



HAL
open science

Validity of Estimated Results from a Wearable Device for the Tests Time Up and Go and Sit to Stand in Young Adults and in People with Chronic Diseases

Kokouvi Geovani Agbohessou, Stephanie Sahuguede, Justine Lacroix, Fadel Hamdan, Emmanuel Conchon, Yannick Dumas, Anne Julien-Vergonjanne, Stephane Mandigout

► **To cite this version:**

Kokouvi Geovani Agbohessou, Stephanie Sahuguede, Justine Lacroix, Fadel Hamdan, Emmanuel Conchon, et al.. Validity of Estimated Results from a Wearable Device for the Tests Time Up and Go and Sit to Stand in Young Adults and in People with Chronic Diseases. *Sensors*, 2023, 23, 10.3390/s23125742 . hal-04674573

HAL Id: hal-04674573

<https://unilim.hal.science/hal-04674573>



Submitted on 21 Aug 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Article

Validity of Estimated Results from a Wearable Device for the Tests Time Up and Go and Sit to Stand in Young Adults and in People with Chronic Diseases

Kokouvi Geovani Agbohessou ^{1,*}, Stephanie Sahuguede ², Justine Lacroix ¹, Fadel Hamdan ², Emmanuel Conchon ², Yannick Dumas ³, Anne Julien-Vergonjanne ² and Stephane Mandigout ^{1,4,*}

¹ HAVAE EA6310 (Handicap, Aging, Autonomy, Environment), University of Limoges, 87042 Limoges, France

² XLIM Laboratory, UMR CNRS 7252, University of Limoges, 87000 Limoges, France

³ Développement de Logiciels, UNOVA, 87000 Limoges, France

⁴ ILFOMER (Institut Limousin de Formation aux Métiers de la Réadaptation), Université de Limoges, 87000 Limoges, France

* Correspondence: geovani.agbohessou@gmail.com (K.G.A.); stephane.mandigout@unilim.fr (S.M.)

Abstract: Background: Health care professionals need a valid tool to assess the physical ability of patients with chronic diseases. We aimed to assess the validity of the results of physical fitness tests estimated by a wrist wearable device in young adults and chronic disease people. Methods: Participants wore a sensor placed on their wrist and performed two physical fitness tests (sit to stand (STS) and time up and go (TUG)). We checked the concordance of sensor-estimated results using Bland–Altman analysis, root-mean-square error, and intraclass coefficient of correlation (ICC). Results: In total, 31 young adults (groups A; median age = 25 ± 5 years) and 14 people with chronic diseases (groups B; median age = 70 ± 15 years) were included. Concordance was high for both STS ($ICC_A = 0.95$, and $ICC_B = 0.90$), and TUG ($ICC_A = 0.75$, $ICC_B = 0.98$). The best estimations were given by the sensor during STS tests in young adults (mean bias = 0.19 ± 2.69 ; $p = 0.12$) and chronic disease people (mean bias = -0.14 ± 3.09 s; $p = 0.24$). The sensor provided the largest estimation errors over 2 s during the TUG test in young adults. Conclusion: This study showed that the results provided by the sensor are consistent with those of the gold standard during STS and TUG in both healthy youth and people with chronic diseases.

Keywords: accelerometer; gyrometer; wearable sensor; reliability; physical activity; chronic disease



Citation: Agbohessou, K.G.; Sahuguede, S.; Lacroix, J.; Hamdan, F.; Conchon, E.; Dumas, Y.; Julien-Vergonjanne, A.; Mandigout, S. Validity of Estimated Results from a Wearable Device for the Tests Time Up and Go and Sit to Stand in Young Adults and in People with Chronic Diseases. *Sensors* **2023**, *23*, 5742. <https://doi.org/10.3390/s23125742>

Academic Editor: Giovanni Saggio

Received: 19 April 2023

Revised: 15 June 2023

Accepted: 16 June 2023

Published: 20 June 2023



Copyright: © 2023 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

Chronic diseases have a significant impact on life expectancy and almost 20% of people have experienced these conditions in France [1]. Chronic diseases represent a major public health issue in France and the most common are cardiovascular diseases (5.1 million people treated in 2019), diabetes (4.0 million), chronic respiratory diseases (3.7 million), and cancers (3.3 million) [2].

In a recent systematic review, studies have mentioned the existence of a positive effect of the practice of physical activity (PA) on the health of these people [3]. Indeed, it has been observed that following PA programs improve the functional abilities of people with chronic diseases [4,5]. However, these programs can be difficult to access for patients, due to lack of availability on the scheduled sessions, lack of economic resources or significant geographical distance. Limited adherence to these PA programs is generally observed in these patients [6]. Concern about the lack of coordination and continuity of care for chronic illness has led to a series of reforms in the health care field. The development of telemedicine in France has encouraged the remote management, evaluation and monitoring of patients with chronic diseases [2].

Monitoring PA levels, to guide rehabilitation, is a challenge for clinicians. Personal activity monitors are becoming increasingly popular and offer the potential to improve rehabilitation protocols at a distance. Measuring its own PA performance is important for adults with chronic diseases to comply with PA recommendations in clinical practice [7,8].

The recent development of small, wearable electromechanical sensors may bridge the gap between the practicality of commonly used clinical tools and more accurate, but expensive and difficult-to-access technological systems for patients. This reduced complexity may encourage more clinicians who are unfamiliar with the technical aspects of the new equipment to integrate sensors into practice [9].

Among these sensors, there are the inertial motion units (IMU) that encompass accelerometer, magnetometer and gyroscope axis. In a five-year update, Picerno et al. showed that IMU, when combined with data processing software, can define person movement, heading and orientation [10]. The accuracy of wearable IMU has improved over time with several studies validating its use for PA measurement such as step count, walked distance, speed measurement and energy expenditure in free-living conditions [10–12].

In clinical practice, health professionals used valid physical fitness tests to assess chronic disease patients. These tests are quick, require no special device or training, and are easily included as part of the routine medical examination. Among the tests commonly used by professionals, sit to stand test (STS) assesses lower body strength, needed for daily life activities. It (the number of STS completed) has shown a good correlation ($r = 0.75$) with the gold standard (dynamometer) in older people [13]. On the other hand, the timed up and go test (TUG) assesses the functional mobility and dynamic balance of persons and allows the detection of older adults at risk of falling [14]. The TUG has also excellent test-retest reliability ($ICC = 0.95$) for individuals with chronic stroke for example [15].

In contrast to healthy subjects, in chronic disease patients, functional heterogeneity such as slowness of the movements, abnormalities in gait, heart rate and metabolism make the assessment of PA particularly challenging using wearable IMU. Research on the validity and reliability of these devices in people with chronic diseases is limited [16]. Above all, it is also important to validate the device in a population of young healthy adults in order to optimize its functioning.

The main aim of the chronic patient rehabilitation process is to optimize patients' performances of motor tasks in daily life. Therefore, clinicians need a reliable and valid measurement that will allow them to establish the patient's functional mobility and monitor the patient's progress at a distance [17].

Previous works using combined data acquired from multiple device placements (e.g., devices placed at hip and trunk, smartphone-based or ultrasound-based sensor) showed that devices could be used to study transitions in STS test [18] and kinematic parameters in TUG tests [19]. These wearable devices require multiple sensors and the sensor to be placed on the L3 region of the subject's lumbar spine [18,19]. Wear several accelerometer devices slightly enhance activity recognition accuracy in chronic disease patients. However, given the extra burden of wearing additional devices for patients, single accelerometers with appropriate placement appear to be sufficient for estimating the number of STS and TUG activity times in chronic disease patients [20,21]. The use of an IMU, therefore, appears to be sufficient to determine people's level of physical fitness through the STS and TUG. To our knowledge, no study has evaluated the accuracy of a simple wrist wearable IMUs sensor for performing these clinical tests.

Most of the existing sensors are closed to reprogramming; we have therefore developed a personal sensor that has allowed us to implement our algorithm capable of evaluating the muscular strength, and dynamic balance of chronic disease people.

The purpose of this study was therefore to evaluate the validity and internal reliability of STS and TUG tests estimated by our personal custom device in young adults and people with chronic diseases.

2. Materials and Methods

2.1. Participants Selection

We have carried out a cross sectional exploratory study. Volunteer participants were allocated to two groups; group A (healthy young adults aged from 18 to 35 years old) and group B (adults with chronic diseases of more than 6 months duration). The healthy group was the test group and consisted of student volunteers who met our inclusion criteria and reported no chronic diseases. For group B (Who had diseases such as diabetes, breast cancer, multiple sclerosis, morbid obesity...), chronic disease participants were contacted through patient associations. We have considered several pathologies to reflect the variability of illnesses among people with chronic diseases.

They were recruited in a laboratory setting and all tests were performed in a gymnasium. Participants with acute cardiac or respiratory pathologies were not included. Participants who had any conditions that might affect the assessment protocol, such as neurological or musculoskeletal problems affecting mobility and hindering proper assessment were not included. The therapist responsible for the protocol informed the participants of the details of the protocol and registered their consent. The research protocol was approved by an ethics committee, "Comité d'Ethique pour la Recherche en STAPS" (CERSTAPS, Notice number: IRB00012476-2023-16-01-220).

2.2. Instrumentation

2.2.1. IMU-Based Custom Device Description

We developed a custom wearable system able to be worn at the wrist (Figure 1), based on the ICM20948 9-axis IMU from STMicroelectronics. The system includes a processing unit (STM32WB55RG) for acquisition, storage and communication, a memory unit, a communication unit based on both USB and Bluetooth connectivity, and a screen and buttons to interact with the user. In order to deal with memory size and consumption, the data from the IMU are stored in the device and transferred by USB at the end of the exercises. The sampling frequency is set to 50Hz in order to be able to memorize values from each of the 9 axes. We performed the same test with the two sampling frequencies of 50 Hz and 100 Hz. It was verified that the results were identical.

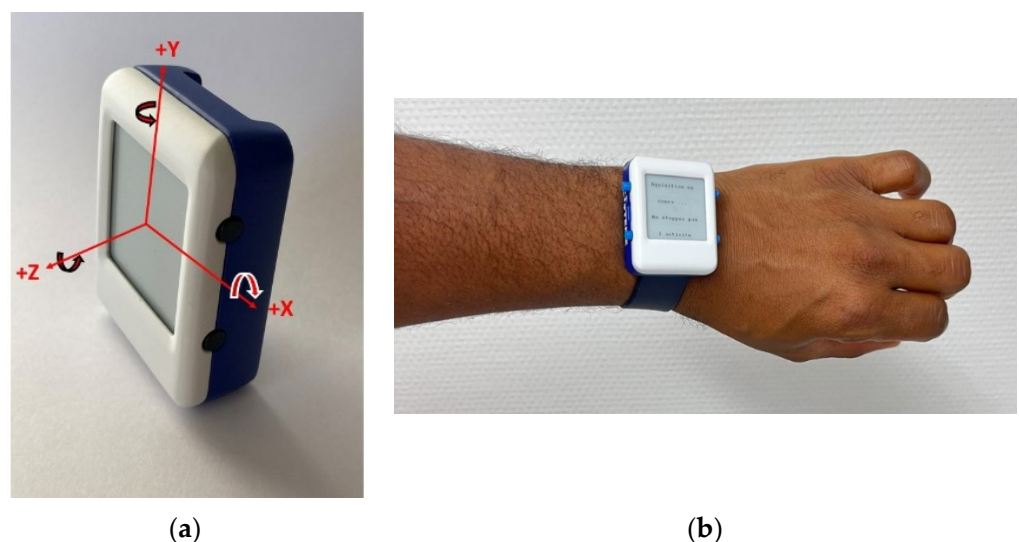


Figure 1. Orientation of the axes and device placement. (a) Orientation of the axes. (b) Device placement and stand-by screen.

Figure 1a shows the orientation of the axes on the prototype and Table 1 shows the main characteristics of the device.

Table 1. Main characteristics of the device.

Characteristics	Value
IMU 9-axis sensor	ICM20948
Microcontroller processing unit	STM32WB55RG
On chip memory (program + data)	1 Mbyte
DAC resolution	12 bits
Sampling frequency	50 Hz
Connectivity	BLE and USB
Power	Lithium ion battery 560 mAh 3.7 V

The data are then post processed on a computer unit, in order to determine the duration of STS and TUG exercises.

2.2.2. The Algorithms

We processed several sets of values in order to identify among all recorded data (accelerometer, gyrometer and magnetometer, each over 3 axes) which are the most suitable for the analysis of the tests (number of sit-to-stand performed during STS test or time duration of TUG test). The algorithms have been set up following an iterative methodology: a preliminary version based on the analysis of the signal of some healthy subjects has been established. Then, we adapted them with the analysis of the signals corresponding to a group of people with chronic diseases. Each algorithm methodology is explained in detail in the following paragraphs.

STS Algorithm

The objective of the processing is to determine the number of sit-to-stand completed in 30 s from the raw data of the recorded values from the IMU.

First, we observed based on preliminary tests that the accelerometer data on each of the 3 axes gives interesting information compared to the other raw data and that it is possible to observe periodic variations linked to the repetitiveness of the exercise. The analysis of a large set of data permitted us to identify in particular that the x-axis of the accelerometer raw data was the one from which information could be extracted the most reliably. Periodic spikes permit identifying the participants' successive chair lifts during the test.

From these raw data, the first processing consisted in carrying out low-pass filtering in order to smooth the values and remove insignificant variations. We used a 4th order Butterworth filter with a cutoff frequency of 0.5 Hz. Indeed, this low value is justified by the fact that it is not possible to perform more than two chair lifts in 1 s.

To determine the number of chair lifts, the strategy then adopted is to count the number of points passing through the median of the curve.

However, for chronic disease, people slow down the speed at which they perform the test because of their diseases, the tracings become irregular and compared to a simple median introduces counting errors. To solve this problem, a sliding median was calculated every 2 s.

Figure 2 shows an example of a plot with a fixed threshold and with a sliding window, for two different persons from the healthy group (Figure 2a) and from the chronic disease one (Figure 2b). These curves highlight the fact that considering a fixed threshold may introduce errors in the counting. Even for a healthy person having regular exercise, some spikes are not identified: in the example (Figure 2a), only one at the beginning of the exercise is missing. In the case where the exercise realization is not regular such as in Figure 2b, it conduces to totally erroneous values. The sliding window with a 2 s width undercomes this issue.

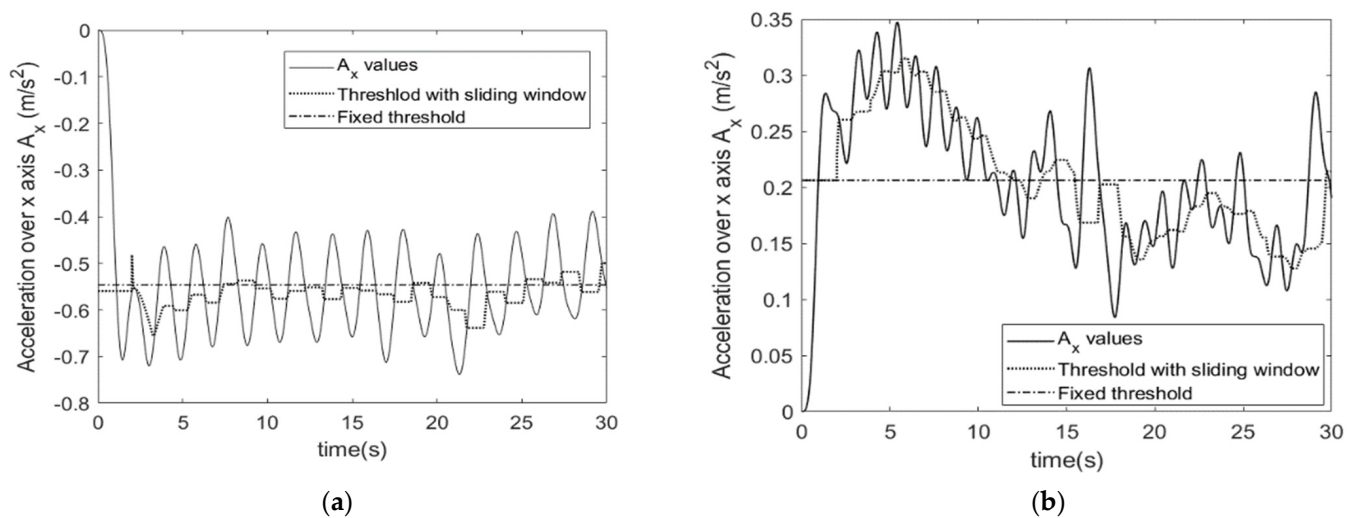


Figure 2. Illustration of sliding window process on x-axis if accelerometer data to identify spikes in the sit to stand (STS) algorithm. (a) Example issued from healthy group. (b) Example issued from chronic disease group.

TUG Algorithm

In this case, the objective is to determine the time taken by the participant to perform the TUG test. It is supposed to be lower than 30 s.

Preliminary tests permitted us to identify that the most significant data to be analyzed during this test is the angular velocity data from the gyrometer, recorded for 30 s, and more precisely the z-axis of gyrometer. In order to remove insignificant variation, the first processing is low-pass filtering as in the STS test.

We used a 4th order Butterworth filter with a cutoff frequency of 5 Hz. Indeed, we have empirically determined from the different recorded sets of values, that the cutoff frequency should be between 4 Hz and 7 Hz. A higher cutting frequency leads to insignificant perturbations in the signal, whereas a lower removes useful data. The efficiency of this cutoff frequency has been verified for people having a mean speed of walking around 0.5–1.5 m/s depending on people's health condition and age.

Figure 3 shows two examples of z-axis gyrometer data recording during the TUG for two very different profiles. The objective is to evaluate the duration of the test T_{TUG} which is between the last extremum on the plot and the interval where the angular velocity is stable, corresponding to the phase where the participant has returned to sit down on the chair and where he is at rest.

The method used then consists in identifying on the recording, the instant at which stability is reached, denoted as T_{stab} and extracting the time T_{ext} where the plot exhibits a local extremum. The end of the test is between these two values T_{ext} and T_{stab} .

Empirically we considered that the duration of the test is equal to:

$$T_{TUG} = \frac{(T_{stab} + T_{ext})}{2} \quad (1)$$

In order to remove artifacts of movements, we determined that the stability instant corresponds to the time at which the variations in angular velocity are less than 10°/s during a duration of 3 s. The points corresponding to T_{ext} and T_{stab} are pointed out in the plot's examples in Figure 3.

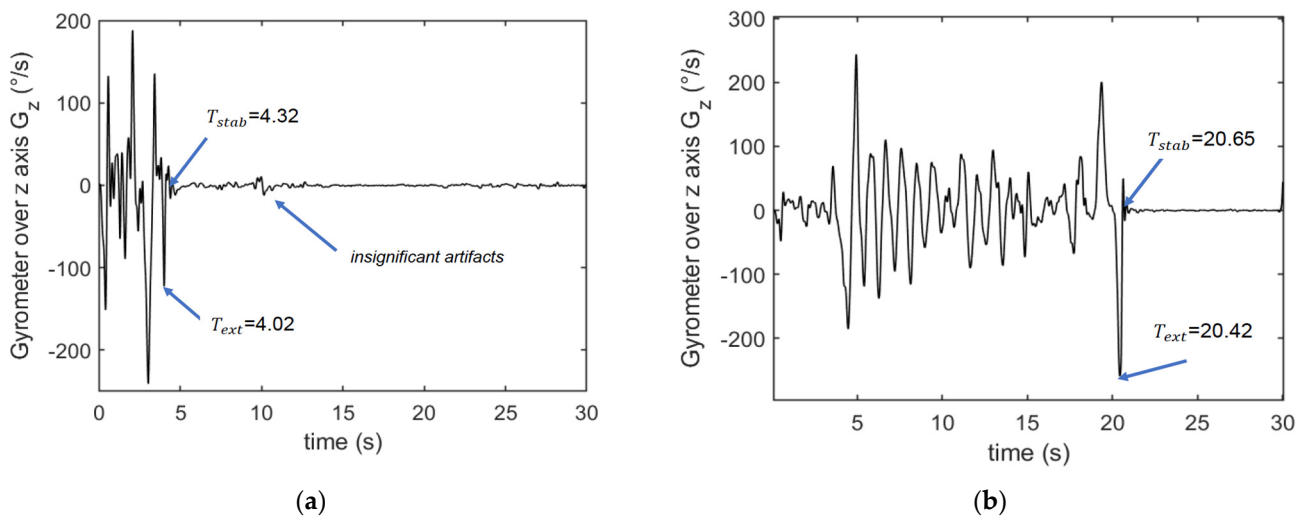


Figure 3. Illustration of z-axis of gyrometer evolution of timed up and go (TUG) exercise. (a) Example issued from healthy group. (b) Example issued from chronic disease group.

2.3. Physical Fitness Tests

We have selected a gold standard for each test that has been validated in chronic disease people.

2.3.1. Sit to Stand Test

The STS aimed to assess lower body strength [22]. Subjects were instructed to stand up fully and sit down in a chair with their back against the rest as fast as possible in 30 s with their arms crossed over their chest throughout. The number of repetitions performed during the 30-s STS test was simultaneously recorded by the examiner (gold standard) and our custom device [13]. STS performance is known to be associated with disability, falls, hip fracture and mortality among older adults and chronic disease people [23].

2.3.2. Timed “Up and Go” Test

The TUG assesses the functional mobility of persons. The subject stands up from a chair, walks 3 m forward, turns around, walks back and sits down again [14]. The end of the walkway was clearly marked with a pin. Subjects were instructed to sit straight on the chair with their hands on their thighs and their backs touching the back of the chair. After the tester gave them the go signal, they arose from the chair, walked at their normal speed, turned around right after passing the pin at the end of the pathway, returned back to the chair, turned around and sat down. The time to complete the test is recorded on the one hand by the investigator using a Handheld Stopwatch (HHS) and on the other hand, is estimated by the custom device.

2.4. Test Protocol and Data Processing

Data collection took place in France between September 2022 and November 2022. Self-reported measures were given by participants. Body mass index (BMI) was calculated from these measures. Participants were advised to wear suitable sports clothing and footwear for the test.

The test protocol involved the following steps:

- A. Warm-up session with stretching, joint mobilization and muscle strengthening exercises.
- B. Installation of the custom device placed on the left wrist of the participant during all tests. The installation of the sensor required a maximum of 3 min. Prior to use by each participant the device was automatically initialized via the supporting software (USB, prototype). To perform a test, the investigator had to choose the test from the

- menu, press the “start” button and 3 s later the device vibrated and the participant could start the test, see Figures 1–3.
- C. Prior to each test, the participant performed a trial run to ensure that they understood the instructions (this step may take about 2 min). The participants performed STS and TUG tests, the results were also measured by the examiner using visual counts for STS and a stop-watch for TUG. The activity mode was automatically disabled at the end of each test. The data collection is carried out during all the tests which is supposed to take no more than 2 times 30 s, so less than 1 min. With the given acquisition frequency of 50 Hz, this corresponds to 3000 measurements. Each measurement corresponds to 9 values issued from the 9-axis accelerometer. With the 12 bits-ADC of the device, a 1-min recording corresponds to less than 40 kbytes.
 - D. Download of the data from the device: Data from the sensor was downloaded via a universal serial bus (USB) and then processed. The members of the research team used a laptop to configure the device, start a test, and read the outcomes at the end of each test. The testing procedure is illustrated in Figure 4.

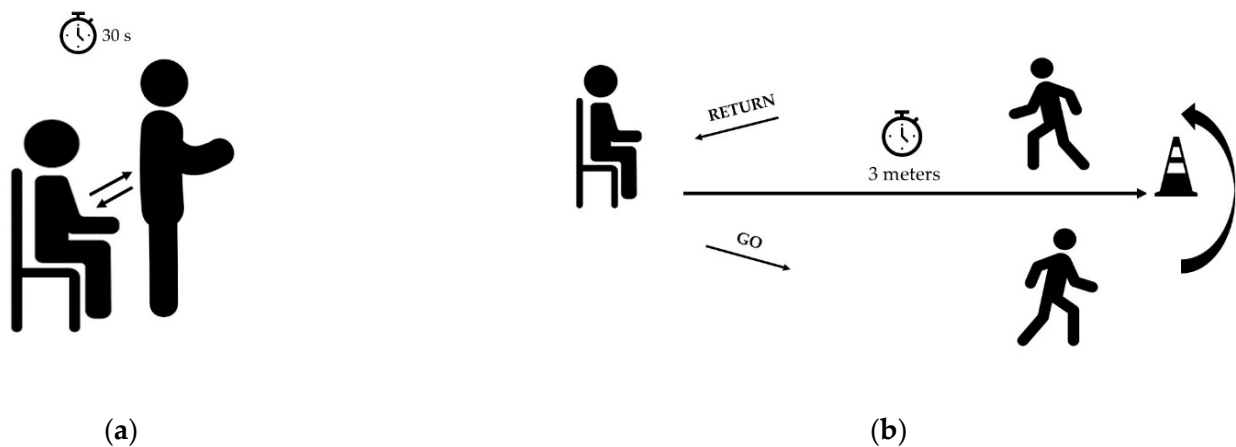


Figure 4. Illustration of the test protocol. (a) Sit to stand (STS) test. (b) Timed up and go (TUG) test.

2.5. Statistical Analysis

Data analysis was performed with R version 4.2.1. The normality of the distribution of data was examined using the Kolmogorov-Smirnov test [24,25].

Descriptive statistics were calculated for all participants’ characteristics. Variables will be expressed as median with interquartile range [IQR], mean \pm standard deviation (SD) or number (percentage), according to the normality of the distribution of data.

Sample size numbers were determined by procedures described by Walter et al., for inter-device reliability [26]. We calculated the required sample size based on a previous study [7]. Twenty-five subjects were deemed to be acceptable to judge the difference between two devices with a minimally acceptable level of 0.5, when $\alpha = 0.05$ and $\beta = 0.20$ (power = 0.8). The risk of data loss related to possible malfunctioning devices was also included in the calculation. We estimated this risk at about 15%, so we needed a total of 30 individuals.

Validity was evaluated for the personal sensor estimations and the measures observed by the investigator during every test. We used the Spearman correlation [27] to determine if both methods of measurement analysis produced comparable results.

Intraclass Correlation Coefficients (ICC) [28] were used for the inter-rater analysis to determine if the custom device can be used with confidence and reliability to assess physical fitness in people. Reliability and validity were assessed using the criteria suggested by Terwee et al. [29]. Based on the 95% confidence interval of the ICC estimate, values less than 0.5, between 0.5 and 0.75, between 0.75 and 0.9, and greater than 0.90 are indicative of poor, moderate, good, and excellent reliability, respectively.

Inter-rater reliability was visualized using Bland–Altman plots with a mean bias (MB), 95% limits of agreement (LoA) calculated using the formula: 95% LoA = mean difference \pm $1.96 \times$ SD [30]. We also appreciated differences between values given by the custom device and the values observed by the investigator using the root-mean-square error (RMSE).

There is currently no conventional cutoff in the literature to define the acceptable levels of inter-observer agreement. A minimum possible error is expected. Based on prior studies and clinical experience, any differences between raters greater than ± 5 units were considered clinically significant, and exceeding ± 10 units were considered unacceptable [31]. The threshold of significance for all tests was 0.05 and 95% confidence intervals (CI) were reported.

3. Results

3.1. Population Characteristics

Fourteen people with a chronic disease and thirty-one healthy young adults were recruited in this study. The median age was 70 [15] years in the chronic disease group and 25 [5] years in the healthy group. The proportion of females was higher in the chronic disease group than in the healthy group (64% vs. 48%). The average BMI was significantly higher in the chronic disease group than in the healthy group (31.7 ± 3.6 vs. 24.6 ± 3.9) $\text{Kg}\cdot\text{m}^{-2}$. In both groups, the average number of STSs with the gold standard was the same as with the customized device.

In addition, in the healthy group, the average number of seconds in the TUG test with the gold standard was slightly higher in the healthy group (5.9 ± 0.8 vs. 5.5 ± 1.2 s), but in the chronic disease group, the gold standard showed lower seconds in the TUG test than the custom device (6.7 ± 3.1 vs. 7.0 ± 3.4 s). Descriptive statistics for all the subjects are presented in Table 2.

Table 2. Demographic characteristics of the population.

Characteristic	Healthy Adults, n = 31	Chronic Disease, n = 14
Sex		
Female	15 (48%)	9 (64%)
Male	16 (52%)	5 (36%)
Weight (Kg)	70.0 [12]	82.0 [10]
Height (m)	1.7 ± 0.1	1.6 ± 0.1
STS observed (n)	18.9 ± 4.5	14.1 ± 4.0
STS sensor (n)	18.7 ± 3.8	14.2 ± 3.0
TUG observed (s)	5.9 ± 0.8	6.7 ± 3.1
TUG Sensor (s)	5.5 ± 1.2	7.0 ± 3.4

The values are expressed in Median [IQR]; Interquartile range, Mean \pm Standard deviation); Kg: Kilogram; m: meter; STS: Sit to stand test; n: number; TUG: Timed Up and Go test; s: second.

3.2. Validity Parameters

In young healthy adults (group A), TUG showed a significant ($p < 0.001$) and good correlation ($r = 0.80$) between the custom device and investigator measures. STS test in the healthy adults' group demonstrated also a significant correlation ($r = 0.96$, $p < 0.001$). The TUG test showed less reliability in the healthy adults ($\text{ICC}_A = 0.75$, $p < 0.001$) than in chronic disease participants ($\text{ICC}_B = 0.98$, $p < 0.001$). Reliability was high for the STS test, in young adults ($\text{ICC}_A = 0.95$, $p < 0.001$) and chronic disease participants ($\text{ICC}_B = 0.90$, $p < 0.001$). Both, TUG ($r = 0.87$, $p < 0.001$) and STS ($r = 0.94$, $p < 0.001$) tests in chronic disease participants (group B) showed a significant correlation. The details of the correlations are summarized in Table 3.

Table 3. Validity parameters of the results estimated by wearable custom device versus results observed by examiner.

Measures	r	p-Value	ICC	95% CI	p-Value
STS group A	0.96	<0.001	0.95	0.9, 0.97	<0.001
STS group B	0.94	<0.001	0.90	0.74, 0.97	<0.001
TUG group A	0.80	<0.001	0.75	0.43, 0.79	<0.001
TUG group B	0.87	<0.001	0.98	0.91, 0.99	<0.001

STS = Sit-to-stand test, TUG = Timed Up and Go test; group A = healthy group; group B = chronic disease group; r = Spearman correlation coefficient; p-value = statistical significance; ICC = Intraclass correlation coefficient; CI = Confidence Intervals.

The parameters of validity of each method are presented in Table 3. In the chronic disease group, the custom device slightly overestimated the number of trials performed for STS (Mean Bias (MB) = -0.14) and the number of seconds performed on the TUG test (MB = -0.36). The best estimations were given by the sensor during STS tests in chronic disease patients (MB = -0.14 s; $p = 0.24$) and in young adults (MB = 0.19 ; $p = 0.12$). The large estimation errors over 2 s ($p < 0.05$) were provided by the custom device during the TUG test in young adults (MB = 0.45 ; 9.24%).

On the other hand, the most accurate estimations were obtained using the custom device for STS tests in young adults (RMSE = 7.22%), chronic disease people (RMSE = 10.83%) but also for the TUG test in chronic disease people (RMSE = 11.02%). The custom device showed lower accuracy for the TUG test in healthy adults (RMSE = 15.75%), (Table 4).

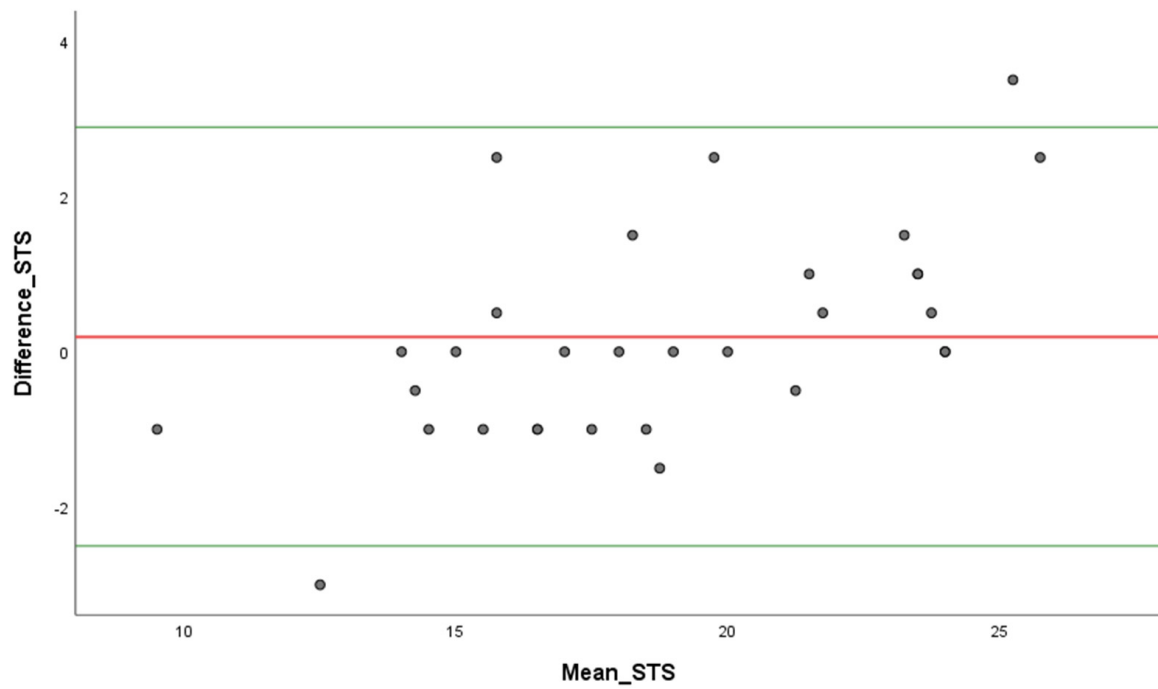
Table 4. Inter-rater reliability parameters of the results estimated by wearable custom device versus results observed by examiner.

Measures	Mean Bias	Percentage Difference (%)	95% LoA Down	95% LoA Up	RMSE	Percentage RMSE (%)
STS group A	0.19	0.08	-2.5	2.89	1.37	7.22
STS group B	-0.14	-3.21	-3.23	2.94	1.52	10.83
TUG group A	0.45	9.24	-1.18	2.08	0.94	15.75
TUG group B	-0.36	-4.56	-1.66	0.95	0.74	11.02

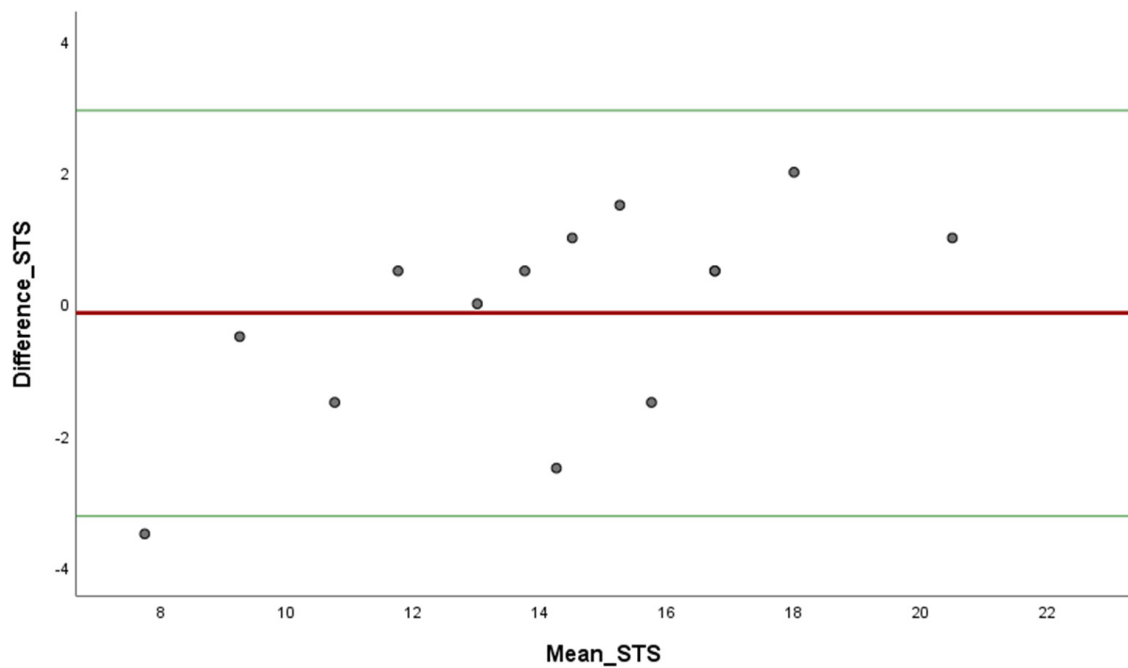
STS = Sit-to-stand test (expressed in number), TUG = Timed Up and Go test (expressed in second); group A = healthy group; group B = chronic disease group; Percentage Difference = mean bias expressed in percentage of results measured and estimated by device; 95% LoA = limits of agreement of Bland–Altman analysis; RMSE = root-mean-square error; Percentage RMSE = RMSE expressed in percentage of result observed by examiner.

We used a Bland–Altman graphical approach to check the assumptions of normality of differences between two measures. We quantified agreement between two quantitative measurements by constructing limits of agreement. These statistical limits are calculated by using the mean and the standard deviations of the differences between two measurements. Inter-rater reliability parameters are illustrated by Bland–Altman plots in Figures 5 and 6.

We found a high concordance with proportional bias in chronic disease people during STS (Figure 5b) and TUG (Figure 6b) tests. On the one hand, we found a high concordance with proportional bias in healthy young adults for the STS test (Figure 5a). On the other hand, as shown in Figure 6a, we found the lowest concordance with non-proportional bias in young adults during the TUG test.

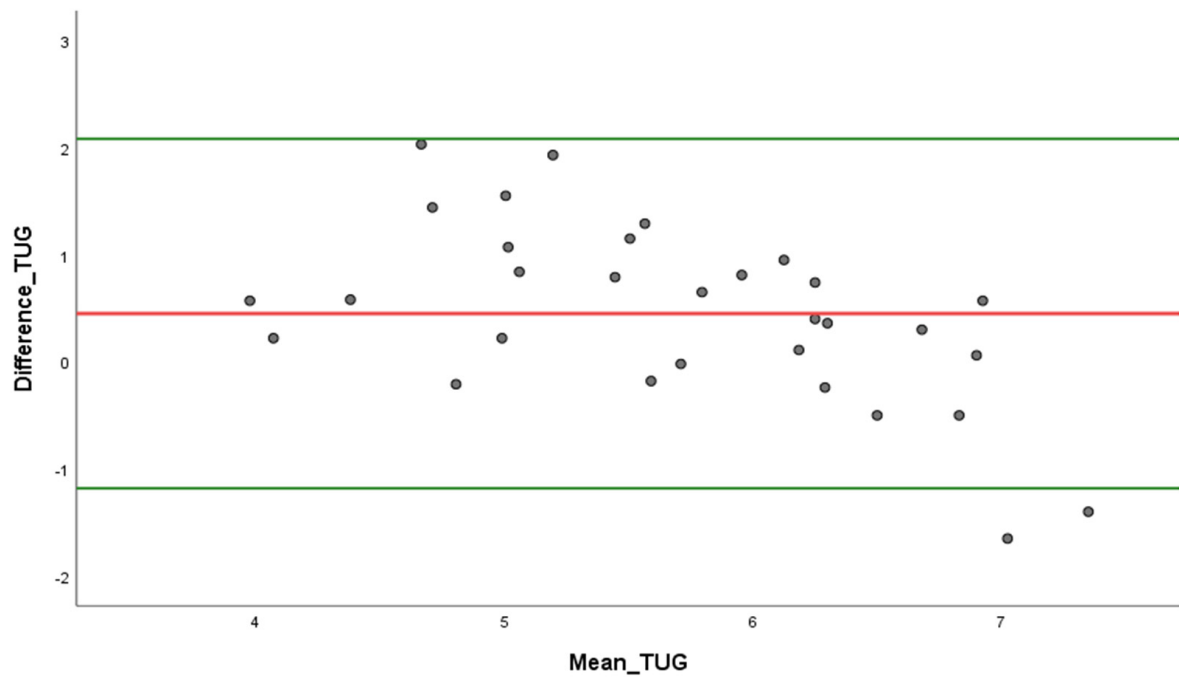


(a)

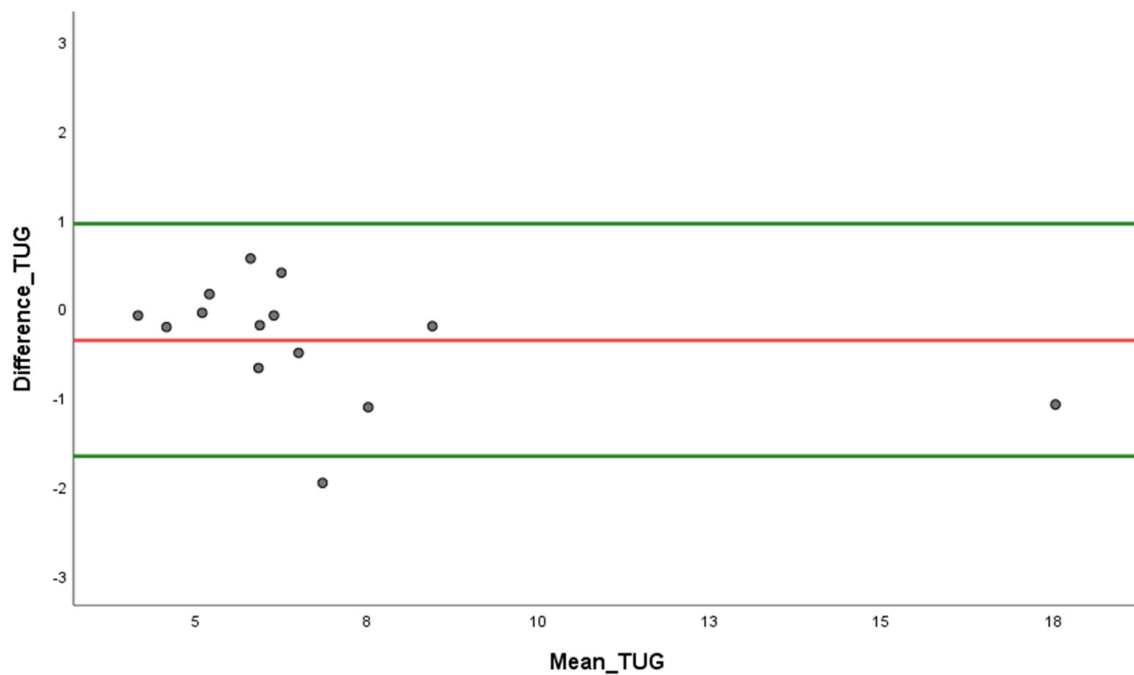


(b)

Figure 5. Bland–Altman plots for sit to stand (STS) test. (a) Healthy young adults. (b) Chronic disease people.



(a)



(b)

Figure 6. Bland–Altman plots for timed up and go (TUG) test. (a) Healthy young adults. (b) Chronic disease people.

4. Discussion

This study evaluated our personal wrist wearable device's reliability and validity in healthy young adults and chronic disease patients. It compared the output from the custom device to the measure recorded by an investigator for STS and TUG tests. The

findings indicate that the custom device provides a valid outcome of a number of STS trials as well as the time for performing the TUG test with high intraclass correlation coefficients compared to estimations observed by the investigator.

Analyses of the relationships between the number of STS provided by the custom device and by observed measures in chronic disease people ($r = 0.94$) and healthy young adults ($r = 0.96$) showed excellent and significant correlation (<0.001). The STS intraclass correlation coefficients were 0.90 in chronic disease people and 0.95 in healthy adults. These coefficients are high enough to conclude that our device provides reliable measurements in the STS test using an ICM20948 IMU sensor and suitable processing thanks to the developed algorithms.

Cobo et al. used an ultrasound-based sensor to assess STS and they found a similar result in young adults (with an ICC = 0.96), and in the elderly group (with an ICC = 0.89) compared to our device [20]. Their device comprises an ultrasound sensor, a web board, and a Bluetooth module. In contrast to our system, their device required a special chair with more sensors and owners have noted that the device is difficult for chronic disease people to use alone.

Despite the high sampling rate of our custom device, it slightly underestimates the results of STS and TUG tests in the young adults group. We observed better performance in the healthy youth group (mean STS = 18.9, mean TUG = 5.9 s) compared to the chronic disease group (mean STS = 14.5, mean TUG = 6.7 s), which confirms an overall functional decrease and certain mobility constraints in the chronic disease group. Their movements become less precise, and, during the experimentation, they might have moved their wrist less prominently than healthy young adults. There is probably an age effect, as people with chronic diseases are older and therefore more likely to have poor physical activity performance compared to young adults [18]. A systematic review of the literature on previous studies highlighted this limitation related to the use of IMU for the assessment of chronic disease and elderly people [18]. The main objective of our work is to adapt our device as much as possible to chronic disease people, so the algorithms are more specific to their functioning. In addition, during the evaluations, the chronic disease patients performed the tests with more applications because they are familiar with them, unlike young adults.

In a recent study, Fudickar et al. used a sensor placed on the chest of patients. They also observed a significant correlation ($r = 0.73$) between the results provided by the device and the stop-watch during the TUG. On the other hand, they noted difficulty with chronic diseases people wearing the sensor and handling it [31], which could indicate age-related (mean, 79.41 years) limitations such as reduced impaired motor and sensory [32,33]. Although the strap was already lengthened in their study, some participants, especially those with higher body weight, still felt it was too narrow and should be lengthened again. In our study, no such difficulties were reported, so the participants found the use of our device very useful.

4.1. Accuracy

In this study, the most accurate estimations were obtained using the custom device for STS tests in young adults (RMSE = 7.22%), chronic disease people (RMSE = 10.83%) but also for TUG test in chronic disease people (RMSE = 11.02%). The custom device showed lower accuracy for the TUG test in healthy adults (RMSE = 15.75%). This lower precision could be explained by the fact that the algorithm was adapted for people with chronic diseases, but we also used it for healthy people.

The findings of our study showed that mean Handheld Stopwatch (HHS) times during TUG was 0.08 s faster than our custom device in the chronic disease group. These values are in close agreement with Hetzler et al., who have reported HHS times to be 0.07 s faster than electronic timing in 200-m sprint [34]. However, in the healthy group, we found that the mean HHS during TUG was 0.13 s lower than our custom device. In another study, Moore et al., have reported HHS times to be 0.08 s slower than electronic times [35]. The

authors attribute these conflicting results to differences in starting procedures used by HHS operators. It is possible that factors associated with the study itself, including the speed of the runners and the starting protocol employed, may have affected the accuracy of HHS times relative to electronic times [34]. It is possible that faster participants (healthy adults) may be more difficult to time accurately using HHS during the TUG, thus increasing the possibility of timing errors.

Additionally, during the TUG test, some healthy adults had to hold their wrists still for a while to look at the screen to ensure the watch was recording activity during the TUG test. They may thus have degraded the reliability. However, we did not collect data to support this hypothesis. We found no other work evaluating the accuracy and error rate following the use of a wrist wearable IMU device to assess STS and TUG in chronic disease people.

Our device is worn on the wrist and remains less bulky compared with existing devices. Complex devices show a low utilization rate in contrast to wrist-worn devices that are easy to wear unaided by chronic disease people. In future studies, we need to evaluate the acceptability of our wrist-worn IMU device. The latter has proven to be accurate, but the ease of use and user experience of chronic disease people remains unknown.

4.2. Limitations

We are aware that the study includes some limitations. The main limitation is the method used to assess the accuracy of the results provided by our custom device. There is no consensus on the acceptable error rate (for MB and RMSE). The authors recommend having the minimum error possible, the error rate varies between 5 and 15% [36].

Furthermore, it is important to remember that, although the groups were divided into “healthy” and “chronically disease” subjects, no intra-group separation was made, which would necessitate taking into consideration the characteristics and differences between men and women, and the specific diseases of the participants when interpreting the results.

4.3. Strengths of the Study

Indeed, there are expensive, complex and cumbersome devices to assess patients' balance, as opposed to our custom device. To our knowledge, to date, no study has investigated the reliability of a wearable wrist sensor to assess patients' condition using physical fitness tests. These findings allow us to evaluate chronic disease patients' physical fitness at a distance using the IMU devices with our algorithms. This wearable device will be integrated into a web platform to allow healthcare professionals to remotely monitor the fitness of chronic disease patients and to guide them in their effort training with adapted exercises. The excellent correlations between our personal measures and observed measures indicate a reliable measure. Our system produced valid results but also an acceptable accuracy with reference to the existing literature.

5. Conclusions

For both young healthy adults and chronic disease people, this integrated wearable device provides valid numbers of STS and seconds in the TUG test with acceptable accuracy. Inter-device reliability is generally good for STS and TUG testing in healthy young adults and people with chronic diseases. The treatments and algorithms developed from the raw data are therefore suitable for people with chronic diseases.

These findings make this wrist-worn device particularly useful for remotely monitoring progress in mobility and strength in patients with chronic conditions. IMU devices can then be easily integrated into a web-based platform so that therapists can automate the physical fitness assessments in order to follow chronic disease patients and improve their life condition. Future studies will be necessary to consolidate the results observed in this first study by working on a larger population. Based on this study, it would be important to conduct future work to integrate other algorithms to assess other physical skills, such as endurance and static balance.

Author Contributions: Conceptualization, K.G.A., S.S., J.L. and S.M.; methodology and investigation, K.G.A., J.L. and S.M.; software, K.G.A., S.S., F.H., Y.D. and A.J.-V.; writing—original draft preparation, K.G.A., S.S., E.C. and J.L.; writing—review and editing, All authors; project administration and funding acquisition, A.J.-V. and S.M. All authors have read and agreed to the published version of the manuscript.

Funding: This work was mainly supported by the University of Limoges and the Région Nouvelle Aquitaine, both of whom are not involved in the analyses. The sponsors were not involved in the design of the study, the collection, analysis, and interpretation of data, and in writing the manuscript.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Ethics Committee of “Comité d’éthique pour la recherche en STAPS” (Ethics committee for research in STAPS) CERSTAPS (protocol code IRB00012476-2023-16-01-220), France.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study. Written informed consent has been obtained from the participants to publish this paper.

Data Availability Statement: In accordance with the RGPD law in effect in France and the requirements of the ethics committee, the data from this study are stored on a secure server. The datasets generated for this study are available on request to the corresponding authors.

Acknowledgments: Authors would like to thank all of our participants for volunteering in this study and “Cisteme” for its technical support and the development of the personal custom device.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Chassang, M.; Gautier, A. Les Maladies Chroniques. 2019. Available online: <https://www.lecese.fr/actualites/maladies-chroniques-le-cese-adopte-son-avis> (accessed on 11 June 2022).
2. Chevreur, K.; Brunn, M.; Durand-Zaleski, I.; Farsi, F.; Knai, C.; Nolte, E. France [Internet]. Assessing Chronic Disease Management in European Health Systems: Country Reports [Internet]. European Observatory on Health Systems and Policies, 2015. Available online: <https://www.ncbi.nlm.nih.gov/books/NBK458739/> (accessed on 13 October 2022).
3. Kunstler, B.E.; Cook, J.L.; Freene, N.; Finch, C.F.; Kemp, J.L.; O’Halloran, P.D.; Gaida, J.E. Physiotherapist-Led Physical Activity Interventions Are Efficacious at Increasing Physical Activity Levels: A Systematic Review and Meta-analysis. *Clin. J. Sport Med. Off. J. Can Acad. Sport Med.* **2018**, *28*, 304–315. [CrossRef] [PubMed]
4. Kujala, U.M. Evidence on the effects of exercise therapy in the treatment of chronic disease. *Br. J. Sports Med.* **2009**, *43*, 550–555. [CrossRef] [PubMed]
5. Jerant, A.F.; von Friederichs-Fitzwater, M.M.; Moore, M. Patients’ perceived barriers to active self-management of chronic conditions. *Patient Educ. Couns.* **2005**, *57*, 300–307. [CrossRef] [PubMed]
6. Gomersall, S.; Maher, C.; English, C.; Rowlands, A.; Olds, T. Time regained: When people stop a physical activity program, how does their time use change? A randomised controlled trial. *PLoS ONE* **2015**, *10*, e0126665. [CrossRef] [PubMed]
7. Roberts-Lewis, S.F.; White, C.M.; Ashworth, M.; Rose, M.R. Validity of Fitbit activity monitoring for adults with progressive muscle diseases. *Disabil. Rehabil.* **2022**, *44*, 7543–7553. [CrossRef]
8. Ndahimana, D.; Kim, E.K. Measurement Methods for Physical Activity and Energy Expenditure: A Review. *Clin. Nutr. Res.* **2017**, *6*, 68–80. [CrossRef]
9. Jones, D.; Crossley, K.; Dascombe, B.; Hart, H.F.; Kemp, J. Validity and reliability of the fitbit flextm and actigraph gt3x+ at jogging and run-ning speeds. *Int. J. Sports Phys. Ther.* **2018**, *13*, 860–870. [CrossRef]
10. Picerno, P.; Iosa, M.; D’Souza, C.; Benedetti, M.G.; Paolucci, S.; Morone, G. Wearable inertial sensors for human movement analysis: A five-year update. *Expert Rev. Med. Devices* **2021**, *18* (Suppl. S1), 79–94. [CrossRef]
11. Li, C.; Chen, X.; Bi, X. Wearable activity trackers for promoting physical activity: A systematic meta-analytic review. *Int. J. Med. Inf.* **2021**, *152*, 104487. [CrossRef]
12. Colley, R.C.; Butler, G.; Garriguet, D.; Prince, S.A.; Roberts, K.C. Comparison of self-reported and accelerometer-measured physical activity in Canadian adults. *Health Rep.* **2018**, *29*, 3–15.
13. Alcazar, J.; Kamper, R.S.; Aagaard, P.; Haddock, B.; Prescott, E.; Ara, I.; Suetta, C. Relation between leg extension power and 30-s sit-to-stand muscle power in older adults: Validation and translation to functional performance. *Sci. Rep.* **2020**, *10*, 16337. [CrossRef]
14. Podsiadlo, D.; Richardson, S. The timed «Up & Go»: A test of basic functional mobility for frail elderly persons. *J. Am. Geriatr. Soc.* **1991**, *39*, 142–148. [PubMed]

15. Ng, S.S.; Hui-Chan, C.W. The Timed Up & Go Test: Its Reliability and Association With Lower-Limb Impairments and Locomotor Capacities in People With Chronic Stroke. *Arch. Phys. Med. Rehabil.* **2005**, *6*, 1641–1647.
16. Gordt, K.; Gerhardy, T.; Najafi, B.; Schwenk, M. Effects of Wearable Sensor-Based Balance and Gait Training on Balance, Gait, and Functional Performance in Healthy and Patient Populations: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. *Gerontology* **2018**, *64*, 74–89. [[CrossRef](#)] [[PubMed](#)]
17. Compagnat, M.; Batcho, C.S.; David, R.; Vuillerme, N.; Salle, J.Y.; Daviet, J.C.; Mandigout, S. Validity of the Walked Distance Estimated by Wearable Devices in Stroke Individuals. *Sensors* **2019**, *19*, 2497. [[CrossRef](#)]
18. Millor, N.; Lecumberri, P.; Gomez, M.; Martinez-Ramirez, A.; Izquierdo, M. Kinematic parameters to evaluate functional performance of sit-to-stand and stand-to-sit transitions using motion sensor devices: A systematic review. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2014**, *22*, 926–936. [[CrossRef](#)]
19. Salarian, A.; Horak, F.B.; Zampieri, C.; Carlson-Kuhta, P.; Nutt, J.G.; Aminian, K. iTUG, a sensitive and reliable measure of mobility. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2010**, *18*, 303–310. [[CrossRef](#)]
20. Cobo, A.; Villalba-Mora, E.; Hayn, D.; Ferre, X.; Pérez-Rodríguez, R.; Sánchez-Sánchez, A.; Bernabe-Espiga, R.; Sanchez-Sanchez, J.L.; Lopez-Diez-Picazo, A.; Moral, C.; et al. Portable Ultrasound-Based Device for Detecting Older Adults' Sit-to-Stand Transitions in Unsupervised 30-Second Chair-Stand Tests. *Sensors* **2020**, *20*, 1975. [[CrossRef](#)]
21. Fudickar, S.; Pauls, A.; Lau, S.; Hellmers, S.; Gebel, K.; Diekmann, R.; Bauer, J.M.; Hein, A.; Koppelin, F. Measurement System for Unsupervised Standardized Assessments of Timed Up and Go Test and 5 Times Chair Rise Test in Community Settings—A Usability Study. *Sensors* **2022**, *22*, 731. [[CrossRef](#)]
22. Hardy, R.; Cooper, R.; Shah, I.; Harridge, S.; Guralnik, J.; Kuh, D. Is chair rise performance a useful measure of leg power? *Aging Clin. Exp. Res.* **2010**, *22*, 412–418. [[CrossRef](#)]
23. Cooper, R.; Kuh, D.; Hardy, R.; Mortality Review Group; FALCon and HALCyon Study Teams. Objectively measured physical capability levels and mortality: Systematic review and meta-analysis. *BMJ* **2010**, *341*, c4467. [[CrossRef](#)]
24. Stephens, M.A. Introduction to Kolmogorov (1933) On the Empirical Determination of a Distribution. In *Breakthroughs in Statistics: Methodology and Distribution [Internet]*; Springer Series in Statistics; Kotz, S., Johnson, N.L., Eds.; Springer: New York, NY, USA, 1992; pp. 93–105. [[CrossRef](#)]
25. Smirnov, N. Table for Estimating the Goodness of Fit of Empirical Distributions. *Ann. Math. Stat.* **1948**, *19*, 279–281. [[CrossRef](#)]
26. Walter, S.D.; Eliasziw, M.; Donner, A. Sample size and optimal designs for reliability studies. *Stat. Med.* **1998**, *17*, 101–210. [[CrossRef](#)]
27. Schober, P.; Boer, C.; Schwarte, L.A. Correlation Coefficients: Appropriate Use and Interpretation. *Anesth. Analg.* **2018**, *126*, 1763–1768. [[CrossRef](#)] [[PubMed](#)]
28. Shrout, P.E.; Fleiss, J.L. Intraclass correlations: Uses in assessing rater reliability. *Psychol. Bull.* **1979**, *86*, 420–428. [[CrossRef](#)] [[PubMed](#)]
29. Terwee, C.B.; Bot, S.D.M.; de Boer, M.R.; van der Windt, D.A.W.M.; Knol, D.L.; Dekker, J.; Bouter, L.M.; de Vet, H.C. Quality criteria were proposed for measurement properties of health status questionnaires. *J. Clin. Epidemiol.* **2007**, *60*, 34–42. [[CrossRef](#)]
30. Giavarina, D. Understanding Bland Altman analysis. *Biochem. Med.* **2015**, *25*, 141–151. [[CrossRef](#)]
31. Beshara, P.; Chen, J.F.; Read, A.C.; Lagadec, P.; Wang, T.; Walsh, W.R. The Reliability and Validity of Wearable Inertial Sensors Coupled with the Microsoft Kinect to Measure Shoulder Range-of-Motion. *Sensors* **2020**, *20*, 7238. [[CrossRef](#)]
32. Cavazzana, A.; Röhrborn, A.; Garthus-Niegel, S.; Larsson, M.; Hummel, T.; Croy, I. Sensory-specific impairment among older people. An investigation using both sensory thresholds and subjective measures across the five senses. *PLoS ONE* **2018**, *13*, e0202969. [[CrossRef](#)]
33. Reuter, E.M.; Voelcker-Rehage, C.; Vieluf, S.; Godde, B. Touch perception throughout working life: Effects of age and expertise. *Exp. Brain Res.* **2012**, *216*, 287–297. [[CrossRef](#)]
34. Hetzler, R.K.; Stickley, C.D.; Lundquist, K.M.; Kimura, I.F. Reliability and Accuracy of Handheld Stopwatches Compared With Electronic Timing in Measuring Sprint Performance. *J. Strength Cond. Res.* **2008**, *22*, 1969. [[CrossRef](#)] [[PubMed](#)]
35. Moore, A.N.; Decker, A.J.; Baarts, J.N.; Dupont, A.M.; Epema, J.S.; Reuther, M.C.; Houser, J.J.; Mayhew, J.L. Effect of competitiveness on forty-yard dash performance in college men and women. *J. Strength Cond. Res.* **2007**, *21*, 385–388. [[PubMed](#)]
36. Falter, M.; Budts, W.; Goetschalckx, K.; Cornelissen, V.; Buys, R. Accuracy of Apple Watch Measurements for Heart Rate and Energy Expenditure in Patients With Cardiovascular Disease: Cross-Sectional Study. *JMIR mHealth uHealth* **2019**, *7*, e11889. [[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.