects of Exercise Training on Handgrip Strength in Older Adults: A Meta-Analytical Review. . *Gerontology*. 2019;65:686-98.
54. Kahraman T, Kahraman BO, Sengul YS, & Kalemci O. Assessment of sit-to-stand movement in nonspecific low back pain: a comparison study for psychometric properties of field-based and laboratory-based methods. . *International Journal of Rehabilitation Research*. 2016;39:165-70.
55. Blankevoort CG, Van Heuvelen MJ, & Scherder EJ. Reliability of six physical performance tests in older people with dementia. . *Physical therapy*. 2013;93(1):69-78.
56. Bohannon RW, Steffl M, Glenney SS, Green M, Cashwell L, Prajerova K, & Bunn J. The prone bridge test: Performance, validity, and reliability among older and younger adults. . *Journal of bodywork and movement therapies*. 2018;22(2):385-9.

LIMITATIONS AND STRENGTHS

LIMITATIONS AND STRENGTHS

The studies included in this International Doctoral Thesis present the general limitations and strengths, which are summarized below.

General Limitations

I. In the systematic review, only apparently healthy populations were included at baseline, so it is possible that some tests may be feasible or safe in some specific pathologies or health conditions (*study I*).

II. About the methodological study it could be mention that this did not include “RPE” registration in all the field-based physical fitness tests. This scale has shown to be a simple measure that can be used as a complement to “HR” monitoring for report the degree of exertion of the participants, particularly in the assessment of cardiorespiratory fitness tests (*study II*).

III. Finally, regarding to the study of responsiveness, the lack of control of factors that could affect the test performance, such as genetics, psychological issues, or/and diet control is the main limitation of this study (*study III*).

General Strengths

I. It is the first systematic review to examine the feasibility and safety of field-based fitness testing from original studies and the most up-to-date evidence for existing field-based fitness testing in people of all ages (*study I*).

II. This is the first study based on scientific evidence to determine the items to assess the degree of feasibility and safety of the most common health-related field-based physical fitness tests used in adults, which was conducted in a large and quite homogeneous sample in terms of sex, age and physical activity level (*study II*).

III. This is the first study that analyses of responsiveness of a wide range of health-related field-based physical fitness tests, and after a thoroughly controlled a supervised training based on multimodal exercises during 12 weeks in healthy adults (*study III*).

CONCLUSIONS [CONCLUSIONES]

CONCLUSIONS

Overall, the findings derived from the present International Doctoral Thesis suggest that:

Study I. *Feasibility and Safety of Field-Based Physical Fitness Test: A Systematic Review.*

I. To our knowledge, there are few studies that have analysed the feasibility and safety of field-based physical fitness tests, and overall, there is little evidence. Most of the tests analysed showed limited evidence, due to the lack of studies, especially in older adults. Based on quality assessment and established levels of evidence for each study, our findings indicate: strong evidence indicated that a) the 2-km walk and 20-m shuttle run tests were feasible to assess cardiorespiratory fitness in adults and children and adolescents, respectively; b) the handgrip strength test was feasible to assess isometric muscular strength in children and adolescents; and c) the standing long jump test was feasible to assess lower body explosive strength in children and adolescents. Regarding safety, only the 2-km walk test showed strong evidence assessing cardiorespiratory fitness in adults.

II. There is no consensus to assess the feasibility and safety of field-based physical fitness tests. It would be desirable that experts in this field generate methodological criteria, that could correctly reflect the feasibility or/and safety of field-based physical fitness test in people of all ages. In this sense, we have made a proposal of the items that we consider should be used to assess the feasibility and safety of field-based physical fitness tests.

Study II. *Feasibility and Safety in health-related field-based physical fitness test in adult population: The ADULT-FIT project.*

III. The most frequently used health-related field-based physical fitness tests are feasible and safe in adult population, regardless sex, age, and physical activity level. The findings of the present study allow to determine what field-based physical fitness tests are feasible and safe in adults, necessary preconditions to propose a health-related field-based physical fitness test battery.

Study III. *Responsiveness of health-related field-based physical fitness tests in adults: The ADULT-FIT project.*

IV. Overall, all the health-related field-based physical fitness tests analysed were found to be responsive after a multicomponent exercise training intervention regardless sex and age; except for changes in anthropometric outcomes and isometric handgrip strength, that were less responsive than the others and might require specifically targeted interventions to respond further. Assessing health-related physical fitness through field-based tests can be a useful assessment tool for determining the efficacy of different exercise program.

Final Conclusion

V. Through the findings of the present International Doctoral Thesis, and those found in a previous International Doctoral Thesis project titled “Validity and Reliability of field-based tests to assess physical fitness related to health in adults. The ADULT-FIT project” and in scientific studies currently under development, it could be concluded that it is possible to make a proposal for a health-related field-based physical fitness test battery in adults, based on scientific evidence based on its validity (i.e., predictive and criterion-related), reliability, feasibility, safety and responsiveness. Therefore, a final conclusion summarizes the evidence derived from these findings (see **Figure 8**):

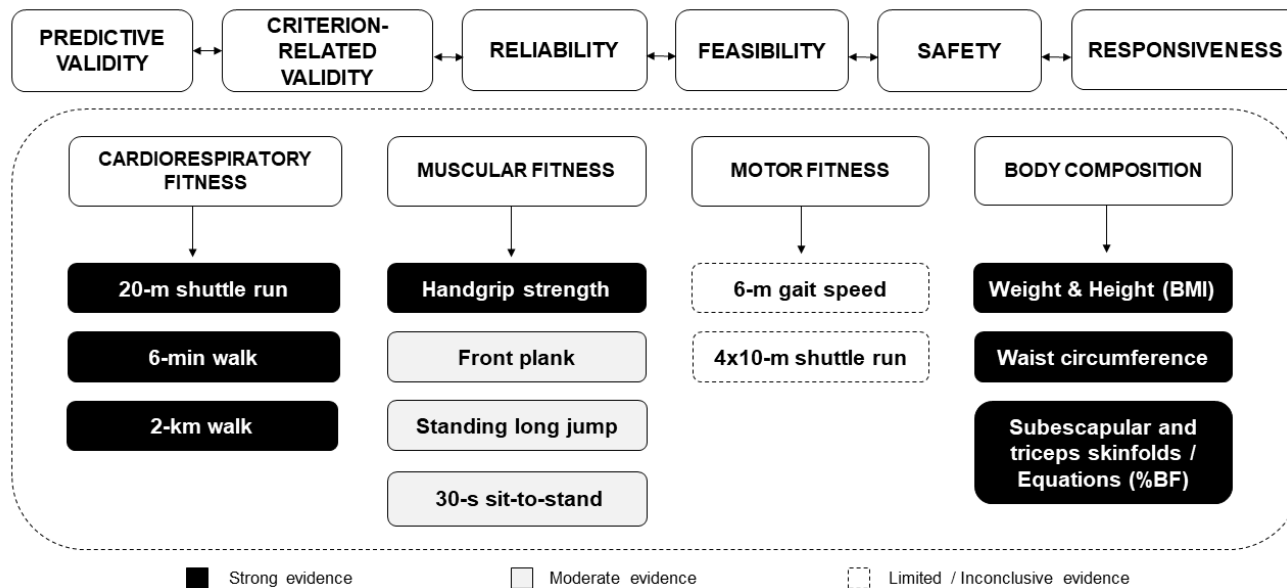


Figure 8. Proposal for an evidence-based health-related field physical fitness test battery for adults: The ADULT-FIT battery.

To assess cardiorespiratory fitness, the 20-m shuttle run, 6-min walk and 2-km walk tests are proposed with strong evidence based on their validity (predictive and criterion-related), reliability, feasibility, safety and responsiveness.

To assess muscular fitness, the handgrip strength is proposed with strong evidence and the front plank, standing long jump and 30-s sit-to-stand tests with moderate evidence based on their validity (predictive and criterion-related), reliability, feasibility, safety and responsiveness.

To assess motor fitness, the 6-m gait speed and 4x10-m shuttle run tests are proposed with limited evidence based on their validity (predictive and criterion-related), reliability, feasibility, safety and responsiveness.

Finally, to assess body composition, the BMI, waist circumference, subscapular and triceps skinfolds and estimation equations implying skinfolds are proposed with strong evidence based on their validity (predictive and criterion-related), reliability, feasibility, safety and responsiveness.

CONCLUSIONES

En general, los hallazgos derivados de la presente Tesis Doctoral Internacional sugieren que:

Estudio I. Viabilidad y seguridad de test de campo para evaluar la condición física: Una revisión sistemática.

I. Hasta donde sabemos, existen pocos estudios que hayan analizado la viabilidad y la seguridad de los test de campo para evaluar la condición física y, en general, las evidencias son limitadas. La mayoría de los test de campo analizadas mostraron evidencias limitadas, debido a la falta de estudios, especialmente en adultos mayores. Según la evaluación de la calidad y los niveles de evidencia establecidos para cada estudio, nuestros hallazgos indican que: los test de campo con fuerte evidencia indicaron que a) los test de caminata de 2-km y carrera de 20-m ida y vuelta fueron viables para evaluar la capacidad cardiorrespiratoria en adultos y niños y adolescentes, respectivamente; b) el test de fuerza de prensión manual fue viable para evaluar la fuerza muscular isométrica en niños y adolescentes; y c) el test de salto de longitud a pies juntos fue viable para evaluar la fuerza explosiva del tren inferior en niños y adolescentes. En cuanto a la seguridad, sólo el test de caminata de 2-km mostró fuerte evidencia para evaluar la capacidad cardiorrespiratoria en adultos.

II. No existe consenso para evaluar la viabilidad y la seguridad de los test de campo para evaluar la condición física. Sería deseable que los expertos en este campo generaran criterios metodológicos que pudieran reflejar correctamente la viabilidad y/o la seguridad de los test de campo para evaluar la condición física en personas de todas las edades. En este sentido, hemos realizado una propuesta de los ítems que consideramos deberían utilizarse para evaluar la viabilidad y seguridad de los test de campo para evaluar la condición física.

Estudio II. Viabilidad y seguridad en test de campo para evaluar la condición física relacionadas con la salud en población adulta: El proyecto ADULT-FIT.

III. Los test de campo para evaluar la condición física relacionada con la salud más utilizados son viables y seguros en población adulta, independientemente del sexo, edad y nivel de actividad física. Los resultados del presente estudio permiten determinar qué

test de campo para evaluar la condición física son viables y seguros en adultos, condiciones previas necesarias para proponer una batería de test de campo para evaluar la condición física relacionadas con la salud.

Estudio III. *Sensibilidad al cambio de rendimiento en test de campo para evaluar la condición física relacionada con la salud adultas: Proyecto ADULT-FIT.*

IV. En general, se observó que todos los test de campo para evaluar la condición física relacionada con la salud que se analizaron, mostraron ser sensibles al cambio tras una intervención de entrenamiento con ejercicios multicomponente independientemente del sexo y la edad; a excepción de los cambios en la composición corporal y la fuerza isométrica de prensión manual que resultaron ser menos sensibles al cambio de rendimiento que los otros y podrían requerir intervenciones específicamente dirigidas para mejorar su sensibilidad al cambio en el futuro. La evaluación de la condición física relacionada con la salud a través de test de campo puede ser una herramienta de evaluación útil para determinar la eficacia de diferentes programas de ejercicio.

Conclusión Final

V. A través de los hallazgos de la presente Tesis Doctoral Internacional, y de los resultados encontrados en un proyecto de Tesis Doctoral Internacional previo titulado “Validez y Fiabilidad de test de campo para evaluar la condición física relacionada con la salud en adultos. Proyecto ADULT-FIT” y en estudios científicos actualmente en desarrollo, se podría concluir que es posible realizar una propuesta de batería de test de campo para evaluar la condición física relacionada con la salud en adultos, basada en la evidencia científica acerca de su validez (es decir, predictiva y de criterio), fiabilidad, viabilidad, seguridad y sensibilidad al cambio de rendimiento. Por consiguiente, una conclusión final resume la evidencia derivada de estos hallazgos (véase **Figura 8**):

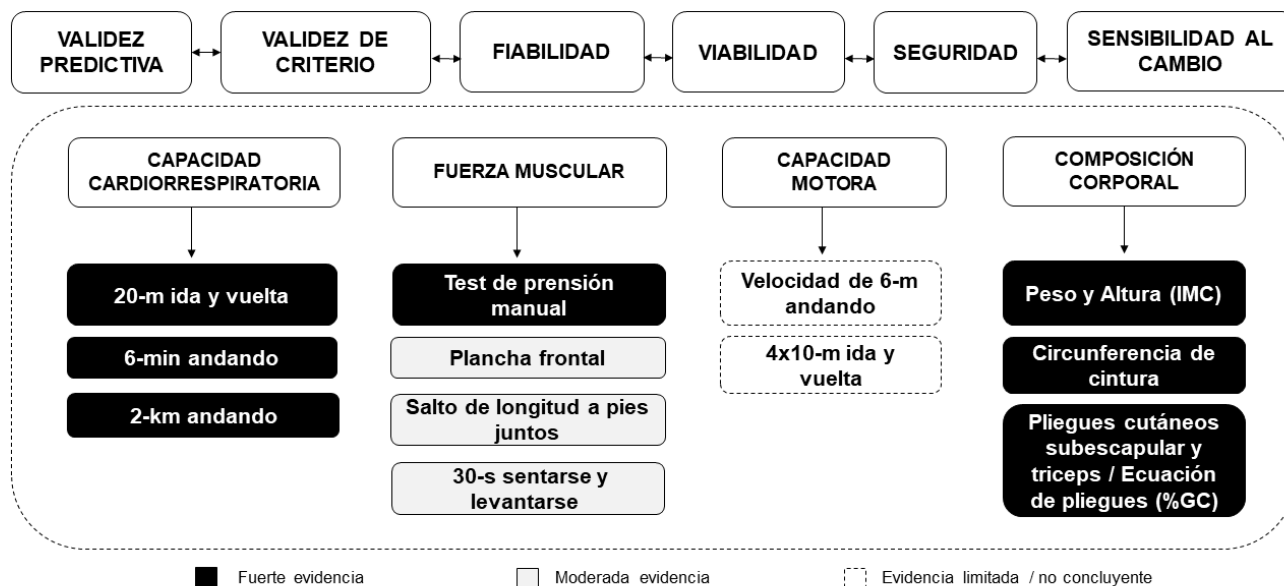


Figura 8. Propuesta de una batería de test de campo para evaluar la condición física relacionada con la salud en adultos basada en la evidencia: La batería ADULT-FIT.

Para evaluar la capacidad cardiorrespiratoria, se proponen los test de 20-m ida y vuelta, 6-min andando y 2-km andando, con fuerte evidencia en base a su validez (predictiva y de criterio), fiabilidad, viabilidad, seguridad y sensibilidad al cambio de rendimiento.

Para evaluar la fuerza muscular, se proponen el test de fuerza de prensión manual con fuerte evidencia, y los test de plancha frontal, salto de longitud a pies juntos y 30-s sentarse y levantarse con evidencia moderada en base a su validez (predictiva y de criterio), fiabilidad, viabilidad, seguridad y sensibilidad al cambio de rendimiento.

Para evaluar la capacidad motora, se proponen los test de velocidad de 6-m andando y 4x10-m ida y vuelta, con evidencia limitada / no concluyente en base a su validez (predictiva y de criterio), fiabilidad, viabilidad, seguridad y sensibilidad al cambio de rendimiento.

Por último, para evaluar la composición corporal, se proponen el IMC, la circunferencia de cintura, los pliegues cutáneos subescapular y tríceps y las ecuaciones de estimación que implican los pliegues cutáneos, con fuerte evidencia en base a su validez (predictiva y de criterio), fiabilidad, viabilidad, seguridad y sensibilidad al cambio de rendimiento.

FUTURE RESEARCH DIRECTIONS

FUTURE RESEARCH DIRECTIONS

The studies derived from the present International Doctoral Thesis provide evidence to propose a health-related field-based physical fitness test battery for adults (*The ADULT-FIT health-related fitness test battery*) based on its feasibility, safety and responsiveness, according to sex, age, and physical activity level.

Some future research directions that can be developed based on the findings of both International Doctoral Thesis and complement the current evidence would be the assessment of the ADULT-FIT battery to: **(i)** establish reference points on the physical fitness of adults according to sex and age and, **(ii)** standardize the assessment of health-related fitness tests among adults, enabling the derivation of normative reference points for physical fitness in this population. **(iii)** Finally, it would be advisable to design a battery of field-based tests to assess health-related physical fitness in older adults based on scientific evidence, valid (predictive and criterion-related), reliable, feasible, safe and responsive according to sex, age and physical activity level. This would enable the assessment of physical fitness pertinent to health among this population, that covers all the aims proposed in the ADULT-FIT study.

REFERENCES

1. Castro-Pinero J, Artero EG, Espana-Romero V, Ortega FB, Sjostrom M, Suni J, Ruiz JR. Criterion-related validity of field-based fitness tests in youth: a systematic review. *British Journal of Sports Medicine*. 2009;44(13):934-43.
2. Castro-Piñero J, Marin-Jimenez, N., Fernandez-Santos, J. R., Martin-Acosta, F., Segura-Jimenez, V., Izquierdo-Gomez, R., ..., Cuenca-García M. Criterion-Related Validity of Field-Based Fitness Tests in Adults: A Systematic Review. *Journal of clinical medicine*. 2021;10:3743.
3. Artero EG, Espana-Romero, V., Castro-Pinero, J., Ortega, F. B., Suni, J., Castillo-Garzon, M. J., Ruiz JR. Reliability of field-based fitness tests in youth. *International journal of sports medicine*. 2011;32:159-69.
4. Cuenca-García M, Marin-Jimenez N, Perez-Bey A, Sánchez-Oliva D, Camiletti-Moiron D, Alvarez-Gallardo IC, Castro-Piñero J. Reliability of Field-Based Fitness Tests in Adults: A Systematic Review. *Sports Medicine*. 2022:1-20.
5. Bruggeman BS, Vincent HK, Chi X, Filipp SL, Mercado R, Modave F, et al. Simple tests of cardiorespiratory fitness in a pediatric population. *Plos one*. 2020;15(9):e0238863.
6. Lamonedá J, Huertas-Delgado, F. J., Cadenas-Sanchez C. Feasibility and concurrent validity of a cardiorespiratory fitness test based on the adaptation of the original 20 m shuttle run: The 20 m shuttle run with music. *Journal of Sports Sciences*. 2021;39:57-63.
7. Laukkanen RM, Oja P, Ojala KH, Pasanen ME, Vuori IM. Feasibility of a 2-km walking test for fitness assessment in a population study. . *Scandinavian journal of social medicine*. 1992;20:119-26.
8. Hébert LJ, Maltais, D. B., Lepage, C., Saulnier, J., Crête, M., Perron M. Isometric muscle strength in youth assessed by hand-held dynamometry: A feasibility, reliability, and validity study: A feasibility, reliability, and validity study. *Pediatric Physical Therapy*. 2011;23:289-99.
9. Suni JH, Miilunpalo SI, Asikainen TM, Laukkanen RT, Oja P, Pasanen ME, Vuori IM. Safety and feasibility of a health-related fitness test battery for adults. *Physical Therapy*. 1998;78:134-48.
10. McAllister LS, Palombaro KM. Modified 30-second sit-to-stand test: reliability and validity in older adults unable to complete traditional sit-to-stand testing. *Journal of Geriatric Physical Therapy*. 2020;43:153-8.
11. José A, Dal Corso S. Step tests are safe for assessing functional capacity in patients hospitalized with acute lung diseases. *Journal of Cardiopulmonary Rehabilitation and Prevention*. 2016;36:56-61.
12. Borel B, Fabre, C., Saison, S., Bart, F., Grosbois JM. An original field evaluation test for chronic obstructive pulmonary disease population: the six-minute stepper test. *Clinical rehabilitation*. 2010;24:82-93.
13. Amado-Pacheco JC, Prieto-Benavides, D. H., Correa-Bautista, J. E., García-Hermoso, A., Agostinis-Sobrinho, C., María Alonso-Martínez, A., ..., Ramírez-Vélez R. Feasibility and reliability of physical fitness tests among colombian preschool children. *International journal of environmental research and public health*. 2019;16:3069.
14. Langhammer B, Stanghelle JK. Senior fitness test; a useful tool to measure physical fitness in persons with acquired brain injury. *Brain injury*. 2019;33:183-8.
15. Oja P, Laukkanen R, Pasanen M, Tyry T, Vuori I. A 2-km walking test for assessing the cardiorespiratory fitness of healthy adults. 1991;12:356-62.

16. Cadenas-Sanchez C, Alcántara-Moral F, Sanchez-Delgado G, Mora-Gonzalez J, Martinez-Tellez B, Herrador-Colmenero M, Ortega FB. Assessment of cardiorespiratory fitness in preschool children: adaptation of the 20 metres shuttle run test. . *Nutricion hospitalaria*. 2014;30:1333-43. .
17. Cadenas-Sanchez C, Martinez-Tellez B, Sanchez-Delgado G, Mora-Gonzalez J, Castro-Pinero J, Lof M, et al. Assessing physical fitness in preschool children: Feasibility, reliability and practical recommendations for the PREFIT battery. *J Sci Med Sport*. 2016;19(11):910-5.
18. Espana-Romero V, Artero EG, Jimenez-Pavon D, Cuenca-Garcia M, Ortega FB, Castro-Pinero J, et al. Assessing health-related fitness tests in the school setting: reliability, feasibility and safety; the ALPHA Study. *Int J Sports Med*. 2010;31(7):490-7.
19. Aadahl M, Zacho M, Linneberg A, Thuesen B, H., Jørgensen T. Comparison of the Danish step test and the watt-max test for estimation of maximal oxygen uptake: the Health2008 study. *European journal of preventive cardiology*. 2013;20:1088-94.
20. Smits-Engelsman B, Bonney E, Neto J. L. C., Jelsma DL. Feasibility and content validity of the PERF-FIT test battery to assess movement skills, agility and power among children in low-resource settings. *BMC Public Health*. 2020;20:1-11.
21. Boyer C, Tremblay M, Saunders T, McFarlane A, Borghese M, Lloyd M, Longmuir P. Feasibility, Validity, and Reliability of the Plank Isometric Hold as a Field-Based Assessment of Torso Muscular Endurance for Children 8-12 Years of Age. *Pediatric Exercise Science*. 2013;25(3):407-22.
22. Ito T, Shirado O, Suzuki H, Takahashi M, Kaneda K, Strax TE. Lumbar trunk muscle endurance testing: an inexpensive alternative to a machine for evaluation. *Archives of physical medicine and rehabilitation*. 1996;77:75-9.
23. Fjørtoft I, PAVSH, Vereijken B. Measuring physical fitness in children who are 5 to 12 years old with a test battery that is functional and easy to administer. *Physical therapy*. 2011;91:1087-95.
24. Cadenas-Sanchez C, Sanchez-Delgado G, Martinez-Tellez B, Mora-Gonzalez J, Löf M, Espana-Romero V, et al. Reliability and validity of different models of TKK hand dynamometers. *The American Journal of Occupational Therapy*. 2016;70(4):7004300010p1-p9.
25. Achten J, Jeukendrup AE. Heart rate monitoring: applications and limitations. *Sports Med*. 2003;33:517-38.
26. Liguori G, American College of Sports M. ACSM's guidelines for exercise testing and prescription . . USA2021.
27. Sanzen LFA, Westlin N. Anterior tibial compartment pressure during race walking. . *The American journal of sports medicine*. 1986;14:136-8.
28. Chleboun GS, Busic AB, Graham KK, Stuckey HA. Fascicle length change of the human tibialis anterior and vastus lateralis during walking. . *Journal of orthopaedic & sports physical therapy*. 2007;37:372-9.
29. Chen MJ, Fan X, Moe ST. Criterion-related validity of the Borg ratings of perceived exertion scale in healthy individuals: a meta-analysis. *Journal of sports sciences* , . 2002;20: 873-99.
30. Cheung KHPA, Maxwell L. Delayed onset muscle soreness. . *Sports medicine*. 2003;33:145-64.
31. Garcia-Hermoso A, Cavero-Redondo I, Ramirez-Velez R, Ruiz JR, Ortega FB, Lee DC, Martinez-Vizcaino V. Muscular Strength as a Predictor of All-Cause Mortality in an Apparently

Healthy Population: A Systematic Review and Meta-Analysis of Data From Approximately 2 Million Men and Women. *Arch Phys Med Rehabil.* 2018;99(10):2100-13 e5.

32. Harber M, Kaminsky L, Arena R, Blair S, Franklin B, Myers J, al. e. Impact of cardiorespiratory fitness on all-cause and disease-specific mortality: advances since 2009. . *Progress in cardiovascular diseases.* 2017;60:11-20.

33. Marin-Jimenez N, Cruz-Leon C, Sanchez-Oliva D, Jimenez-Iglesias J, Caraballo I, Padilla-Moledo C, Castro-Piñero J. Criterion-Related Validity of Field-Based Methods and Equations for Body Composition Estimation in Adults: A Systematic Review. . *Current Obesity Reports.* 2022;1-14.

34. Marín-Jiménez N, Cruz-León, C., Perez-Bey, A., Conde-Caveda, J., Grao-Cruces, A., Aparicio, V. A., ..., Cuenca-García M. Predictive Validity of Motor Fitness and Flexibility Tests in Adults and Older Adults: A Systematic Review. *Journal of Clinical Medicine.* 2022;11:328.

35. Cruz-León C, Marín-Jiménez N, Sánchez-Parente S, Borges Cosic M, Grao-Cruces A, Castro-Piñero J, Cuenca-Garcia M. Feasibility and Safety in health-related field-based physical fitness tests in adult population: The ADULT-FIT project. *Scandinavian Journal of Medicine & Science in Sports.* 2024;Under Rev.

36. Hopkins WG. A Spreadsheet for Monitoring an Individual's Changes and Trend. *Sportscience.* 2017;21:5-9.

37. Guo L, Chen J, Yuan W. The effect of HIIT on body composition, cardiovascular fitness, psychological well-being, and executive function of overweight/obese female young adults. *Frontiers in Psychology.* 2023;13:1095328.

38. Keating SE, Johnson NA, Mielke GI, Coombes JS. A systematic review and meta-analysis of interval training versus moderate-intensity continuous training on body adiposity. . *Obesity reviews.* 2017;18:943-64.

39. Wewege M, Van Den Berg R, Ward RE, Keech A. The effects of high-intensity interval training vs. moderate-intensity continuous training on body composition in overweight and obese adults: a systematic review and meta-analysis. . *Obesity Reviews* 2017;18:635-46.

40. Wilson JM, Marin PJ, Rhea MR, Wilson SM, Loenneke JP, Anderson JC. Concurrent training: a meta-analysis examining interference of aerobic and resistance exercises. . *The Journal of Strength & Conditioning Research.* 2012;26:2293-307.

41. Sultana RN, Sabag A, Keating SE, Johnson NA. The effect of low-volume high-intensity interval training on body composition and cardiorespiratory fitness: a systematic review and meta-analysis. . *Sports Medicine.* 2019;49:1687-721.

42. Clark JE. Diet, exercise or diet with exercise: comparing the effectiveness of treatment options for weight-loss and changes in fitness for adults (18–65 years old) who are overfat, or obese; systematic review and meta-analysis. . *Journal of Diabetes & Metabolic Disorders.* 2015;14:1-28.

43. Stiegler P, Cunliffe A. The role of diet and exercise for the maintenance of fat-free mass and resting metabolic rate during weight loss. . *Sports medicine.* 2006;36:239-62.

44. Ahmaidi SB, Varray AL, Savy-Pacaux AM, Pnefault CG. Cardiorespiratory fitness evaluation by the shuttle test in asthmatic subjects during aerobic training. . *Chest.* 1993;103:1135-41.

45. Laukkanen RMT, Kukkonen-Harjula TK, Oja P, Pasanen ME, Vuori IM. Prediction of change in maximal aerobic power by the 2-km walk test after walking training in middle-aged adults. . *International journal of sports medicine.* 2000;21:113-6.

46. Holland AE, Spruit MA, Troosters T, Puhan MA, Pepin V, Saey D, ... & Singh SJ. An official European Respiratory Society/American Thoracic Society technical standard: field walking tests in chronic respiratory disease. . *European Respiratory Journal*. 2014;44:1428-46.
47. O'Keeffe S, Lye M, Donnellan C, Carmichael D. Reproducibility and responsiveness of quality of life assessment and six minute walk test in elderly heart failure patients. . *Heart*. 1998;80:337.
48. Bellet RN, Adams L, Morris NR. The 6-minute walk test in outpatient cardiac rehabilitation: validity, reliability and responsiveness: a systematic review. . *Physiotherapy*. 2012;98:277-86.
49. Wilken JM, Darter BJ, Goffar SL, Ellwein JC, Snell RM, Tomalis EA, Shaffer SW. Physical performance assessment in military service members. *JAAOS-Journal of the American Academy of Orthopaedic Surgeons*. 2012;20:S42-S7.
50. Marín-Jiménez N, Sánchez-Parente S, Expósito-Carrillo P, Jiménez-Iglesias J, Álvarez-Gallardo IC, Cuenca-García M, & Castro-Piñero J. Criterion-related validity and reliability of the 2-km walk test and the 20-m shuttle run test in adults: The role of sex, age and physical activity level. . *Journal of Science and Medicine in Sport*. 2023;26:267-76.
51. Alfonso-Rosa RM, del Pozo-Cruz B, del Pozo-Cruz J, Sanudo B, & Rogers ME. Test-retest reliability and minimal detectable change scores for fitness assessment in older adults with type 2 diabetes. *Rehabilitation Nursing Journal*. 2014;39:260-8.
52. Aktar B, Balci B, Oztura I, & Baklan B. The test-retest reliability and minimal detectable change of the six-minute walk test, timed up and go test, and 30-second chair stand test in people with epilepsy. . *Physiotherapy theory and practice*. 2023;1-10.
53. Perera S, Mody SH, Woodman RC, & Studenski SA. Meaningful change and responsiveness in common physical performance measures in older adults. . *Journal of the American Geriatrics Society*. 2006;54:743-9.
54. Bohannon RW. Test-Retest Reliability of Measurements of Hand-Grip Strength Obtained by Dynamometry from Older Adults: A Systematic Review of Research in the PubMed Database. . *The Journal of frailty & aging*. 2017;6:83-7.
55. Labott B, H. B, Morat M, Morat T, Donath L. Effects of Exercise Training on Handgrip Strength in Older Adults: A Meta-Analytical Review. . *Gerontology*. 2019;65:686-98.
56. Soysal P, Hurst C, Demurtas J, Firth J, Howden R, Yang L, et al. Handgrip strength and health outcomes: Umbrella review of systematic reviews with meta-analyses of observational studies. . *Journal of Sport and Health Science*. 2021;10:290-5.
57. Kahraman T, Kahraman BO, Sengul YS, & Kalemci O. Assessment of sit-to-stand movement in nonspecific low back pain: a comparison study for psychometric properties of field-based and laboratory-based methods. . *International Journal of Rehabilitation Research*. 2016;39:165-70.
58. Blankevoort CG, Van Heuvelen MJ, & Scherder EJ. Reliability of six physical performance tests in older people with dementia. . *Physical therapy*. 2013;93(1):69-78.
59. Bohannon RW, Steffl M, Glenney SS, Green M, Cashwell L, Prajerova K, & Bunn J. The prone bridge test: Performance, validity, and reliability among older and younger adults. . *Journal of bodywork and movement therapies*. 2018;22(2):385-9.

ANNEXES

SHORT CV

CURRICULUM VITAE [SHORT CV]

Personal Data

Name: Carolina Vaeny Cruz León

E-mail: carolina.cruz@uca.es

Birth Date: 19/03/1988

Birth Place: Viña del Mar

Nationality: Spanish

ORCID profile: [0000-0001-9070-5693](https://orcid.org/0000-0001-9070-5693)

Research Gate: Carolina Cruz-León

Academic background

- 2019-2024 PhD candidate in Health Sciences. University of Cádiz, Spain.
- 2015-2016 Master's degree in Physical Activity for Health. Pontifical University of Valparaíso, Chile.
- 2008-2014 Bachelor's Degree in Physical Education. Pontifical University of Valparaíso, Chile.

Research contracts

- 2022-2023 Technical Research Graduate / University of Cadiz / Spain.
- Funding Entity Municipal Parking Company, Scientific, technical or artistic work in collaboration with other entities or individuals under Article 83 of the Organic Law of Universities (OTRI Contract).
- 2019-2021 Research Technician / University of Cadiz / Spain.
- Funding Entity Spanish Ministry of Economy, Industry and Competitiveness (Plan nacional I+D+i 2017).

International interships

- 2022 Research mobility: Faculty of Rehabilitation Science, Andres Bello University, Chile. Dr. Christian Campos Jara (4 months).

Participation in research projects

Researcher 2023

Project Title & Design: Programa de fomento de desplazamiento activo en bicicleta para escolares de la ciudad de Cádiz (BiciCole).

Principal Investigator: Daniel Camiletti Moirón and Rocio Izquierdo Gómez

Department of Physical Education and Sport, University of Cadiz, Spain.

Funding: Municipal Parking Company, Scientific, technical or artistic work in collaboration with other entities or individuals under Article 83 of the Organic Law of Universities (OTRI Contract) (62.532€).

Researcher 2019–...

Project Title & Design: Health-related physical fitness test battery design for the adult population: The ADULT-FIT project.

Principal Investigator: José Catro Piñero and Magdalena Cuenca García.

Department of Physical Education and Sport, University of Cadiz, Spain.

Funding: Spanish Ministry of Economy, Industry, and Competitiveness (112.000€).

Researcher 2018

Project Title & Design: Asociación entre desplazamiento activo condición física y rendimiento académico en estudiantes universitarios.

Principal Investigator: Maribel Parra Saldías.

Department of Physical Education and Sport, Pontifical University of Valparaiso, Chile.

Funding: Physical Education Department (\$400.000).

Researcher 2017

Project Title & Design: Estudio sobre la contribución de las sesiones de las escuelas deportivas integrales de iniciación deportiva al incremento del tiempo de actividad física en niños y niñas participantes del programa de escuela deportivas integrales.

Principal Investigator: Fernando Rodríguez-Rodríguez.

Department of Physical Education and Sport, Pontifical University of Valparaíso, Chile.

Funding: National Sports Institute (\$49.000.000).

Scientific and organizing committee member activities

2022 1st INIBICA MEETING FOR YOUNG RESEARCHER ABROAD (MYRA).

Virtual event, Cadiz, Spain. Organising committee

Publications (*not included as part of this Doctoral Thesis*)

1. Castro-Piñero, J., Aragón-Aragón, P., **Cruz-León, C.**, Jiménez-Iglesias, J., Camiletti-Moirón, D., Izquierdo-Gómez, R., & Cuenca-García, M. (2023). Time measurement validity and reliability of the 4× 10-m shuttle run test in adult population: The ADULT-FIT project. *Journal of Science and Medicine in Sport*, 26(10), 553-560. DOI: 10.1016/j.jsams.2023.08.176
2. Grao-Cruces, A., Pérez-Bey, A., Jiménez-Iglesias, J., **Cruz-León, C.**, Cabanas-Sánchez, V., Veiga, O. L., & Castro-Piñero, J. (2023). Influence of Volume and Bouts of Sedentary Time and Physical Activity on School-Aged Youth's Physical Fitness: The UP & DOWN Longitudinal Study. *Journal of Physical Activity and Health*, 20(2), 142-148. DOI: 10.1123/jpah.2022-0406
3. Marín-Jimenez, N., **Cruz-León, C.**, Sanchez-Oliva, D., Jimenez-Iglesias, J., Caraballo, I., Padilla-Moledo, C., ... & Castro-Pinero, J. (2022). Criterion-related validity of field-based methods and equations for body composition estimation in adults: a systematic review. *Current Obesity Reports*, 11(4), 336-349. DOI: 10.1007/s13679-022-00488-8
4. Marín-Jiménez, N., **Cruz-León, C.**, Perez-Bey, A., Conde-Caveda, J., Grao-Cruces, A.,

Aparicio, V. A., ... & Cuenca-García, M. (2022). Predictive Validity of Motor Fitness and Flexibility Tests in Adults and Older Adults: A Systematic Review. *Journal of Clinical Medicine*, 11(2), 328. DOI: 10.3390/jcm11020328

5. Barranco-Ruiz, Y., **Cruz-León, C.**, Villa-González, E., Palma Leal, X., Chillón, P., & Rodríguez-Rodríguez, F. (2019). Active commuting to university and its association with sociodemographic factors and physical activity levels in Chilean students. *Medicina*, 55(5), 152. DOI: 10.3390/MEDICINA55050152

Submitted publications (not included as part of this Doctoral Thesis)

1. Pérez-Bey, A., Sánchez-Delgado, A., **Cruz-León, C.**, Padilla-Moledo, C., Izquierdo-Gómez, R., Aadland, E., L Veiga, O., Castro-Piñero, J. (2023). Muscular fitness, neck circumference, and blood pressure: An insight into their relationship through cross-lagged panel analysis. *The UP&DOWN Study. The Indian Journal of Pediatrics*, sometido (Q1).

2. Marín-Jiménez, N., Pérez-Bey, A., **Cruz-León, C.**, Conde-Caveda, J., Segura-Jiménez, V., Castro-Piñero, J., Cuenca-García, M. (2023). Criterion-related validity and reliability of the standing long jump test in adult population: The Adult-Fit project. *Scandinavian Journal of Medicine and Science in Sports*, sometido (Q1).

3. Rodríguez-Perea, A., Aragón-Aragón, P., Cuenca-García, M., **Cruz-León, C.**, Torres-Banduc, M., Sánchez-Parente, S., Castro-Piñero, J. Criterion-related validity and reliability of the front plank test in adults: The ADULT-FIT Project. (2023). *Journal of Electromyography and Kinesiology*, sometido (Q3).

4. Jimenez-Iglesias, J., Owen, A., **Cruz-León, C.**, Campos-Vázquez, MA., Sánchez-Parente, S., Gonzalo-Skok, O., Cuenca-García, M., Castro-Piñero, J. (2023). Improving Change of Direction in Male Soccer Players Through Plyometric Training: A Systematic Review. *Journal of Sports Sciences for Health*, sometido (Q3).

Other merits

Five communications or posters both in national and international conferences.

ACKNOWLEDGMENTS [AGRADECIMIENTOS]

ACKNOWLEDGMENTS [AGRADECIMIENTOS]

Departamento de Didáctica de la Educación Física, Plástica y Musical

Facultad de Ciencias de la Educación
Universidad de Cádiz

