



Article

Development of a Support System for Physicians and Patients during Rehabilitation

Luisa Barrera-Leon ¹, Massimo Canonico ^{1,*}, Francesco Desimoni ^{1,*}, Alessandro de Sire ^{2,3},
Marco Invernizzi ^{4,5} and Lorenzo Lippi ^{4,5}

¹ Department of Science and Technological Innovation, University of Piemonte Orientale, 15121 Alessandria, Italy; luisa.barrera@uniupo.it

² Department of Medical and Surgical Sciences, Magna Graecia University of Catanzaro, 88100 Catanzaro, Italy; alessandro.desire@unicz.com

³ Research Center on Musculoskeletal Health, Magna Graecia University of Catanzaro, 88100 Catanzaro, Italy

⁴ Department of Health Sciences, University of Eastern Piedmont 'Amedeo Avogadro', 28100 Novara, Italy; marco.invernizzi@med.uniupo.it (M.I.); lorenzolippi.mt@gmail.com (L.L.)

⁵ Azienda Ospedaliera Nazionale Santi Antonio e Biagio e Cesare Arrigo, 15121 Alessandria, Italy

* Correspondence: massimo.canonico@uniupo.it (M.C.); francesco.desimoni@uniupo.it (F.D.)

Abstract: Musculoskeletal disorders are common among older adults, affecting mobility and quality of life. Effective rehabilitation is essential, but the implementation of programs faces challenges. Traditional methods often necessitate in-person assessments, which can be difficult for older adults with mobility limitations. Telerehabilitation offers a solution, bringing therapy closer to patients. However, the accurate remote monitoring of health and performance remains a challenge. This study addresses this gap by developing and validating the System for Tracking and Evaluating Performance (STEP). STEP is a hardware-software system that automates physical performance tests, eliminating the need for constant expert supervision. The system focuses on three standard tests: the Six-Minute Walking Test (6MWT), the Ten-Meter Walking Test (10MWT), and the 30-s Sit-to-Stand Test (30STS). Validation compared results from the STEP app with in-person assessments by physicians for patients undergoing rehabilitation after knee or hip arthroplasty. The study found strong positive correlations between the app's results and the physicians' assessments for all tests. These findings demonstrate the STEP system's potential as a reliable tool for remote physical performance assessment. Further research is needed to explore its integration into clinical practice and cost-effectiveness in reducing the need for operator assistance in monitoring patients with physical limitations.

Keywords: wearable devices; telemedicine; cloud computing; sensors



Citation: Barrera-Leon, L.; Canonico, M.; Desimoni, F.; de Sire, A.; Invernizzi, M.; Lippi, L. Development of a Support System for Physicians and Patients during Rehabilitation. *Biomechanics* **2024**, *4*, 520–541. <https://doi.org/10.3390/biomechanics4030037>

Academic Editor: Tibor Hortobágyi

Received: 16 April 2024

Revised: 11 July 2024

Accepted: 19 August 2024

Published: 4 September 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The global population is rapidly aging, leading to a significant rise in the elderly population [1,2]. An increase in life expectancy is often accompanied by a decline in health [3–6]; thus, maximizing the quality of life for elderly people becomes crucial. Rehabilitation plays a vital role in achieving this goal. An important part of it is tracking recovery from injuries and surgeries, which are common among the elderly. Through continuous assessment of physical performance and health status, rehabilitation protocols can be tailored to individual needs, ultimately improving quality of life.

Standardized tools like *submaximal exercise tests* (e.g., Six-Minute Walking Test) [7] offer a safe and practical way to assess various health aspects in patients with limitations during rehabilitation programs. These tests operate below a person's maximum capacity, ensuring safety and practicality for a broader range of patients, especially the elderly population targeted in this study. In contrast, other exercise tests, such as *maximal tests* (e.g., ergometer test), push patients to their physical limits and require close in-person monitoring [7]. Given these considerations, the physicians involved in the study selected three submaximal

tests for implementation in STEP: *Six-Minute Walking Test* (6MWT), *Ten-Meter Walking Test* (10MWT), and *Sit-to-Stand Test* (30STS).

Although alternative submaximal tests can provide more granular data points, the selected tests prioritize patient safety and offer a practical approach to evaluating key aspects of patient health, including endurance, aerobic capacity, lower body strength, and fall risk [8–13]. These three tests require minimal equipment and are readily accessible for patients to perform. In contrast, the submaximal Wingate test [14] provides a more precise estimate of aerobic capacity and lower body strength but requires specialized equipment and expertise. These reasons make the 6MWT, the 10MWT, and the 30STS ideal for use in various settings and with patients with diverse physical capabilities.

Despite this, standard tests often require supervision, limiting testing frequency and potentially missing performance fluctuations [15–19]. However, testing remains essential because patients can be unreliable in self-reporting [20–24]. The rise of *telemedicine* (remote healthcare delivery using technology) [25,26] and home-based rehabilitation technologies offers promise for overcoming these limitations [27–31].

Telemedicine enables testing in daily environments, potentially enhancing real-world applicability while maintaining result accuracy and allowing for the collection of additional data [32]. In this context, accurate sensor technology is critical. While some existing solutions require controlled settings [33,34], smartphones with built-in sensors (gyroscope, accelerometer, camera) offer a promising alternative due to their widespread availability and diverse functionalities, suitable for various submaximal tests [28–30,35–37].

Concerning the submaximal test development for telemedicine, studies have focused on the widespread use of inertial sensors, like those in smartphones, to monitor gait and other technologies for vital signs, specifically during the 6MWT [38]. However, developing mobile applications specifically designed to automate 10MWT and 30STS tests, with accessibility as a core principle, is an under-explored area within the current research literature. Aside from the limitations of the individual studies, we could not find any study focusing on multiple tests and integrating solutions for data exchange and analysis.

Considering the need to autonomously perform submaximal tests in health facilities and at home through easily accessible technologies, we propose a *System for Tracking and Evaluating Patient Performance* (STEP). STEP is a system composed of hardware and software components, aiming to automate the execution of submaximal tests using the resources available on a standard smartphone, eliminating the need for expert supervision. The STEP system also uploads the collected results and information online to easily keep track of the patient.

To assess reliability, the STEP system was validated during the rehabilitation of knee and hip replacement patients. Test results from the app were compared to those collected in person by physicians. Minimal discrepancies between the two methods were considered indicative of STEP's effectiveness.

This paper is structured as follows: Section 2 describes the selected tests along with some related work on their automation; Section 3 describes the development of the software, the study execution, the approach validation, and the data analysis; Section 4 presents the obtained results, while Section 5 discusses a few notes in detail and goes over future works; Section 6 presents the study conclusions.

2. Background

The STEP system revolves around submaximal tests, safe and practical methods for assessing various aspects of a patient's health without reaching maximum capacity [39]. This approach is preferred for patients with physical limitations due to age or medical procedures [7]. For the STEP system, the physicians involved in the study selected three submaximal tests:

- The *Six-Minute Walking Test* (6MWT), shown in Figure 1a, in which the patient is asked to walk back and forth on a flat surface (usually 30 m long) for six minutes. The patient is instructed to cover the greatest possible distance within the time limit,

turning around at the end of the walkway and stopping to rest as needed before resuming the walk [9]. The test result is the walked distance in meters.

- The *Ten-Meter Walking Test* (10MWT), shown in Figure 1b, in which the patient is asked to walk on a 10 m path, the first 2 m of which are used for acceleration and the last 2 for deceleration [40]. The test result is the walking speed calculated on the six central meters.
- The *Sit-to-Stand Test* (30STS), shown in Figure 1c, in which the patient is asked to stand up from a chair and sit back down for as many repetitions as possible, measuring the completed repetitions in a given time (or the time the patient took to complete a given amount of repetitions) [10]. The test result is the amount of time or the number of repetitions.

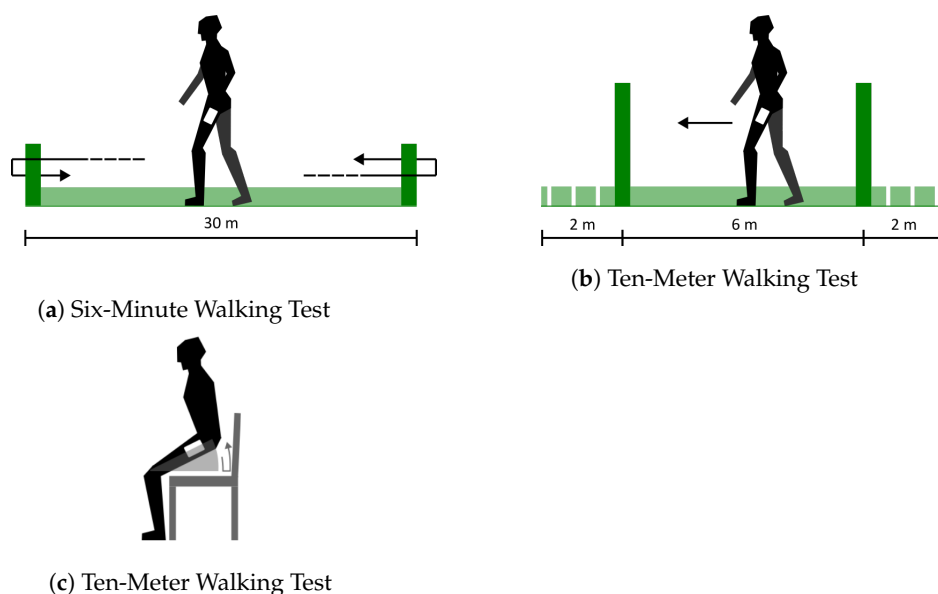


Figure 1. Graphic depiction of selected submaximal tests.

A substantial body of research has investigated the three tests we will be focusing on, with the 6MWT receiving the most attention. In a recent study [41], the authors asked healthy university students to perform the 6MWT using the EasyMeasure app to measure distances through the camera. This study encountered several issues which often prevented users from completing the test. In a similar study [42], the authors asked participants to perform the 6MWT while using an app to count steps to estimate the total distance using regression with other variables like age, height, and weight. Despite the promising results on healthy subjects, this approach may not translate to patients with an irregular gait, as sole reliance on step count during the test can be unreliable in these cases.

Other research studies have focused on developing mobile applications that can perform the tests autonomously. For instance, Cheng et al. [43,44] developed and validated an app to perform the 6MWT autonomously by detecting gait features. The authors of another study [45] also developed an app for the same purpose, using the accelerometer on a smartphone to count steps and collecting additional information from patients, such as heart rate and rate of perceived exertion on the Borg scale [46,47]. Scherrenberg et al. [48] tackled the same problem using GPS and Google Fit data. Capela et al. [49] detected walk distance and gait features through the smartphone's accelerometer and gyroscope sensors.

Another study automates the 6MWT using GPS outdoors and smartphone sensors for indoor direction changes [50]. Mainly, in order to calculate the indoor distance, the app uses smartphone sensors to detect direction changes, which are then multiplied by the length of the path on which the test is performed. In contrast, the remaining distance walked after the last direction change is estimated using the time median or the average step length. Despite smartphone-based 6MWT automation in these studies, results show

limitations. For example, one of the studies mentioned earlier [44] reports a 5.87% walk distance average error, which, on an average distance for 70 to 79 years old patients of 350–400 m [51], means an error of about 22 m. Additionally, studies lack an evaluation of patients exhibiting gait abnormalities.

The existing literature regarding the 10MWT and the 30STS is diminished, but still shows some interesting studies. A recent study [52] used an application to measure gait parameters during a 10MWT performed under supervision. In contrast, Saito et al. [53] attempted to create an app to automate the 10MWT for the general public. The implementation requires time-critical patient interaction that could prove problematic in a rehabilitation scenario and skew the results.

Several studies [54–56] utilize the non-automated 30STS for data collection. The tests in these studies were performed under supervision, but the sensors worn by patients provided more data for analysis. The supervised 30STS is also used in more complex health-monitoring systems [57,58]. In the existing literature, there are a few studies [34,59] where a wearable sensor and a smartwatch are used, respectively, to allow the test to be performed autonomously. While these are steps in the right direction, they only focused on one test, and the devices used were less widespread than a generic smartphone.

A comprehensive review of the existing literature reveals critical research gaps: the absence of applications capable of automating the 10MWT and 30STS tests using readily accessible devices and the lack of solutions that integrate all the tests, which would facilitate a more comprehensive monitoring of rehabilitation patients.

3. Materials and Methods

We developed a native Android mobile application that automates the tests using the device's sensors. The app allows users to customize a few parameters to better fit each patient and the environment in which the tests are performed. Moreover, it provides different solutions to simplify the process, like voice commands to begin and end a test. We also developed a system around the STEP app to ease data storage and retrieval for analysis and validation.

The entire architecture is depicted in Figure 2. The specifics of each architectural component are detailed in the following section, Section 3.1, and the algorithms implemented for each test are explained in Section 3.2.

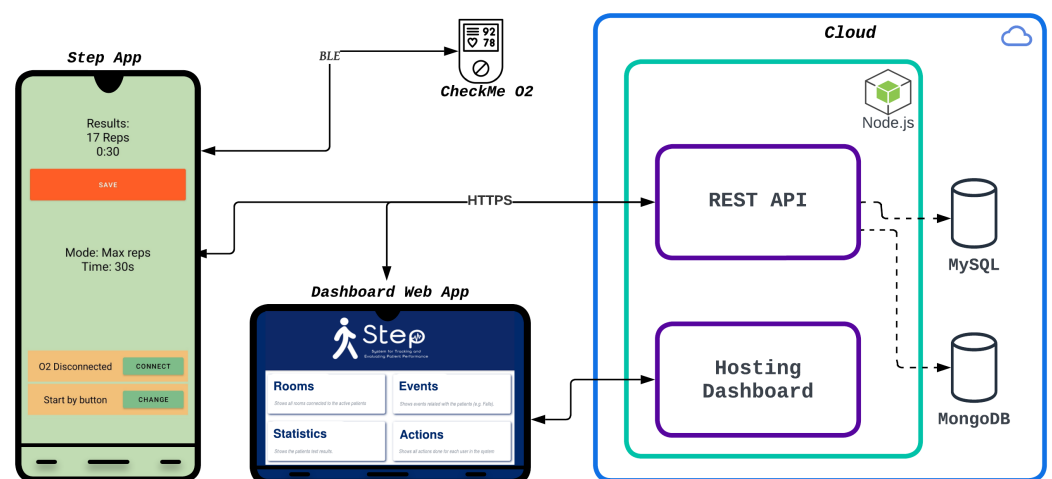


Figure 2. STEP system in architecture. The image depicts the STEP system's components: a cloud-based server (**right**) for data storage, the STEP app (**left**) for test execution and result collection, and a dashboard web app to visualize the results.

3.1. Hardware and Infrastructure

The main app runs on a smartphone, shown on the left side of Figure 2. Specifically, although the app was tested on multiple devices, we mostly used a Xiaomi Redmi

Note 11. This smartphone is equipped with a triaxial accelerometer and a triaxial gyroscope. Unfortunately, the exact models of the sensors are not available online.

The app is designed to connect to other nearby devices using *Bluetooth Low Energy* (BLE) to collect additional data about the patients. In this paper, we describe the use of one specific device in this role: the Viatom CheckMe O2. The Viatom Checkme O2 is a wrist oximeter that measures blood oxygen levels and heart rate in real time, among other features that are not crucial in this context. Hardware-wise, the device uses a ring sensor to collect oxygen saturation and heart rate data. At the same time, the main unit is a bracelet containing a BLE 4.0 chip, a 3.7v 130 mAh rechargeable battery, an accelerometer and a small OLED display. It also has an *Ingress Protection* (IP) rating of 22 and an internal memory able to store up to 40 h of recording data.

The app can receive data from this device through a BLE connection and the *Generic ATtribute* (GATT) [60] protocol. BLE is a wireless communication technology designed for low-power devices over short distances. It supports two types of devices: peripheral devices, which broadcast data packets (advertisements), and central devices, which scan for these advertisements to discover nearby peripherals (and connect to them if necessary). GATT defines how BLE devices exchange data. It organizes data into attributes (e.g., heart rate) grouped into services. Each service contains characteristics (e.g., the characteristic of heart rate) identified by *Universally Unique Identifiers* (UUIDs). Operations include reading, writing and enabling real-time updates from characteristics.

In this framework, the smartphone on which the app is installed acts as a central device. The smartphone discovers, connects, and sends commands to the CheckMe O2, acting as a peripheral device, which in turn answers by sending the current heart rate and oxygen saturation values. This feature is used during the execution of the 6MWT to collect values throughout the entire duration of the test and track meaningful variations.

The app can be configured with patient data and send test results by communicating with the server through *HyperText Transfer Protocol over Secure Socket Layer* (HTTPS) (see Figure 2, right side). HTTPS is an extension of the *Hypertext Transfer Protocol* (HTTP), a stateless application-level protocol used for exchanging data over the internet that operates in a client-server model. HTTPS ensures secure communication through encryption and authentication mechanisms. With HTTPS, data is encrypted during transmission, protecting against eavesdropping and tampering. Additionally, it relies on trusted third-party digital certificates to verify the server's authenticity.

The system's server consists of a series of Node.js [61] modules and two databases. Node.js is a runtime environment for the execution of server-side JavaScript code. While, at the core, it is built using the microservices architecture principles to ease maintenance and scalability, the server components run on a single Ubuntu 18.04 virtual machine hosted on Chameleon Cloud [62] for the limited scope of this study. Chameleon Cloud is a computing platform available for research purposes that provides computational resources and infrastructure for experimentation.

The backend logic resides within the Node.js modules, with the primary ones outlined below.

- *Main application programming interface* (API)—Two modules, based on the *Representational State Transfer* (RESTful) principles, handle all HTTPS requests and interact with the rest of the system accordingly. The API is built using Express.js [63]. Express.js is a Node.js library that provides a minimalistic framework for API creation. Its popularity made it a standard for Node.js backends, and it is part of several stacks like MEAN (MongoDB, Express, Angular, Node).
- *Database interaction*—Each main module that interacts with a database does so through one of these modules, created as an abstraction layer between logic and data. They connect to the database, perform queries, and return the results.
- *Authentication*—This module is an abstraction layer between the server logic and Firebase Auth [64], the third-party authentication service we used. It uses the Node.js library provided by Firebase to verify tokens sent to the server in HTTPS requests.

- *Hosting*—This module is the web application used to manage patients and view results in a clinical setting. It provides a simple API to receive HTTPS requests and return the web application as a response.

On the rightmost side of Figure 2, two databases are shown: the first one is a relational MySQL [65] database used for structured data about patients, the second one is a document MongoDB [61] database used to store test results and time series from the heart rate and oxygen saturation data.

Although the backend (the server side of the application that manages data, performs calculations, and interacts with databases) is used for data upload and download, it is optional for automating tests, since that is carried out entirely on the smartphone. The result is shown at the end of each test, as seen in Figure 2. The smartphone mockup shows an example of what is visible at the end of a test, specifically the 30STS. The number of completed repetitions is displayed at the top, along with the duration of the test. Since this test was executed in the *as many repetitions as possible in a given time* mode, the time was set at 30 s. The starting configuration for this test is visible at the center of the screen. The orange button at the top saves the results to the server if the user wishes. The buttons at the bottom connect to other devices (to collect additional data) and set the preferred method for starting the test.

While the system has been designed to be used in a clinical study setting, in order to validate the test automation methods, it can easily be adapted to be used in daily life settings by both physicians and patients, with the latter being able to seamlessly continue the rehabilitation process at home without the need for assistance. In greater detail, users could use the app “as a service” by taking advantage of the infrastructure we created, creating their own, or using one created by a third party following the same principles described here.

Although available on Google Play, the app is undergoing internal testing at the moment of writing and can only be downloaded by the people involved. It will be made public on the same platform at the end of this process. Furthermore, even if we developed an Android application, the same functionalities can be achieved on any device matching the hardware requirements.

3.2. Algorithms

Although similar, the three tests are based on fundamentally different movements and constraints. Therefore, while the smartphone remains the primary tool for automating the tests, how it is employed changes for each test. This section will review the algorithms implemented in each case, how the smartphone and its sensors are used, and how the final result is computed.

The STEP algorithms are applied to the sensor data without additional processing or filtering. In more detail, the Android platform performs a simple calibration on the sensor data, which, for the accelerometer, consists of temperature compensation, online bias calibration, and online scale calibration. At the same time, the gyroscope includes temperature compensation, online scale compensation, and online bias calibration. Preliminary testing showed no issues in processing the samples as returned by the Android system; therefore, we did not process them any further. While the Android API offers methods for retrieving uncalibrated sensor data, we determined it would not benefit our specific use case.

3.2.1. Six-Minute Walking Test

Our approach for this test is based on a few algorithms proposed in the existing literature [49,50]. Using the accelerometer, gyroscope, and timer present on a smartphone, an estimation of the total walking distance is returned at the end of 6 min.

The accelerometer is used to detect steps. While Android often provides a step counter, especially on newer devices, we implemented this functionality so that the app could work on older devices as well. This sensor is sampled at 30 Hz, obtaining a value for each axis at each interval. By applying a vector sum to these values, we get the magnitude representing

the total acceleration affecting the device at any given time. The resulting signal includes the influence of gravity, which is not filtered out. The algorithm works accordingly, knowing the magnitude will have a value of 1 G or 9.81 m/s^2 when the smartphone is at rest. Since the user is not expected to perform any other daily activity during the test, such as sitting down, the resulting acceleration signal is only used to detect steps and breaks. Walking generates a distinctive pattern in the signal, with clear peaks, and a simple threshold-based algorithm is reliable enough for this scenario.

Figure 3 shows an example of a signal along with the two thresholds. When the magnitude crosses the high threshold on the way up, a new step is counted, and no other steps will be counted until the signal crosses the low threshold on the way down. At that point, the algorithm will wait for the signal to cross the high threshold again. A simple calibration procedure, described below, was implemented in the app to find the best pair of thresholds for the user automatically.

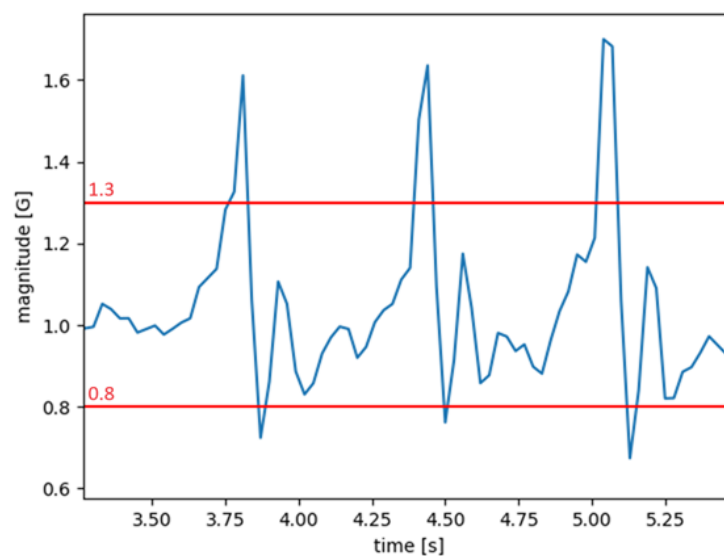


Figure 3. Overall acceleration (magnitude) generated during a walk (blue) with two example thresholds (red) that could be used in the algorithm. The three peaks correspond to three steps.

Using the gyroscope, sampled at 20 Hz, the direction changes at the ends of the path can be detected. The sample returned by the gyroscope, made up of three values, is the speed of the rotation along each of the three axes, which, if integrated, can return the rotation completed by the device in a given time window. Examining only the samples from the axis perpendicular to the ground and including them in a time window allows the algorithm to detect if a user performed a 180° turn within a given time. When this event is detected, a new direction change will be counted.

To minimize missed detections (false negatives), initial tests indicated a threshold below 180° was necessary. After further testing, a value of 140° proved to be a good balance between catching true turns and avoiding false positives. This value is thus used as the default, but, as discussed below, it can also be customized for each patient by manually changing the parameter. Similarly, a time window of 6 s proved to be the right fit for most patients during testing and was made the default, but this parameter can also be manually changed. For instance, the algorithm may need more time to accurately detect turns for a patient struggling to turn at the end of the path.

The walked distance, and therefore the test result, is estimated in the following way. Since the length of the path on which the test is performed is known (and customizable, along with other parameters, as explained below), at each change of direction, a full length can be added to the partial result. We can also estimate the length of a step by dividing, at each change of direction, the walked distance up to that point by the number of steps

taken since the test started or, to account for the variability of steps as the test goes on, the path length can be divided by the number of steps taken between the two most recent completed rotations. At the end of the six minutes, the remaining distance covered since the last rotation can be estimated by multiplying the number of steps taken after the rotation by the estimated step length.

This approach offers numerous benefits and means to better fit the algorithm to different environments and patients. Firstly, patients can wear the smartphone anywhere on their body according to their preference and comfort, as long as it is in a vertical position. Moreover, several customizable parameters can be set, such as path length, test duration, minimum angle needed to detect a rotation, maximum time window to detect a rotation, and the two thresholds used in the step-counting algorithm, whose value can be set manually or by performing a short calibration.

The calibration procedure must be performed at least once before the 6MWT but can be performed multiple times, independently from the tests, and the user is guided through the different stages. In order to perform the calibration, the patient has to wear a smartphone the same way they would during a 6MWT. After starting the calibration, the patient is asked to walk freely in the same conditions in which a 6MWT would take place (on a straight, flat path). During this stage, the patient can take as many steps as they wish, as long as they take at least 30. A higher amount would improve the accuracy of the step counter used in the 6MWT. The user should also count the number of steps or have somebody assisting them in order to type in an accurate estimate when prompted by the app at the end of this stage. After this, the new thresholds will be computed and shown to the user, and the app will ask whether to save or discard them. Internally, the calibration algorithm uses the signal recorded by the accelerometer during the free walk and the number of steps manually inserted to find the pair of thresholds that, if used on the signal by the same algorithm used during the 6MWT for step detection, would return the closest number of steps to the one typed in. The calibration algorithm tries different pairs of thresholds in a finite range, returning the best-performing pair as a result.

3.2.2. Ten-Meter Walking Test

Since no existing study proposed an algorithm compatible with our goals, we explored different approaches and found the following to be the best: using the smartphone camera, the STEP app can detect two colored plastic implements positioned against a wall and aligned with the three sections of the test (as shown in Figure 4b), to return, at the end of it, the time to complete the central portion of the path and, therefore, the walking speed.

The algorithm for this test works by detecting the right colors in the right place. Two items of the same color must be placed against a wall on the 2 and 8-m mark, one at each mark. To prevent the width of the items from interfering with the results, the item at the 2-m mark should be placed with one edge on the mark and the other one towards the start line, while the other item should be placed with one edge on the 8-m mark and the other one towards the finish line. The target color has to be chosen according to the different colors that can be found when looking in the direction of the wall so as to avoid the detection of false positives. The easiest way to accomplish this is to find a color that does not appear in the field of view and that is in contrast to any other color visible in that direction, cut two rectangular pieces out of a plastic sheet of that color, and hang them on a wall parallel to the walking direction.

For the 10MWT, the patient wears the smartphone with the camera facing out in a direction perpendicular to the walking path. The camera is used to record a short video from before the beginning of the test to after its end. The video will then be split into single frames and analyzed after the test. Each frame can be classified as either a *Frame containing the Target Color* (FTC) or a frame without it. Considering the duration of each frame is known, counting the number of frames without the target color between two FTCs enables us to compute how long it took the patient to walk from the first colored implement to the second one. In order to find which frames contain the target color, the algorithm

compares a subset of the pixels in each frame with the target color in the HSV color space. The colors are processed in this color space to prevent variable light conditions from having an effect on the comparison. Using the HSV color space allows us to set different values and thresholds for each dimension (hue, saturation, and value). A color is classified as the target color if the following constraints are met:

- The smallest distance between the current hue and the target one is below the hue threshold.
- The distance between the current saturation and the target one is below a saturation/value threshold.
- The distance between the current and target values is below a saturation/value threshold.

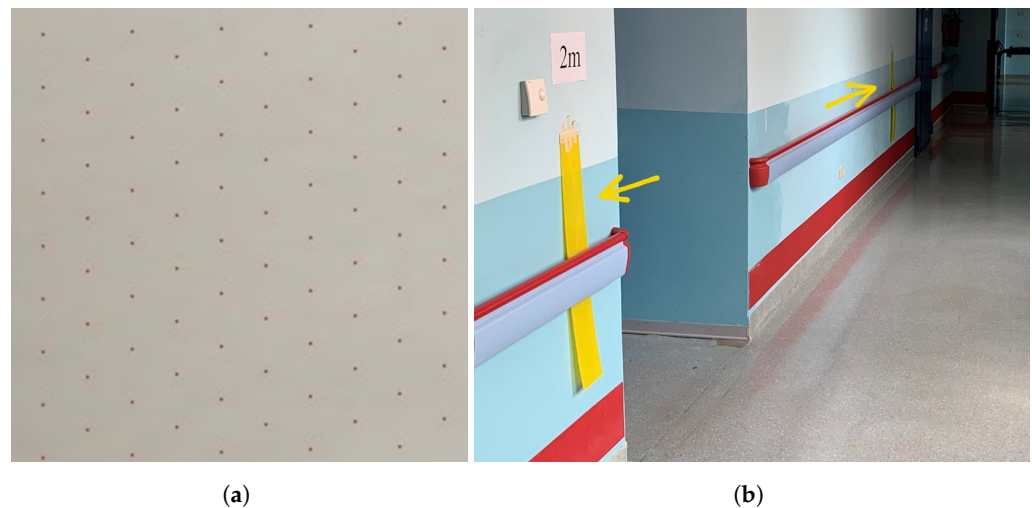


Figure 4. (a) The pattern in which a subset of the pixels in a frame are selected for comparison. (b) A hallway where patients performed the Ten-Meter Walking Test during the study. The two colored plastic implements are highlighted. Different lighting conditions around the implements are also visible.

In order to speed up the analysis at the end of the test, only some of the pixels in each frame are considered according to the size of the frame. Figure 4a shows the pattern used to select the pixels to compare. We found this sample to represent the entire frame since, if the colored implement is big enough and regularly shaped, it is unlikely to miss all of the pixels evaluated. Only one pixel has to match for the entire frame to be classified as FTC.

Each time a FTC is found, its timestamp in the video is estimated using the frame duration and its position in chronological order. The returned time between the two implements is the longest timestamp difference between two consecutive FTCs. The following points help explain why this method returns the correct time.

- When the patient starts the test at the 0 m mark, the target color is not in the camera's field of view.
- When the first FTC is recorded at the 2 m mark, there are no previous FTCs. If the test ended at this point, an error would be returned, because at least two FTCs are needed.
- When the following FTCs are recorded, they are all at a frame duration distance in time from each other. If the longest time between two consecutive FTCs is equal to a frame duration, an error is returned because, most likely, only one of the implements was detected during the test.
- When the patient is walking between the two implements, no FTCs should be recorded, and this is what increases the time between consecutive FTCs.
- When an FTC is re-recorded at the 8 m mark, it means the patient has reached the second implement. The returned result is the timestamp difference between this FTC and the previous one.
- After the 8 m mark, no more FTC should be recorded.

This method allows the patient to wear the smartphone anywhere on their body, as long as the camera is facing out towards the wall on which the colored implements are positioned, and no body parts cover the camera, even temporarily. A noteworthy benefit of this method is that, once the test is started, the patient can focus on performing at their best without having to interact with the device until after the end of the test. The app allows for the customization of a few parameters to improve the algorithm's reliability in different environments, such as path length, target color (which can be set using the camera), hue threshold, and saturation/value threshold.

3.2.3. Sit-to-Stand Test

One way to track repetitions during the 30STS involves monitoring the thigh's angle of movement. This method offers two benefits: it can distinguish between sitting and standing movements and help assess whether the range of motion during a repetition is sufficient to be considered valid.

Determining the angle of the thigh can be more challenging when the sensors are placed on different body parts. Likewise, attempting to detect repetitions through alternative data, like acceleration, poses similar difficulties. Placing the smartphone on the thigh, instead, allows for monitoring all the rotations experienced by the device, accurately determining the precise stage of the current repetition performed by the user. We have developed an algorithm that uses a gyroscope to do this.

In greater detail, the gyroscope is sampled at 30 Hz. Since the smartphone is worn lengthwise on the thigh (either flat in the front or on the side), the rotational values along the two axes orthogonal (perpendicular) to the thigh can be derived by employing the same processing approach explained in Section 3.2.1.

The patient can wear the smartphone in a pocket on the front or the side of the thigh and can also choose to start the test sitting or standing. When the test starts, the initial position is regarded as the baseline position. Therefore, any deviation from this will be regarded as part of a repetition. For each axis, an integral, which is equal to zero at the starting position, is updated over time by adding the angular velocity of each sample. Thus, by integrating angular velocity over time, we obtain angular displacement. This way, two opposite movements cancel each other out. After a new sample is obtained and the two integrals are updated, their absolute values are then added together to get the *Total Rotation* (TR). The algorithm monitors this value and compares it with different thresholds.

The entire movement is divided into four stages to ease the detection of a complete repetition. To better describe each stage, an example of a test where the patient starts in the sitting position will be used. As the patient stands up and moves through the range of motion, the TR increases until it reaches the *Repetition Threshold* (RT). At that point, the first half of the repetition is completed, and the algorithm moves from the first stage to the second one. In the second stage, the algorithm waits until a change of direction is detected, which happens when the patient reaches the standing position and reverses the motion.

The shift in direction is detected when the difference between the maximum value of TR reached during the repetition, and the current one is greater than the *Direction Threshold* (DT) set to 10° . When that happens, the algorithm moves on to stage three, sets the integrals and the TR back to zero, and adds to them the difference between the value they had at the direction change and the ones they have now, so as to not waste a portion of the repetition equal to the DT. The third stage mirrors the first one and waits until the RT is crossed again. This will happen when the patient is almost sitting; at this point, a complete repetition is counted before moving on to stage four. This last stage mirrors stage two and waits for a change in direction, at which point the algorithm will move on to the first stage of the new repetition.

This implementation allows patients to perform the test in two ways: maximum repetitions in a given time or minimum time to complete a given number of repetitions. Either way, the gyroscope and a timer keep track of repetitions and time. The algorithm uses a few parameters that can be customized to better fit different patients and environments,

such as test mode (for maximum reps or minimum time) with either the time or repetition limit and the RT.

3.3. Study

The application and its supporting system have been validated on real patients during their post-surgery rehabilitation. Individuals undergoing hip or knee replacement recovery, with their altered gait patterns, presented an excellent evaluation group for the STEP system. This enabled a direct comparison of the system's results with those obtained in person by physicians across diverse ranges and conditions.

3.3.1. Participants

A consecutive series of patients referred to the Physical and Rehabilitative Medicine Department of the Azienda Ospedaliera Nazionale SS. Antonio e Biagio e Cesare Arrigo, Alessandria, Italy, were assessed for eligibility from January 2023 to June 2023. To be eligible, the study participants had to meet the following inclusion criteria:

- A. Older people (over 60 years old), according to the definition by World Health Organization [66].
- B. Recent hip or knee replacement surgery within the past two weeks.
- C. Ability to walk without operator assistance.
- D. Ability to stand up from a chair without operator assistance.
- E. Ability to read and understand the Italian language.
- F. Read and signed the informed consent.

The exclusion criteria were:

- A. Dementia or cognitive impairment assessed with a Mini-Mental State Examination score < 24 [67].
- B. Pain at rest assessed with Numerical Rating Scale scoring below 4 points.
- C. History of metabolic, cardiovascular, pulmonary, neurological or other pathological comorbidities affecting physical performance.
- D. Anemia.
- E. Premorbid bed-bound.

The eligibility was assessed by a multidisciplinary team composed of an expert physician specializing in physical medicine and rehabilitation and a physiotherapist with years of expertise in orthopedic rehabilitation. The clinical trial protocol was performed in accordance with the Declaration of Helsinki [68] and pertinent National and International regulatory requirements. The study was approved by the Institutional Review Board of the Azienda Ospedaliera Nazionale SS. Antonio e Biagio e Cesare Arrigo, Alessandria, Italy (ASO.RRF.23.01; protocol number 14570). All the participants carefully read and signed a written informed consent before the study began.

3.3.2. Protocol

All patients were treated with a standard rehabilitation protocol. The 6MWT, the 10MWT and the 30STS were performed at baseline (T0), where a specialized physician also assessed useful information about the patients, and after a rehabilitation period (T1). All tests were overseen by a seasoned physiotherapist specializing in the comprehensive evaluation of patients with musculoskeletal conditions, who also collected the results personally by timing and measuring. These results constitute the *In-Person Assessments* (IPA) used as reference in the validation. All parameters were initially set, and calibrations were conducted before the tests, with provisions for additional calibrations at a later stage if required. Following each test, all *Results Returned by the App* (RRA) underwent immediate error checks to provide improved feedback and a comprehensive understanding of any potential malfunctions. Subsequently, these results were uploaded to the server. Following data collection for each test, validation involved a straightforward comparison to quantify

the agreement between the STEP app's results and the reference values. A smaller difference indicated stronger performance by the STEP app.

The 6MWT was performed following the American Thoracic Society guidelines [69]. In addition, the rate of perceived exertion (assessed by the Modified Borg Scale) and the peripheral capillary oxygen saturation were measured [70]. During the test, the patients wore both the smartphone (on the thigh) to compute the results and the Viatom CheckMe O2 to collect heart rate and oxygen saturation data during the test, as shown in Figure 5.

During the 10MWT, the patient's walk from the 2 m line to the 8 m line was recorded. Similarly, the app presented the duration rather than the speed, for direct comparison. Calculating the speed based on the time and known distance was considered straightforward and fell outside the scope of the validation. The walking speed (m/s) was determined by dividing the distance covered by the time taken to traverse it. The test was performed four times, two times at a comfortable speed (slow 10MWT) and two times at the fastest possible speed (fast 10MWT). The final results used for validation and analysis were the means of each pair of results [71].

The 30STS was performed using a chair with a seat height of 43.2 cm from the floor. The chair was positioned against a wall in order to prevent movements during the test. The patients had to complete as many repetitions as possible in 30 s. The score was given as the total number of full stands (body erect and straight) performed by the patient [10].

Lastly, the safety was recorded by the registration of adverse events, and the Global Perceived Effect scale was used to characterize patient satisfaction regarding this treatment [72].



Figure 5. A patient standing at one end of the 30 m hallway, getting ready to begin the Six-Minute Walking Test(6MWT). The patient is wearing a smartphone running the STEP app (A) and the CheckMe O2 device with the ring sensor (B).

4. Results

This section presents the data analysis collected from the study described in the previous section.

The data sample used for the analysis was composed of 15 females (51.7%) and 14 males (48.3%), characterized by a mean age of 70.0 ± 11.0 years and mean BMI of 27.1 ± 4.19 kg/m². Most of the patients included were surgically treated by total hip arthroplasty (21 patients; 72.4%), while seven patients (24.1%) were treated with knee arthroplasty and 1 with unicompartmental knee arthroplasty. The causes of the surgery were coxarthrosis (17; 58.6%), gonarthrosis (7; 24.1%), and osteonecrosis (1; 3.4%) or fracture (4; 13.8%). In order to validate the algorithms used by the STEP app, we performed a statistical comparison of IPA and RRA. In addition, we performed a statistical analysis of test results and medical records to find patterns or behaviors that show clinically relevant information.

4.1. Data Validation and Reliability

In this section, we present the validation results, performed by comparing RRA and IPA and including a correlation analysis. Then, we conclude this section by reporting the data reliability assessment. The sample size was estimated following the recommendation by Hobart et al. about validation studies [73]. As a result, the sample size of at least 25 patients was needed to have an alpha error of 0.05 and a power (1-Beta) of 0.85.

To validate the algorithm and ensure RRA could be considered clinically comparable to IPA, several metrics were calculated by comparing RRA with IPA to enhance performance quantification. Table 1 shows the most significant metrics among the ones computed for validation. These metrics were calculated in the following ways. The *Mean absolute difference* is the average of the absolute values of the differences between each RRA and IPA. The *Absolute difference standard deviation* is the standard deviation computed on the absolute values of the differences between each RRA and IPA. The *Smallest absolute difference* is the smallest value among the absolute values of the differences between each RRA and IPA. For each test, it is the error in the RRA closest to the IPA. The *Largest absolute difference* is the largest value among the absolute values of the differences between each RRA and IPA. In other words, for each test, it represents the error in the RRA furthest from the IPA. The *Mean percentage error* is the average of the absolute values of the differences between each RRA and IPA expressed as a percentage of the IPA. The relative RRA is subtracted from each IPA, and the absolute value of the result is computed. This is then expressed as a percentage of the IPA. The average of these percentages for each test is then calculated. These results are discussed in Section 5.

Table 1. Validation results for the Six-Minute Walking Test (6MWT), Ten-Meter Walking Test (10MWT), fast and slow, and 30-Second Sit-to-Stand Test (30STS).

Metric	6MWT (m)	10MWT Fast (s)	10MWT Slow (s)	30STS (reps)
<i>Mean absolute difference</i>	1.90	1.08	1.20	0.06
<i>Absolute difference standard deviation</i>	1.90	0.83	0.79	0.31
<i>Smallest absolute difference</i>	0	0.25	0.26	0
<i>Largest absolute difference</i>	10.09	4.34	4.78	2
<i>Mean percentage error</i>	1.35	11.55	9.72	0.67

[reps] = Repetitions.

To validate STEP app results further, we performed a correlation analysis for each test to assess the strength of the relationship. Pearson correlation coefficient (r) with 95% confidence intervals was used to assess the correlation between IPA and RRA, which was considered as poor ($r < 0.5$), moderate (r ranging between 0.5 to <0.75), good (r ranging between 0.75 to 0.9), or excellent ($r > 0.90$). A p -value lower than 0.05 was considered statistically significant. The correlation analysis showed strong positive relationships between RRA and IPA for all tests. Figure 6 shows the graphs representing the analysis results, also discussed below. In addition, for each test, physicians involved in the project provided data that further corroborated the previously observed consistency between app-collected data and their In-Person data, specifically at the patient improvement analysis level.

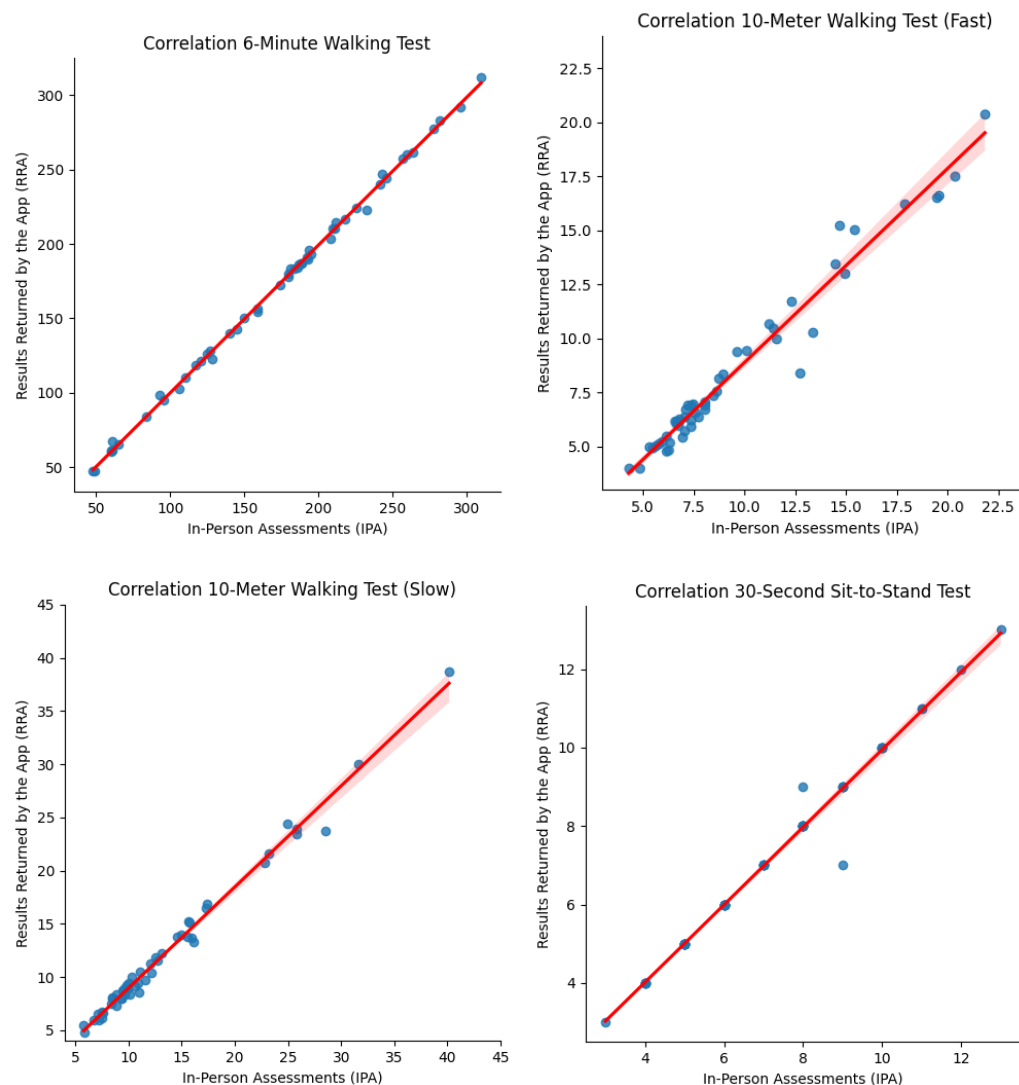


Figure 6. Correlation plots between In-Person Assessment (IPA) and Results Returned by the App (RRA) data.

An excellent correlation (r) was shown between 6MWT-IPA and 6MWT-RRA both at T0 ($r = 0.9981$; $p = 0.000067$) and T1 ($r = 0.9981$; $p = 0.000078$). Also, according to the physicians involved, an excellent correlation was presented in terms of distance improvement (T1-T0) measured after the rehabilitation treatments ($r = 0.9982$; $p = 1.248 \times 10^{-34}$).

In accordance, an excellent correlation was also found between fast 10MWT-IPA and fast 10MWT-RRA at both time points (T0: $r = 0.9769$; $p = 0.0021$; T1: $r = 0.9669$; $p = 6.148 \times 10^{-14}$). Concurrently, an excellent correlation was also confirmed in terms of velocity improvements (T1-T0) after the rehabilitation intervention ($r = 0.9336$; $p = 1.713 \times 10^{-20}$) as analyzed by physicians.

Moreover, a strong correlation was observed between slow 10MWT-IPA and slow 10MWT-RRA at both time points (T0: $r = 0.9943$; $p = 9.447 \times 10^{-37}$; T1: $r = 3.643 \times 10^{-15}$; $p < 0.0001$). Furthermore, the physician's analysis showed a significant correlation regarding velocity improvements (T1-T0) following the rehabilitation intervention ($r = 0.9966$; $p = 2.257 \times 10^{-23}$).

Lastly, the statistical correlation between 30STS-IPA and 30STS-RRA is elevated, as in the other tests at T0 ($r = 0.9821$, $p = 3.973 \times 10^{-21}$) and after the rehabilitation intervention ($r = 0.9960$, $p = 1.604 \times 10^{-23}$), indicating a strong and direct correlation between the two variables. In addition, based on the physicians' analysis, the improvements between T0 and T1 in terms of repetition number for both 30STS-IPA and 30STS-RRA were significantly

correlated ($r = 9952$; $p = 1.248 \times 10^{-34}$), indicating a clear and consistent relationship between the variables.

Regarding data reliability, we performed a test–retest procedure for each type of test to verify the consistency of repeated measurements. Unfortunately, patients could not accomplish this during the study because of the quick onset of fatigue, so the test–retest was performed by a healthy subject. Each type of test was performed five times by the same subject to achieve a consistent score on each occasion. For each execution, both RRA and IPA were collected. Considering the controlled execution environment, it is unsurprising that the errors observed comparing RRA and IPA were minimal, negligible, and consistently smaller than those encountered during the study. We focus on measurement variability rather than discussing the errors already addressed during validation. The inherent variability in test execution prevents results from being consistently identical.

Although minor behavioral changes by the subject during testing can influence the outcome, it is crucial to recognize that these alterations impact both IPA and RRA equally. Therefore, even without multiple identical test results for direct comparison, we can analyze the test–retest data to determine whether the differences observed in RRA data align with those in the IPA data, considering the ground truth results. For this analysis, we compare the *standard deviation* of RRA with that of IPA, computed separately for each type of test and reported in Table 2.

Table 2. Standard deviation for the Six-Minute Walking Test (6MWT), Ten-Meter Walking Test (10MWT), and 30-s Sit-to-Stand Test (30STS) for both measuring methods, computed from the results collected during a test–retest procedure.

Measurement Type	6MWT (m)	10MWT (s)	30STS (reps)
IPA	4.02	0.15	0.84
RRA	5.9	0.15	0.84

[IPA] = In-Person Assessments, [RRA] = Results Returned by the App, [reps] = repetitions.

The results demonstrate high comparability between RRA and IPA, showing consistent measurements by the app. In more detail, the 6MWT showed a higher variability, which can be attributed to the longer test duration affecting both RRA and IPA. The 10MWT, on the contrary, is a short test and shows little variability among repeated executions. Finally, while the 30STS exhibits some variability, both measurement types share the same standard deviation value because all results in the test–retest data were identical.

Finally, we compared the measurements from both the accelerometer and gyroscope sensors on the smartphone with the ones collected by the same type of sensors on another device. Specifically, we used the B-L475E-IOT01A IoT node by STMicroelectronics [74], a microcontroller board featuring an ARM Cortex-M4-based STM32L4 low-power MCU and, among other components, an accelerometer and a gyroscope (LSM6DSL). We assessed the consistency over time by comparing data collected at two different points under identical conditions, and we observed a high correlation, indicating strong reliability.

4.2. Additional Data Analysis: Exploring Correlations between Patient Data and App Data

This section delves into the application’s broader potential by exploring patient and app data correlations. This analysis represents an exploratory extension of the initial project scope, leveraging the application’s data collection capabilities. The correlation analysis was performed using linear regression. Linear regression is a fundamental statistical method used to uncover and measure connections among multiple factors [75]. By fitting a straight line to a collection of data points, this approach enables us to discern the underlying trend and generate forecasts based on this pattern. When applied as a knowledge tool in healthcare treatments, linear regression results provide a data-driven foundation for tailoring interventions and optimizing patient outcomes [76]. This analytical approach could help healthcare operators fine-tune treatment strategies to improve patient well-being.

The STEP app was validated in a real-world setting with data collected from actual patients. Therefore, these data could provide clinically relevant information. Consequently, we use linear regression to investigate whether any RRAs correlate with the patient's physical factors (e.g., age, surgeries). *Backward elimination* was used to select the patient physical factors most strongly associated with each sub-maximal test. Backward elimination is a statistical method for selecting variables in a *multiple-regression model* [77]. It starts with all of the model's variables and then removes them one by one until the remaining variables are statistically significant [78]. For each test, we performed a multiple linear regression with *each test* (e.g., 6MWT) as the dependent variable and *age, weight, height, diagnosis, surgery, and body mass index (BMI)* as the independent variables.

Table 3 presents an example of the results of a linear regression analysis applied to the 6MWT. The listed variables are those that, based on their *p*-values, are statistically significantly associated with the test (*p*-value ≤ 0.05). From a statistical perspective, this indicates that the 6MWT test can be affected by the patient's medical characteristics, such as the type of diagnosis. However, this relationship is weak, as the standard error exceeds the commonly used range of 10% of the coefficient value. Therefore, it is important to clarify that the relationships between the data we will present below are only possible interpretations that require confirmation with more data collected by the application.

Table 3. Linear Regression Six-Minute Walking Test.

Variable	Coeff	Std. Error	t-Ratio	p-Ratio
Const	0.36	0.21	1.71	0.1008
Age	-0.73	0.29	-2.57	0.0176
Overweight	0.28	0.13	2.14	0.0435
Normalweight	0.49	0.13	3.72	0.0012
Gonarthrosis	0.51	0.16	3.14	0.0047
Coxarthrosis	0.36	0.15	2.39	0.0256

According to the model presented in Table 3, the 6MWT score decreases by 0.73 points for each additional year of patient age. The relationship between the 6MWT score and the other categorical variables is diverse, depending on the presence or absence of each variable. For example, a patient with a *normal weight* has a 6MWT score of 0.21 points higher than someone *overweight*. Similarly, a patient diagnosed with *gonarthrosis* has a 6MWT score of 0.15 points higher than someone with *coxarthrosis*.

The above analysis was also performed using all data obtained from the other tests. However, only one independent variable had a statistically significant value compared to the rest of the tests. Table 4 shows the related independent variable that obtained the minimum *p*-value for each test and the corresponding correlation coefficient.

Table 4. Independent variables statistically relevant for each *test*.

Test	Independent Variable	p-Value	Coefficient
6MWT1	[BMI] Normal Weight	0.0012	0.49
6MWT2	[BMI] Overweight	0.2024	-0.13
10MWT1F	[D] Fracture	0.0948	0.30
10MWT2F	[BMI] Overweight	0.0871	0.16
10MWT1S	[D] Fracture	0.0053	0.38
10MWT2S	[BMI] Normal Weight	0.0529	-0.18
30STS1	[BMI] Normal Weight	0.0148	0.25
30STS2	[BMI] Normal Weight	0.0368	0.28

[BMI] = body mass index, [D] = diagnosis.

In general, the results of the analysis suggest a notable association between body mass index (BMI) and results from most submaximal tests. This is plausible, given that a person's morbidity can influence the results of these tests [79]. This type of information obtained

from the STEP application could be helpful for clinicians, as it allows for the adaptation of tests to be better aligned with or more tailored to the specific characteristics of patients.

5. Discussion

The STEP application demonstrated its ability to automate the tests reliably and accurately, emphasizing its potential as a dependable and valuable tool. As shown in the Results section, the study revealed compelling results in reliability, validation, and correlation between Results Returned by the App and In-Person Assessments.

While several studies concentrate on healthy individuals or patients with cardiopulmonary diseases (both usually exhibiting standard gait patterns), our study was conducted on patients surgically treated for total hip or knee arthroplasty. The nature of the tests makes them harder to perform correctly when the patients are injured or recovering after surgery. Despite these challenges, the application demonstrated remarkable efficacy across this heterogeneous patient cohort, providing further evidence for the robustness and reliability of the implemented algorithms.

A comprehensive review of the relevant literature revealed no prior studies validating the use of wearable digital tools for gait assessment in patients undergoing total hip or knee arthroplasty. This suggests that the current investigation is the pioneer of such validation efforts. Although this can be seen as a merit of the study, more patient diversity is needed to generalize the results to other populations.

The study faced challenges due to the patients' use of walking aids. Some patients switched aids between trials, significantly impacting the gait signal. This necessitated recalibration to optimize the 6MWT algorithm. For example, a patient using a walker initially exhibited a low error (0.23 m) which increased (4.41 m) after switching to crutches for the second trial without recalibration. This highlights the impact of walking aids on algorithm performance. Implementing recalibration after any aid change improved outcomes. The process customizes the parameters for each individual, enhancing resilience in suboptimal conditions. A recent case with comparable conditions, which included a recalibration before the second trial, showed errors of 0.71 m and 0.03 m despite using different aids in the two trials.

The validation of the data obtained automatically by the app showed significant consistency with those obtained by the physicians. Beyond initial validation, the quality of the results can be further verified by comparing them with the minimum detectable change of each test reported in the existing literature. Compared to the minimal detectable changes reported in the literature (e.g., 33.5–82 m for 6MWT [80,81], three repetitions for 30STS [82], 0.22–0.23 m/s for TMT [83]), our mean absolute errors were significantly lower (1.9 m, 0.06 repetitions, 0.06 m/s, 0.1 m/s), indicating high agreement with physician measurements. No clinically significant errors were found.

Regarding the STEP limitations, the first one is related to the small number of patients, which made it harder to obtain meaningful results from the data analysis. After collecting a larger dataset, we will be able to explore novel analysis techniques and leverage machine learning algorithms. This comprehensive practice has the potential to augment the established analyses outlined in the previous section, potentially leading to the generation of more statistically robust and dependable findings.

A potential source of discrepancy in STEP 6MWT results is the inherent limitation of visual assessments using hallway markers for the collection of the In-Person Assessments used as reference. Since these measurements are rounded to the nearest meter, any decimal value generated by the app appears as an error. Future validation efforts could incorporate measuring wheels for control measurements to address this limitation and achieve a more accurate comparison.

Our validation process highlighted the Ten-Meter Walking Test (10MWT) as the area with the most significant potential for improvement. While the underlying algorithm demonstrates reliability, its usability could be enhanced. The results indicate that error rates increase with test duration, with the app consistently underestimating the actual walking

time. These discrepancies can be attributed to factors such as excessive patient-to-wall distance or improper smartphone placement. Under these conditions, the initial colored marker may appear in later frames, even after the patient has crossed the 2 m line. At the same time, the second marker may appear earlier. Consequently, the final frame with the first marker and the initial frame with the second marker appear closer in time than the actual test duration. Despite these challenges, the results remain sufficiently close to control measurements, reaffirming the algorithm's overall reliability.

Ensuring accurate distance measurement to set up the environment remains a challenge, especially in home use scenarios, and the current 10MWT implementation requires a one-time setup that could be complex for some users. Involving a relative or caregiver in the initial setup process can significantly alleviate this hurdle. Following this initial assistance, patients should be able to perform subsequent tests independently. Future iterations of the app could explore visual cues or on-screen instructions to simplify the setup process further or establish a reliable method that involves an easier setup. Moreover, the space required to perform these tests might limit the widespread adoption of this technology. Lastly, the app's suitability for independent home use by patients depends on individual circumstances. Physicians should evaluate each patient to determine if the app is appropriate for remote monitoring.

This study identified areas for improvement that will be addressed in future research. To ensure robust and statistically significant data, we will test the app across diverse demographic cohorts and target a broader patient population in our recruitment efforts. These future studies, along with our results, will serve two purposes: application improvement and suitability verification for home-based healthcare settings, fulfilling the project's initial objectives.

Considering the application's utility as a telemedicine tool, our future endeavors involve the integration of additional features to further enhance its capability in accurately assessing the user's health status. To this end, new tests, including balance assessments, could be automated. Moreover, the application of sensors during these tests could facilitate the extraction of additional features, such as vertical acceleration in the 30STS or gait analysis in the 6MWT, in addition to their primary role in computing and delivering the test results.

Upon the system's successful validation, physicians and patients can extend this application beyond routine clinical usage. Its utility in research studies could facilitate seamless data collection of physical performance and its multilevel interactions with patient characteristics, diet, and disease severity, which could have crucial implications in rehabilitation management.

6. Conclusions

Assessing physical performance is crucial for rehabilitation, but frequent, accurate, and efficient monitoring remains challenging. While technology offers solutions, ensuring reliable test results can be difficult. This paper presented the System for Tracking and Evaluating Performance (STEP), a telemedicine tool developed to automate submaximal tests, minimizing the need for assistance and expert supervision in rehabilitation protocols. STEP integrates a smartphone application leveraging built-in sensors for automated test execution and a server infrastructure for data storage and visualization, streamlining the process for patients and healthcare providers. The validation involved a comparative study, where patients performed tests using the app while an expert monitored and manually recorded results. STEP demonstrated high accuracy, consistently producing minimal errors below clinically significant thresholds. These findings suggest STEP's potential to significantly improve test accessibility for patients, particularly those with mobility limitations, and potentially reduce healthcare resource utilization. Further research is warranted to confirm these benefits in larger and more diverse populations. Additionally, future efforts will focus on simplifying setup and execution procedures.

Overall, this study reinforces existing literature on remote physical performance assessment challenges while introducing a novel, validated solution with promising implications for future research and improved rehabilitation monitoring.

Author Contributions: Conceptualization, M.C., F.D., M.I. and L.L.; Data curation, F.D. and L.L.; Formal analysis, L.B.-L., F.D. and L.L.; Investigation, F.D. and L.L.; Methodology, L.L.; Project administration, M.C., A.d.S. and M.I.; Resources, L.L.; Software, F.D.; Supervision, M.C. and L.L.; Validation, L.B.-L. and F.D.; Visualization, L.B.-L. and F.D.; Writing—original draft, L.B.-L., M.C., F.D., A.d.S., M.I. and L.L.; Writing—review and editing, L.B.-L., M.C., F.D. and L.L. All authors have read and agreed to the published version of the manuscript.

Funding: This manuscript is part of the project NODES, which has received funding from the MUR—M4C2 1.5 of PNRR funded by the European Union-NextGenerationEU (Grant agreement no. ECS00000036).

Institutional Review Board Statement: The study was approved by the Ethics Committee of Comitato Etico Interaziendale A.O. “SS. Antonio e Biagio e Cesare Arrigo” di Alessandria (Approval Code: ASO.RRF.23.01 and Approval Date: 5 June 2023).

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Beard, J.R.; Officer, A.; De Carvalho, I.A.; Sadana, R.; Pot, A.M.; Michel, J.P.; Lloyd-Sherlock, P.; Epping-Jordan, J.E.; Peeters, G.G.; Mahanani, W.R.; et al. The World report on ageing and health: A policy framework for healthy ageing. *Lancet* **2016**, *387*, 2145–2154. [[CrossRef](#)] [[PubMed](#)]
2. Lutz, W.; Sanderson, W.; Scherbov, S. The coming acceleration of global population ageing. *Nature* **2008**, *451*, 716–719. [[CrossRef](#)] [[PubMed](#)]
3. Kennedy, B.K.; Berger, S.L.; Brunet, A.; Campisi, J.; Cuervo, A.M.; Epel, E.S.; Franceschi, C.; Lithgow, G.J.; Morimoto, R.I.; Pessin, J.E.; et al. Geroscience: Linking aging to chronic disease. *Cell* **2014**, *159*, 709–713. [[CrossRef](#)] [[PubMed](#)]
4. Lippi, L.; de Sire, A.; Mezian, K.; Curci, C.; Perrero, L.; Turco, A.; Andaloro, S.; Ammendolia, A.; Fusco, N.; Invernizzi, M. Impact of exercise training on muscle mitochondria modifications in older adults: A systematic review of randomized controlled trials. *Aging Clin. Exp. Res.* **2022**, *34*, 1495–1510. [[CrossRef](#)]
5. Lippi, L.; Uberti, F.; Folli, A.; Turco, A.; Curci, C.; d’Abrosca, F.; de Sire, A.; Invernizzi, M. Impact of nutraceuticals and dietary supplements on mitochondria modifications in healthy aging: A systematic review of randomized controlled trials. *Aging Clin. Exp. Res.* **2022**, *34*, 2659–2674. [[CrossRef](#)]
6. Moggio, L.; de Sire, A.; Marotta, N.; Demeco, A.; Ammendolia, A. Exoskeleton versus end-effector robot-assisted therapy for finger-hand motor recovery in stroke survivors: Systematic review and meta-analysis. *Top. Stroke Rehabil.* **2022**, *29*, 539–550. [[CrossRef](#)]
7. Noonan, V.; Dean, E. Submaximal exercise testing: Clinical application and interpretation. *Phys. Ther.* **2000**, *80*, 782–807. [[CrossRef](#)]
8. Moore, J.L.; Potter, K.; Blankshain, K.; Kaplan, S.L.; O’Dwyer, L.C.; Sullivan, J.E. A core set of outcome measures for adults with neurologic conditions undergoing rehabilitation: A clinical practice guideline. *J. Neurol. Phys. Ther.* **2018**, *42*, 174. [[CrossRef](#)]
9. Society, A.T. ATS Statement: Guideline for the six-minute walk test. ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. *J. Respir. Crit. Care Med.* **2002**, *166*.
10. Jones, C.J.; Rikli, R.E.; Beam, W.C. A 30-s Chair-Stand Test as a Measure of Lower Body Strength in Community-Residing Older Adults. *Res. Q. Exerc. Sport* **1999**, *70*, 113–119.
11. Nilsagård, Y.; Andreasson, M.; Carling, A.; Vesterlin, H. Examining the validity and sensitivity to change of the 5 and 10 sit-to-stand tests in people with multiple sclerosis. *Physiother. Res. Int. J. Res. Clin. Phys. Ther.* **2017**, *22*, e1681. [[CrossRef](#)] [[PubMed](#)]
12. Møller, A.B.; Bibby, B.M.; Skjerbæk, A.G.; Jensen, E.; Sørensen, H.; Stenager, E.; Dalgas, U. Validity and variability of the 5-repetition sit-to-stand test in patients with multiple sclerosis. *Disabil. Rehabil.* **2012**, *34*, 2251–2258.
13. Csuka, M.; McCarty, D.J. Simple method for measurement of lower extremity muscle strength. *Am. J. Med.* **1985**, *78*, 77–81. [[CrossRef](#)]
14. Maud, P.J.; Shultz, B.B. Norms for the Wingate anaerobic test with comparison to another similar test. *Res. Q. Exerc. Sport* **1989**, *60*, 144–151. [[CrossRef](#)] [[PubMed](#)]
15. Powell, D.J.; Lioffi, C.; Schlotz, W.; Moss-Morris, R. Tracking daily fatigue fluctuations in multiple sclerosis: Ecological momentary assessment provides unique insights. *J. Behav. Med.* **2017**, *40*, 772–783. [[CrossRef](#)] [[PubMed](#)]

16. Kasser, S.L.; Goldstein, A.; Wood, P.K.; Sibold, J. Symptom variability, affect and physical activity in ambulatory persons with multiple sclerosis: Understanding patterns and time-bound relationships. *Disabil. Health J.* **2017**, *10*, 207–213. [[CrossRef](#)] [[PubMed](#)]
17. Takayanagi, N.; Sudo, M.; Yamashiro, Y.; Lee, S.; Kobayashi, Y.; Niki, Y.; Shimada, H. Relationship between daily and in-laboratory gait speed among healthy community-dwelling older adults. *Sci. Rep.* **2019**, *9*, 3496. [[CrossRef](#)]
18. Hillel, I.; Gazit, E.; Nieuwboer, A.; Avanzino, L.; Rochester, L.; Cereatti, A.; Croce, U.D.; Rikkert, M.O.; Bloem, B.R.; Pelosin, E.; et al. Is every-day walking in older adults more analogous to dual-task walking or to usual walking? Elucidating the gaps between gait performance in the lab and during 24/7 monitoring. *Eur. Rev. Aging Phys. Act.* **2019**, *16*, 1–12. [[CrossRef](#)]
19. Giannouli, E.; Fillekes, M.P.; Mellone, S.; Weibel, R.; Bock, O.; Zijlstra, W. Predictors of real-life mobility in community-dwelling older adults: An exploration based on a comprehensive framework for analyzing mobility. *Eur. Rev. Aging Phys. Act.* **2019**, *16*, 1–13. [[CrossRef](#)]
20. Prince, S.; Adamo, K.; Hamel, M.; Hardt, J.; Gorber, S.; Tremblay, M. A Comparison of direct versus self-report measures for assessing physical activity in adults: A systematic review. *Int. J. Behav. Nutr. Phys. Act.* **2008**, *5*, 56. [[CrossRef](#)]
21. Harris, T.; Owen, C.; Victor, C.; Adams, R.; Ekelund, U.; Cook, D. A Comparison of Questionnaire, Accelerometer, and Pedometer. *Med. Sci. Sport. Exerc.* **2009**, *41*, 1392–1402. [[CrossRef](#)] [[PubMed](#)]
22. Bobholz, J.; Rao, S. Cognitive dysfunction in multiple sclerosis: A review of recent developments. *Curr. Opin. Neurol.* **2003**, *16*, 283–288. [[CrossRef](#)] [[PubMed](#)]
23. Macko, R.F.; Haeuber, E.; Shaughnessy, M.; Coleman, K.L.; Boone, D.A.; Smith, G.V.; Silver, K.H. Microprocessor-based ambulatory activity monitoring in stroke patients. *Med. Sci. Sport. Exerc.* **2002**, *34*, 394–399. [[CrossRef](#)] [[PubMed](#)]
24. Coughlin, S.S. Recall bias in epidemiologic studies. *J. Clin. Epidemiol.* **1990**, *43*, 87–91. [[CrossRef](#)] [[PubMed](#)]
25. de Sire, A.; Marotta, N.; Agostini, F.; Drago Ferrante, V.; Demeco, A.; Ferrillo, M.; Inzitari, M.T.; Pellegrino, R.; Russo, I.; Ozyemisci Taskiran, O.; et al. A telerehabilitation approach to chronic facial paralysis in the COVID-19 pandemic scenario: What role for electromyography assessment? *J. Pers. Med.* **2022**, *12*, 497. [[CrossRef](#)]
26. Turolla, A.; Rossetini, G.; Viceconti, A.; Palese, A.; Geri, T. Musculoskeletal physical therapy during the COVID-19 pandemic: Is telerehabilitation the answer? *Phys. Ther.* **2020**, *100*, 1260–1264. [[CrossRef](#)]
27. Scherrenberg, M.; Wilhelm, M.; Hansen, D.; Völler, H.; Cornelissen, V.; Frederix, I.; Kemps, H.; Dendale, P. The future is now: A call for action for cardiac telerehabilitation in the COVID-19 pandemic from the secondary prevention and rehabilitation section of the European Association of Preventive Cardiology. *Eur. J. Prev. Cardiol.* **2021**, *28*, 524–540. [[CrossRef](#)]
28. Lippi, L.; Turco, A.; Folli, A.; D’Abrosca, F.; Curci, C.; Mezzan, K.; de Sire, A.; Invernizzi, M. Technological advances and digital solutions to improve quality of life in older adults with chronic obstructive pulmonary disease: A systematic review. *Aging Clin. Exp. Res.* **2023**, *35*, 953–968. [[CrossRef](#)]
29. Moore, S.; Jayewardene, D. The use of smartphones in clinical practice. *Nurs. Manag.* **2014**, *21*, 18–22. [[CrossRef](#)]
30. Marotta, N.; Calafiore, D.; Curci, C.; Lippi, L.; Ammendolia, V.; Ferraro, F.; Invernizzi, M.; de Sire, A. Integrating virtual reality and exergaming in cognitive rehabilitation of patients with Parkinson disease: A systematic review of randomized controlled trials. *Eur. J. Phys. Rehabil. Med.* **2022**, *58*, 818. [[CrossRef](#)]
31. Iovanel, G.; Ayers, D.; Zheng, H. The Role of Wearable Technology in Measuring and Supporting Patient Outcomes Following Total Joint Replacement: Review of the Literature. *JMIR Perioper. Med.* **2023**, *6*, e39396. [[CrossRef](#)]
32. Van Lummel, R.C.; Walgaard, S.; Maier, A.B.; Ainsworth, E.; Beek, P.J.; van Dieën, J.H. The instrumented sit-to-stand test (iSTS) has greater clinical relevance than the manually recorded sit-to-stand test in older adults. *PLoS ONE* **2016**, *11*, e0157968. [[CrossRef](#)]
33. Reissman, M.E.; Gordon, K.E.; Dhaher, Y.Y. Manipulating post-stroke gait: Exploiting aberrant kinematics. *J. Biomech.* **2018**, *67*, 129–136. [[CrossRef](#)] [[PubMed](#)]
34. Tulipani, L.J.; Meyer, B.; Allen, D.; Solomon, A.J.; McGinnis, R.S. Evaluation of unsupervised 30-second chair stand test performance assessed by wearable sensors to predict fall status in multiple sclerosis. *Gait Posture* **2022**, *94*, 19–25. [[CrossRef](#)]
35. Bäcker, H.C.; Wu, C.H.; Pfürringer, D.; Petersen, W.; Stöckle, U.; Braun, K.F. A review of functional outcomes after the app-based rehabilitation of patients with TKA and THA. *J. Pers. Med.* **2022**, *12*, 1342. [[CrossRef](#)] [[PubMed](#)]
36. Ozdalga, E.; Ozdalga, A.; Ahuja, N. The smartphone in medicine: A review of current and potential use among physicians and students. *J. Med Internet Res.* **2012**, *14*, e1994. [[CrossRef](#)]
37. Ponciano, V.; Pires, I.M.; Ribeiro, F.R.; Spinsante, S. Sensors are capable to help in the measurement of the results of the timed-up and go test? a systematic review. *J. Med Syst.* **2020**, *44*, 1–16. [[CrossRef](#)] [[PubMed](#)]
38. Pires, I.M.; Denysyuk, H.V.; Villasana, M.V.; Sá, J.; Marques, D.L.; Morgado, J.F.; Albuquerque, C.; Zdravetski, E. Development technologies for the monitoring of six-minute walk test: A systematic review. *Sensors* **2022**, *22*, 581. [[CrossRef](#)]
39. Sabour, S.; Ghassemi, F. Validity of Submaximal Step Tests to Estimate Maximal Oxygen Uptake in Healthy Adults. *Sport. Med.* **2016**, *46*, 1381–1392. [[CrossRef](#)]
40. Shirley Ryan AbilityLab. 10 Meter Walk Test. 2014. Available online: <https://www.sralab.org/rehabilitation-measures/10-meter-walk-test> (accessed on 28 August 2024).
41. Smith-Turchyn, J.; Adams, S.C.; Sabiston, C.M. Testing of a self-administered 6-minute walk test using technology: Usability, reliability and validity study. *JMIR Rehabil. Assist. Technol.* **2021**, *8*, e22818. [[CrossRef](#)]

42. Jesus, M.O.d.; Ostolin, T.L.V.D.P.; Proença, N.L.; Silva, R.P.d.; Dourado, V.Z. Self-administered six-minute walk test using a free smartphone app in asymptomatic adults: Reliability and reproducibility. *Int. J. Environ. Res. Public Health* **2022**, *19*, 1118. [CrossRef] [PubMed]
43. Cheng, Q.; Juen, J.; Li, Y.; Prieto-Centurion, V.; Krishnan, J.A.; Schatz, B.R. GaitTrack: Health monitoring of body motion from spatio-temporal parameters of simple smart phones. In Proceedings of the International Conference on Bioinformatics, Computational Biology and Biomedical Informatics, Honolulu, HI, USA, 4–6 March 2013; pp. 897–906.
44. Juen, J.; Cheng, Q.; Prieto-Centurion, V.; Krishnan, J.A.; Schatz, B. Health monitors for chronic disease by gait analysis with mobile phones. *Telemed. E-Health* **2014**, *20*, 1035–1041. [CrossRef]
45. Brooks, G.C.; Vittinghoff, E.; Iyer, S.; Tandon, D.; Kuhar, P.; Madsen, K.A.; Marcus, G.M.; Pletcher, M.J.; Olgin, J.E. Accuracy and usability of a self-administered 6-minute walk test smartphone application. *Circ. Heart Fail.* **2015**, *8*, 905–913. [CrossRef] [PubMed]
46. Borg, G. *Borg's Perceived Exertion and Pain Scales*; Human Kinetics: Champaign, IL, USA, 1998.
47. Borg, G.A. Psychophysical bases of perceived exertion. *Med. Sci. Sport. Exerc.* **1982**, *14*, 377–381. [CrossRef]
48. Scherrenberg, M.; Bonneux, C.; Yousif Mahmood, D.; Hansen, D.; Dendale, P.; Coninx, K. A Mobile application to perform the six-minute walk test (6MWT) at home: A random walk in the park is as accurate as a standardized 6MWT. *Sensors* **2022**, *22*, 4277. [CrossRef] [PubMed]
49. Capela, N.A.; Lemaire, E.D.; Baddour, N. Novel algorithm for a smartphone-based 6-minute walk test application: Algorithm, application development, and evaluation. *J. Neuroeng. Rehabil.* **2015**, *12*, 1–13. [CrossRef] [PubMed]
50. Salvi, D.; Poffley, E.; Orchard, E.; Tarassenko, L. The mobile-based 6-minute walk test: Usability study and algorithm development and validation. *JMIR MHealth UHealth* **2020**, *8*, e13756. [CrossRef]
51. Salbach, N.M.; O'Brien, K.K.; Brooks, D.; Irvin, E.; Martino, R.; Takhar, P.; Chan, S.; Howe, J.A. Reference values for standardized tests of walking speed and distance: A systematic review. *Gait Posture* **2015**, *41*, 341–360. [CrossRef]
52. Clavijo-Buendia, S.; Molina-Rueda, F.; Martin-Casas, P.; Ortega-Bastidas, P.; Monge-Pereira, E.; Laguarda-Val, S.; Morales-Cabezas, M.; Cano-de-la Cuerda, R. Construct validity and test-retest reliability of a free mobile application for spatio-temporal gait analysis in Parkinson's disease patients. *Gait Posture* **2020**, *79*, 86–91. [CrossRef]
53. Saito, Y.; Nakamura, S.; Tanaka, A.; Watanabe, R.; Narimatsu, H.; Chung, U.i. Evaluation of the validity and reliability of the 10-meter walk test using a smartphone application among Japanese older adults. *Front. Sport. Act. Living* **2022**, *4*, 904924. [CrossRef]
54. Marques, D.L.; Neiva, H.P.; Pires, I.M.; Zdravevski, E.; Mihajlov, M.; Garcia, N.M.; Ruiz-Cárdenas, J.D.; Marinho, D.A.; Marques, M.C. An experimental study on the validity and reliability of a smartphone application to acquire temporal variables during the single sit-to-stand test with older adults. *Sensors* **2021**, *21*, 2050. [CrossRef]
55. Marques, D.L.; Neiva, H.P.; Pires, I.M.; Marinho, D.A.; Marques, M.C. Accelerometer data from the performance of sit-to-stand test by elderly people. *Data Brief* **2020**, *33*, 106328. [CrossRef]
56. Cerrito, A.; Bichsel, L.; Radlinger, L.; Schmid, S. Reliability and validity of a smartphone-based application for the quantification of the sit-to-stand movement in healthy seniors. *Gait Posture* **2015**, *41*, 409–413. [CrossRef] [PubMed]
57. Teixeira, E.; Bohn, L.; Guimarães, J.P.; Marques-Aleixo, I. Portable digital monitoring system for sarcopenia screening and diagnosis. *Geriatrics* **2022**, *7*, 121. [CrossRef] [PubMed]
58. Montemurro, A.; Ruiz-Cárdenas, J.D.; Martínez-García, M.d.M.; Rodríguez-Juan, J.J. Validity of an iPhone App to Detect Prefrailty and Sarcopenia Syndromes in Community-Dwelling Older Adults: The Protocol for a Diagnostic Accuracy Study. *Sensors* **2022**, *22*, 6010. [CrossRef] [PubMed]
59. Jovanov, E.; Wright, S.; Ganegoda, H. Development of an Automated 30 Seconds Chair Stand Test Using Smartwatch Application. In Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, 23–27 July 2019; pp. 2474–2477.
60. Bluetooth® Technology Website. GATT. 2023. Available online: <https://www.bluetooth.com/specifications/specs/gatt-specification-supplement/> (accessed on 28 August 2024).
61. Sathesh, M.; D'mello, B.J.; Krol, J. *Web Development with MongoDB and Nodejs*; Packt Publishing Ltd.: Birmingham, UK, 2015.
62. Keahey, K.; Anderson, J.; Zhen, Z.; Riteau, P.; Ruth, P.; Stanzione, D.; Cevik, M.; Colleran, J.; Gunawi, H.S.; Hammock, C.; et al. Lessons Learned from the Chameleon Testbed. In Proceedings of the 2020 USENIX Annual Technical Conference (USENIX ATC '20), Online, 15–17 July 2020; USENIX Association: Berkeley, CA, USA, 2020.
63. OpenJS Foundation Website. Express. 2023. Available online: <https://expressjs.com/> (accessed on 28 August 2024).
64. Moroney, L.; Moroney, L. Using authentication in firebase. In *The Definitive Guide to Firebase: Build Android Apps on Google's Mobile Platform*; Apress: Berkeley, CA, USA, 2017; pp. 25–50.
65. Šušter, I.; Ranisavljević, T. Optimization of MySQL database. *J. Process Manag. New Technol.* **2023**, *11*, 141–151. [CrossRef]
66. World Health Organization. *Decade of Healthy Ageing: Baseline Report*; World Health Organization: Geneva, Switzerland, 2020. Available online: <https://www.who.int/publications/i/item/9789240017900> (accessed on 28 August 2024).
67. Kurlowicz, L.; Wallace, M. The mini-mental state examination (MMSE). *J. Gerontol. Nurs.* **1999**, *25*, 8–9. [CrossRef]
68. World Medical Association. World Medical Association Declaration of Helsinki. Ethical principles for medical research involving human subjects. *Bull. World Health Organ.* **2001**, *79*, 373.

69. ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: Guidelines for the six-minute walk test. *Am. J. Respir. Crit. Care Med.* **2002**, *166*, 111–117. [CrossRef]
70. Mahler, D.A.; Horowitz, M.B. Perception of breathlessness during exercise in patients with respiratory disease. *Med. Sci. Sport. Exerc.* **1994**, *26*, 1078–1081. [CrossRef]
71. Amatachaya, S.; Kwanmongkolthong, M.; Thongjumroon, A.; Boonpew, N.; Amatachaya, P.; Saensook, W.; Thaweewannakij, T.; Hunsawong, T. Influence of timing protocols and distance covered on the outcomes of the 10-meter walk test. *Physiother. Theory Pract.* **2020**, *36*, 1348–1353. [CrossRef]
72. Meisingset, I.; Stensdotter, A.; Woodhouse, A.; Vasseljen, O. Predictors for global perceived effect after physiotherapy in patients with neck pain: An observational study. *Physiotherapy* **2018**, *104*, 400–407. [CrossRef]
73. Hobart, J.C.; Cano, S.J.; Warner, T.T.; Thompson, A.J. What sample sizes for reliability and validity studies in neurology? *J. Neurol.* **2012**, *259*, 2681–2694. [CrossRef] [PubMed]
74. STMicroelectronics. Discovery Kit IoT Node. 2023. Available online: <https://www.st.com/en/evaluation-tools/b-1475e-iot01a.html> (accessed on 28 August 2024).
75. Schneider, A.; Hommel, G.; Blettner, M. Linear Regression Analysis. *Dtsch Arztebl Int.* **2010**, *107*, 776–782. [CrossRef] [PubMed]
76. Vetter, T.; Schober, P. Regression: The Apple Does Not Fall Far From the Tree. *Anesth. Analg.* **2018**, *127*, 1. [CrossRef]
77. Montgomery, D.C.; Peck, E.A.; Vining, G.G. Backward Elimination: A Stepwise Regression Procedure. *J. Bus. Econ. Stat.* **1996**, *14*, 499–504. [CrossRef]
78. Kutner, M.H.; Nachtsheim, C.J.; Neter, J.; Wasserman, W. *Applied Linear Regression Models*, 4th ed.; McGraw-Hill/Irwin: New York, NY, USA, 1996.
79. Goldfarb, A.; Coggan, D. Obesity and exercise testing: A review. *JAMA* **1995**, *274*, 1235–1242.
80. Ries, J.D.; Echternach, J.L.; Nof, L.; Gagnon Blodgett, M. Test-retest reliability and minimal detectable change scores for the timed “up & go” test, the six-minute walk test, and gait speed in people with Alzheimer disease. *Phys. Ther.* **2009**, *89*, 569–579.
81. Steffen, T.; Seney, M. Test-retest reliability and minimal detectable change on balance and ambulation tests, the 36-item short-form health survey, and the unified Parkinson disease rating scale in people with parkinsonism. *Phys. Ther.* **2008**, *88*, 733–746. [CrossRef]
82. Petersen, C.; Steffen, T.; Paly, E.; Dvorak, L.; Nelson, R. Reliability and minimal detectable change for sit-to-stand tests and the functional gait assessment for individuals with Parkinson disease. *J. Geriatr. Phys. Ther.* **2017**, *40*, 223–226. [CrossRef]
83. Lang, J.T.; Kassan, T.O.; Devaney, L.L.; Colon-Semenza, C.; Joseph, M.F. Test-retest reliability and minimal detectable change for the 10-meter walk test in older adults with Parkinson’s disease. *J. Geriatr. Phys. Ther.* **2016**, *39*, 165–170. [CrossRef] [PubMed]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.