TESIS DE LA UNIVERSIDAD DE ZARAGOZA
2021
242

Juan de la Torre Pardos

Mejora de la aplicación clínica de métodos e instrumentación para evaluación de trastornos del equilibrio: Monitorización de pacientes y apoyo al diagnóstico por medio de modelos predictivos

Director/es
Marín Zurdo, José Javier
Gómez Trullén, Eva M

ISSN 2254-7606

Prensas de la Universidad
Universidad Zaragoza

ISSN 2254-7606

Universidad Zaragoza
1542

Tesis Doctoral

MEJORA DE LA APLICACIÓN CLÍNICA DE MÉTODOS E INSTRUMENTACIÓN PARA EVALUACIÓN DE TRASTORNOS DEL EQUILIBRIO: MONITORIZACIÓN DE PACIENTES Y APOYO AL DIAGNÓSTICO POR MEDIO DE MODELOS PREDICTIVOS

Autor

Juan de la Torre Pardos

Director/es

Marín Zurdo, José Javier
Gómez Trullén, Eva M

UNIVERSIDAD DE ZARAGOZA
Escuela de Doctorado

2021

Universidad Zaragoza
1542

Tesis Doctoral

MEJORA DE LA APLICACIÓN CLÍNICA DE MÉTODOS E INSTRUMENTACIÓN PARA EVALUACIÓN DE TRASTORNOS DEL EQUILIBRIO: MONITORIZACIÓN DE PACIENTES Y APOYO AL DIAGNÓSTICO POR MEDIO DE MODELOS PREDICTIVOS

Autor

Juan de la Torre Pardos

Director/es

Marín Zurdo, José Javier
Gómez Trullén, Eva Mª

Instituto de Investigación en Ingeniería de Aragón

Ingeniería Biomédica

2021

A Bea, por
creer en mí y
acompañarme durante
todo el camino.

“Ever tried. Ever failed.
No matter. Try again.
Fail again. Fail better”
Samuel Beckett

AGRADECIMIENTOS

En primer lugar, quiero agradecer a mis **directores de tesis**, **Eva Gómez** y **Javier Marín**, su apoyo incondicional, sus consejos y sus conocimientos transmitidos durante todo este proceso. No hubiera sido posible realizar esta tesis sin vosotros.

Quiero agradecer también a todos los **compañeros del grupo de investigación** el apoyo que me han dado, compartiendo tantos proyectos, y ayudándome a superar obstáculos y frustraciones.

Por otra parte, también quiero agradecer a **Thomas Robert** su acogida y dedicación durante la estancia de investigación que realicé en el *Laboratoire de biomécanique et mécanique des chocs* en Lyon. También a los compañeros de laboratorio que tuve la suerte de conocer, los cuales me integraron como a uno más desde el primer día.

También quiero agradecer a los distintos **médicos y profesionales sanitarios** del Hospital Universitario Miguel Servet de Zaragoza, del Hospital Comarcal de Alcañiz, del Hospital MAZ de Zaragoza y del Instituto de Medicina Legal de Zaragoza, que han colaborado en las distintas investigaciones. Vuestro punto de vista y enfoque nos ha permitido conocer de primera mano los problema que debéis enfrentar en el día a día, así como las necesidades que existen en los distintos servicios del hospital.

Igualmente quiero agradecer a los **expertos** que han evaluado los distintos artículos que componen esta tesis, cuyos comentarios y revisiones han permitido elaborar unos trabajos de calidad y mejorar el conocimiento aportado a la sociedad con esta investigación.

No quiero pasar por alto lo mucho que me han ayudado las personas que me han acompañado en mi día a día durante esta etapa de mi vida. Estoy muy agradecido a mis **amigos**, los cuales nunca han dudado que lograría terminar la tesis, y que siempre me han acompañado tanto en los buenos momentos, como en los que no lo han sido tanto.

Gracias también a **mis padres Juan y Gloria, a mi hermana María, y a mi tía Chus**. Los cuales, pese a no terminar de entender completamente “de que iba" esta tesis, siempre me han apoyado y nunca me han dejado caer. Gracias por creer en mí.

También quiero agradecer a **Molly** su apoyo invisible durante estos años, siempre consigues sacarme una sonrisa, y me alegras hasta en los momentos más difíciles.

Y por supuesto a ti, **Bea**. Gracias por creer en mí y compartir tu vida conmigo.

ÍNDICE

PRESENTACIÓN

La presente tesis doctoral, con título "Mejora de la aplicación clínica de métodos e instrumentación para evaluación de trastornos del equilibrio: monitorización de pacientes y apoyo al diagnóstico por medio de modelos predictivos", comprende un conjunto de trabajos científicos publicados.

Los artículos recogen los resultados del conjunto de investigaciones y desarrollos llevados a cabo para mejorar la evaluación de los trastornos del equilibrio, así como el tratamiento de la información generada en el campo de la salud mediante el uso de modelos predictivos.

La tesis se estructura de forma secuenciada. En primer lugar, la motivación de la investigación llevada a cabo, seguida por una introducción donde se incluyen los diferentes temas abordados. A continuación, se presenta la justificación de la misma, así como los diferentes objetivos e hipótesis, seguido por un resumen de las diferentes acciones de investigación. Tras ello, se presentan los artículos científicos publicados, con un resumen de los mismos, así como una copia de los propios trabajos científicos. Finalmente, se presenta una discusión de los objetivos de la tesis, así como un apartado de conclusiones globales. Los materiales, los métodos, los resultados y las discusiones particulares de cada investigación, quedan expuestos mediante la inclusión de dichos trabajos.

Los resultados obtenidos gracias a la realización de estos estudios han aportado información relevante y novedosa sobre el tema y han sido recogidos en cinco artículos originales, publicados en diferentes revistas de amplia difusión internacional:

Artículo 1

De la Torre, J.; Marin, J.; Marin, J. J.; Auria, J. M.; Sanchez-Valverde, M. B. **Balance study in asymptomatic subjects: Determination of significant variables and reference patterns to improve clinical application**. *J. Biomech*. **2017**, *65*, 161–168, doi:10.1016/j.jbiomech.2017.10.013.

Factor de Impacto JCR (2019):

2.320

Áreas Temáticas:

Biomedical Engineering: Q3 (48/87)

Biophysics: Q3 (38/71)

Artículo 2

De la Torre, J.; Marin, J.; Polo, M.; Marín, J.J. **Applying the Minimal Detectable Change of a Static and Dynamic Balance Test Using a Portable Stabilometric Platform to Individually Assess Patients with Balance Disorders**. *Healthcare (S.I. Ageing Effects on Kinematics, Kinetics and Balance)* **2020**, *8*, 402, https://doi.org/10.3390/healthcare8040402.

Factor de Impacto JCR (2019):

1.916

Áreas Temáticas:

Health Care Sciences & Services: Q3 (62/102)

Artículo 3

De la Torre J, Marin J, Polo M, Gómez-Trullén EM, Marin JJ. **2021. MCQ-Balance: a method to monitor patients with balance disorders and improve clinical interpretation of posturography.** PeerJ 9:e10916 https://doi.org/10.7717/peerj.10916

Factor de Impacto JCR (2019):

2.379

Áreas Temáticas:

Multidisciplinary Sciences: Q2 (32/71)

Artículo 4

De la Torre, J.; Marin, J.; Ilarri, S.; Marin, J.J. **Applying Machine Learning for Healthcare: A Case Study on Cervical Pain Assessment with Motion Capture**. *Appl. Sci (S.I. Medical Informatics and Data Analysis)* **2020**, *10*, 5942, doi: 10.3390/app10175942.

Factor de Impacto JCR (2019):

2.474

Áreas Temáticas:

Multidisciplinary Engineering: Q2 (32/91)

Applied Physics: Q2 (63/155)

Multidisciplinary Materials Science: Q3 (161/314)

Multidisciplinary Chemistry: Q2 (88/177)

Artículo 5

De la Torre, J.; Bonnet, V.; Mauti, R.; Chabaud, P.; Robert, T. **Implementing a test to assess reaction, attention and inhibition capacity in elderly**. *Computer Methods in Biomechanics and Biomedical Engineering*, 2020, vol. 23 (sup1), S294-S296, doi: 10.1080/10255842.2020.1815313.

Factor de Impacto SJR (2019):

0.45

Áreas Temáticas:

Bioengineering: Q3

Biomedical Engineering: Q3

Computer Sciences Applications: Q2

Human-Computer Interaction: Q2

Medicine (miscellaneous): Q2

RESUMEN

Entre las patologías más frecuentes en la actualidad, destacan los trastornos del equilibrio, teniendo especial prevalencia en personas mayores y en países desarrollados. Los trastornos del equilibrio son una patología que se caracteriza principalmente por síntomas de vértigo y mareo, y se calcula que un cuarto de la población tendrá un episodio vertiginoso en su vida. Las caídas y las lesiones inducidas por las mismas son problemas de salud pública de primer orden a nivel mundial. De manera complementaria, también existen alteraciones o patologías que pueden afectar a los sistemas sensoriales que sustentan el equilibrio, existiendo la necesidad de monitorizarlos para evaluar su progresión en el tiempo. Sin embargo, no resulta sencillo obtener información objetiva del diagnóstico o progresión de un paciente afectado por trastornos del equilibrio; destacando la complejidad para obtener informes concisos con información comprensible y fácil de interpretar por facultativos sanitarios, que deben prescribir tratamientos a dichos pacientes.

Asimismo, y en relación con la necesidad de obtener métodos o procesos para poder aprovechar mejor la información en el campo de la salud, destaca el incremento exponencial de los datos que los centros sanitarios generan y requieren manejar, así como la compleja gestión de dicha información, siendo la disciplina científica *Machine Learning* una posible solución para abordar esta problemática. Sin embargo, aplicar las técnicas de este tipo de disciplina en el campo de la salud conlleva un alto grado de complejidad.

La presente tesis pretende dar respuesta a las incertidumbres y problemáticas expuestas a través de varias investigaciones que han tratado de proporcionar información objetiva sobre el estado de pacientes que presentan trastornos del equilibrio, así como monitorizar este tipo de patologías mediante diversos métodos e instrumentación; igualmente, se han definido los procesos y particularidades para la generación de modelos predictivos como herramienta de apoyo al diagnóstico en el ámbito de la salud. En concreto, los objetivos de la tesis son: generar y proporcionar información objetiva sobre el estado de pacientes que presentan trastornos del equilibrio; monitorizar y analizar la progresión de este tipo de patologías en pacientes inmersos en tratamientos de rehabilitación; generar modelos predictivos para mejora del diagnóstico en el ámbito de la salud.

Esta tesis se ha complementado con una estancia de investigación de cuatro meses en el *Laboratoire de biomécanique et mécanique des chocs* (Université Gustave Eiffel) en Lyon (Francia).

El cuerpo principal de esta tesis consiste en una compilación de cinco artículos, los cuales han sido publicados en revistas científicas indexadas en el JCR.

ABSTRACT

Among the most common pathologies today, balance disorders stand out, having a special prevalence in older people and in developed countries. Balance disorders are pathologies that are characterized mainly by symptoms of vertigo and dizziness, and it is estimated that a quarter of the population will have a vertiginous episode in their lifetime. Falls and fall-induced injuries are major public health problems worldwide. In a complementary way, there are also alterations or pathologies that can affect the sensory systems that support balance, and there is a need to monitor them to assess their progression over time. However, it is not easy to obtain objective information on the diagnosis or progression of a patient with balance disorders, highlighting the complexity of obtaining concise reports with information easily interpretable by health professionals who must prescribe treatments for these patients.

Likewise, and in relation to the need to obtain methods or processes to be able to take better advantage of the information in the health field, the exponential increase in the data that health centres generate and require handling, as well as the complex management of the information, being the scientific discipline of *Machine Learning* a possible solution to address this problem. However, applying the techniques of this type of discipline in the health field involves a high degree of complexity.

The present thesis aims to respond to the uncertainties and problems exposed through certain investigations that have tried to provide objective information on the state of patients with balance disorders, as well as to monitor this type of pathologies through various methods and instrumentation; Likewise, the processes and particularities have been defined for the generation of predictive models as a tool to support diagnosis in the health field. Specifically, the objectives of the thesis are: to generate and provide objective information on the state of patients with balance disorders; monitor and analyse the progression of this type of pathology in patients immersed in rehabilitation treatments; and generate predictive models to improve diagnosis in the health field.

This thesis was complemented by a four-month research stay at the L*aboratoire de biomécanique et mécanique des chocs* (Université Gustave Eiffel) in Lyon (France).

The main body of this thesis consists of a compilation of five articles, which have been published in scientific journals indexed in the JCR.

ABREVIATURAS

ANOVA: Analysis of Variance (análisis de varianza)

AP: Antero-Posterior (dirección anteroposterior)

BMI: Body Mass Index (índice de masa corporal)

BPPV: Benign Paroxysmal Peripheral Vertigo (vertigo periférico paroxistico benigno)

BSS: Balance Sensory System (sistemas sensoriales que sustentan el equilibrio)

CEICA: Research Ethics Committee of the Community of Aragon (comité de bioética de Aragón)

COP: Centre Of Pressure (centro de presiones)

CQ: Quantification of the Change (cuantificación del cambio)

CRISP-DM: CRoss-Industry Standard Process for Data Mining (proceso de minería de datos CRoss-Industry Standard)

CSRT: Choice Stepping Reaction Time Test (test Choice Stepping Reaction Time)

CTSIB-M: Modified Clinical Test of Sensory Interaction in Balance (test de interacción del equilibrio)

DLA: Daily life activities (actividades de la vida diaria)

DM: Data Mining (minería de datos)

EMG: Surface Electromyography (electromiografía de superficie)

ES: Eye-Sight System (Sistema visual)

FL: Flanker Task (test de Flanker)

GBA: Gradient Boosting Algorithm (algoritmo de Machine Learning)

ICC: Intraclass Correlation Coefficient (índice de correlación intraclase)

IDERGO: Research and Development in Ergonomics (grupo de investigación de investigación y desarrollo en ergonomía)

ISPGR: International Society for Posture and Gait Research (sociedad internacional de la investigación sobre el equilibrio y la marcha)

κ: Cohen's Kappa Statistical Coefficient (coeficiente Kappa)

KDD: Knowledge Discovery in Databases (método de Machine Learning)

KNN: K-Nearest Neighbors (algoritmo de Machine Learning)

LOS: Limits of Stability (límites de estabilidad)

MAE: Mean Absolute Error (error medio)

MBD: Magnitude-Based Decision (método de decisión basado en magnitudes)

MCQ: Measure, Classify and Qualify (medir, clasificar y cualificar)

MDC: Minimal Detectable Change (mínimo cambio detectable)

MID: Minimal Important Difference (minima diferencia importante)

ML: Machine Learning (aprendizaje automático)

MLP: MultiLayer Perceptron (perceptrón de varias capas)

MoCap: Motion Capture (captura de movimiento)

PCA: Principal component analysis (algoritmo de Machine Learning)

PM&R: Physical Medicine and Rehabilitation Service (servicio de rehabilitación)

PoC: Probability of Change (probabilidad de cambio)

PS: Proprioceptive System (sistema propioceptivo)

RMS: Root Mean Square (raíz cuadrada media)

RMSE: Root Mean Squared Error (error de raíz cuadrada media)

ROM: Range of Movement (rango de movimiento)

RSEC: Rigid Surface with Eyes Closed (superificie rígida con ojos cerrados)

RSEO: Rigid Surface with Eyes Open (superificie rígida con ojos abiertos)

RWS: Rhythmic Weight Shift (prueba de control rítmico direccional)

SD: Standard Deviation (desviación estándar)

SDpool: Pooled Average between the Standard Deviation of the Test and Retest (desviación estándar promedio entre el test-retest)

SEM: Standard Error of Measurement (error standard de medida)

SEMMA: Sample, Explore, Modify, Model, and Assess (muestrear, explorar, modificar, modelar y evaluar)

SSEC: Soft Surface with Eyes Closed (superificie blanda con ojos cerrados)

SSEO: Soft Surface with Eyes Open (superificie blanda con ojos abiertos)

ST: Stepping Time (tiempo de paso)

STTE: Short-Term Typical Error (error típico)

SVM: Support Vector Machine (algoritmo de Machine Learning)

VS: Vestibular System (Sistema vestibular)

WDQ: Whiplash Scale (escala de latigazo cervical)

Xdif: Difference Between the measures taken in two Temporal Point (diferencia entre dos medidas tomadas en diferentes instantes temporales)

LISTA DE FIGURAS

LISTA DE TABLAS

1. INTRODUCCIÓN

1.1. MOTIVACIÓN DE LA INVESTIGACIÓN

Entre las patologías más frecuentes en la actualidad destacan los trastornos del equilibrio, con el vértigo y el mareo como síntomas cardinales, teniendo especial prevalencia en personas mayores y en países desarrollados (Rubenstein & Josephson, 2002). Se calcula que un cuarto de la población tendrá un episodio vertiginoso en su vida. Las caídas y las lesiones inducidas por las mismas son problemas de salud pública de primer orden a nivel mundial, ya que alrededor de un 30% de las personas mayores de 65 años, y más del 50% de las personas en centros sanitarios o residencias, sufren una o más caídas al año (Tinetti, 2003). Además, el envejecimiento, las alteraciones sensoriales, los trastornos musculoesqueléticos o neurológicos, las enfermedades cardiovasculares, las infecciones o las alteraciones del metabolismo con frecuencia también están relacionados con los trastornos del equilibrio (Ma, Wong, Lam, Wan, & Lee, 2016). Asimismo, el equilibrio también está relacionado con el deterioro físico derivado del envejecimiento, y conectado con el concepto de *fragilidad*. La fragilidad es un síndrome geriátrico resultante de la edad, relacionada con la disminución acumulativa en múltiples sistemas fisiológicos, reserva homeostática deteriorada y capacidad reducida para resistir el estrés. La evaluación del grado de fragilidad de un sujeto resulta de interés para procurar un envejecimiento activo ligada con la línea de investigación europea relativa al término *"Healthy Ageing"* (Mohler, Wendel, Taylor-Piliae, Toosizadeh, & Najafi, 2016).

De manera complementaria, también existen alteraciones o patologías que pueden afectar a los *Balance Sensory Systems* (BSS: sistemas sensoriales del equilibrio), existiendo la necesidad de monitorizarlos (realizar un seguimiento) para evaluar su progresión en el tiempo (Shumway-Cook, Brauer, & Woollacott, 2000). Entre las patologías existentes que pueden afectar a dichos sistemas, destacan las siguientes: alteraciones neurológicas, como son los accidentes cerebrovasculares, lesiones medulares incompletas, enfermedades desminilizantes; cervicalgias de distinto origen, patologías vestibulares uni- o bilaterales o incluso patologías visuales. Debido a la existencia de tan diversos orígenes, es necesario monitorizar la progresión de dichas patologías que afectan al equilibrio de los pacientes a lo largo del tiempo, a fin de comprobar los efectos de los tratamientos de rehabilitación en cada paciente.

Sin embargo, no resulta sencillo obtener información objetiva del diagnóstico o progresión de un paciente afectado de trastornos del equilibrio, destacando la complejidad para obtener informes concisos con información comprensible e interpretable por facultativos sanitarios que deben prescribir tratamientos a dichos pacientes. Esto es debido a la falta de estandarización y a la necesidad de adaptación de los equipos, como lo posturografía, a las necesidades clínicas (Visser, Carpenter, van der Kooij, & Bloem, 2008; Von Lubitz & Wickramasinghe, 2006).

De igual modo, y en relación con la necesidad de obtener métodos o procesos para poder aprovechar mejor la información en el campo de la salud, destaca el incremento exponencial de los datos que los centros sanitarios generan y requieren manejar, así como la compleja gestión de dicha información, siendo la disciplina científica *Machine Learning* (ML) una posible solución para abordar esta problemática. Sin embargo, aplicar las técnicas de este tipo de disciplina en el campo de la salud conlleva un alto grado de complejidad, entre otras razones, por la gestión de la información personal y sensible de pacientes, así como de las técnicas necesarias para poder hacer uso de dicha información (Kayyali, Knott, & Van Kuiken, 2013; Koh & Tan, 2011; Tomar & Agarwal, 2013).

La presente tesis pretende dar respuesta a las incertidumbres y problemáticas planteadas a través de ciertas investigaciones que han tratado de proporcionar información objetiva sobre el estado de pacientes que presentan trastornos del equilibrio, monitorizar este tipo de patologías mediante diversos métodos e

instrumentación y definir los procesos y particularidades para la generación de modelos predictivos como herramienta de apoyo al diagnóstico en el ámbito de la salud.

Las investigaciones realizadas durante el desarrollo de esta tesis doctoral, se han llevado a cabo en el seno del grupo de investigación IDERGO (Investigación y Desarrollo en Ergonomía) de la Universidad de Zaragoza. IDERGO es un grupo de investigación en el área de la tecnología reconocido como grupo consolidado de investigación aplicada por el Gobierno de Aragón desde 2008. En la actualidad, es un grupo de investigación de referencia reconocido por el Gobierno de Aragón en la Resolución del 13 de marzo de 2020, durante el periodo 2020-2022 (ref: T38_20R). Pertenece al Instituto Universitario de Investigación en Ingeniería de Aragón de la Universidad de Zaragoza (I3A), encuadrado en el área de ingeniería biomédica.

Las investigaciones realizadas en esta tesis han sido posible por la colaboración entre los investigadores y doctorandos del grupo IDERGO, así como por el trabajo colaborativo con otros grupos de investigación (AMB de la Universidad de Zaragoza) y diferentes hospitales de referencia de la Comunidad de Aragón (Hospital Universitario Miguel Servet, Hospital Comarcal de Alcañiz, Hospital MAZ de Zaragoza y el Instituto de Medicina Legal de Zaragoza).

Asimismo, parte de los desarrollos y las experimentaciones llevadas a cabo, han sido financiados por el Fondo Social Europeo (FSE) del Gobierno de Aragón, así como el doctorando ha estado contratado como investigador en la modalidad N3, durante las etapas principales de desarrollo de la presente tesis.

1.2. ANTECEDENTES

1.2.1. TRASTORNOS DEL EQUILIBRIO E IMPACTO

El control postural es esencial cuando se realizan actividades cotidianas, y sus posibles trastornos tienen un efecto muy significativo en la autonomía personal, ya que son indicativos de patologías físicas o neuronales incipientes (García, Corresa, Bertomeu, & Suárez-Varela, 2012). Estos trastornos son disfunciones comunes encontradas tanto por médicos generales como por especialistas (Baydal-Bertomeu et al., 2004). Estos trastornos, junto con los trastornos de la marcha, representan la segunda causa principal de caídas después de los accidentes (Rubenstein & Josephson, 2002). Las caídas y las lesiones provocadas por ellas corresponden a problemas de salud pública muy importantes a nivel global. Alrededor del 30% de las personas mayores de 65 años y más del 50% de las personas en centros de salud u hogares de ancianos sufren / sufrirán una o más caídas cada año; además, aproximadamente entre el 20% y el 30% de la población mundial ha tenido o tendrá un episodio vertiginoso en su vida (Baydal-Bertomeu et al., 2004; Lin, Seol, Nussbaum, & Madigan, 2008; Ma et al., 2016; Salehi, Ebrahimi-Takamjani, Esteki, Maroufi, & Parnianpour, 2010; Sánchez-Sánchez et al., 2018; Tinetti, 2003; Wolf et al., 1996a). Asimismo, el envejecimiento, los cambios sensoriales, los trastornos musculoesqueléticos o neurológicos, las enfermedades cardiovasculares, las infecciones o los trastornos metabólicos están estrechamente relacionados con los trastornos del equilibrio (Ma et al., 2016).

Los trastornos del equilibrio son una de las patologías más comunes entre las personas mayores en los países desarrollados (Penger, Strobl, & Grill, 2017). En combinación con el aumento gradual en el índice de envejecimiento de la población (Gavrilov & Heuveline, 2003; Muir, Kiel, Hannan, Magaziner, & Rubin, 2013; Vaupel & Loichinger, 2006), ha llevado a un aumento de las patologías que afectan equilibrio, causando un mayor riesgo de caídas en la población de edad avanzada (Aftab, Robert, & Wieber, 2016; Khalaj, Osman, Mokhtar, Mehdikhani, & Abas, Wan Abu Bakar Wan, 2014).

El vértigo se define como una ilusión de movimiento, ya sea del mundo externo que gira en torno al individuo o del individuo que gira en el espacio (Medical Subject Headings (MeSH), 2020). Es el síntoma cardinal de los trastornos del equilibrio y suele asociarse a los síntomas del síncope vasovagal, lo que conlleva una reducción significativa de la calidad de vida y un aumento de la discapacidad, la ansiedad y la depresión (Neuhauser, 2016). Del porcentaje estimado de la población mundial que tiene / tendrá un episodio vertiginoso de diversos orígenes y gravedad a lo largo de la vida (da Costa Barbosa & Vieira, 2017; Lin et al., 2008; Tinetti, 2003; Wolf et al., 1996a), el 20% de ellos no reciben un diagnóstico claro (Swanenburg, de Bruin, Favero, Uebelhart, & Mulder, 2008).

El vértigo suele ser causado por una disfunción resultante de una lesión periférica o central (Stanton M, 2020 Apr 28); por tanto, dependiendo del origen, se puede clasificar como vértigo de origen periférico o central (Baumgartner B, 2019 Jun 3; Strupp, Dieterich, & Brandt, 2013). El vértigo periférico se caracteriza por crisis repentinas de corta duración, a veces con sintomatología auditiva asociada y con un cortejo vegetativo mayor que en el vértigo central, estando representado principalmente por el vértigo de origen vestibular (Wipperman, 2014). Igualmente, el vértigo también puede tener un origen cervical en los casos en que el sistema nervioso central no se encuentre afectado (excepto el síndrome de la médula espinal), incluyéndose como vértigo periférico, con una relación demostrada entre cervicalgia y vértigo. El vértigo de origen cervical se identifica con mayor frecuencia en las consultas de rehabilitación, y ocurre con mayor frecuencia en pacientes de entre 30 y 50 años (Dieterich & Eckhardt-Henn, 2004; Reiley, Vickory, Funderburg, Cesario, & Clendaniel, 2017). Por el contrario, el vértigo central suele ser continuo y

prolongado con el tiempo. En general, su apariencia es progresiva, con inestabilidad asociada, cortejo vegetativo leve y recuperación lenta. Se asocia con vértigo de origen neurológico (Solomon, 2000).

En línea con lo expuesto, en relación al vértigo de origen cervical, una parte importante de las alteraciones del equilibrio provienen de trastornos originados en dicha zona. Los trastornos en la columna cervical tienen una alta incidencia y prevalencia, y se consideran un problema de salud pública, especialmente en los países desarrollados. Asimismo, las lesiones cervicales (ocasionadas generalmente por latigazo cervical tras un accidente de tráfico) tienen un diagnóstico difícil porque las lesiones traumáticas de la columna cervical y sus síntomas asociados son diversos (Alejandro J. Moreno, Gonzalo Utrilla, Javier Marin, Jose J. Marin, Maria B. Sanchez-Valverde, Ana C. Royo., 2017).

1.2.2. SISTEMAS SENSORIALES DEL EQUILIBRIO

El diagnóstico del vértigo requiere medir el grado de alteración, de forma aislada o combinada, de cada sistema sensorial del equilibrio (BSS). Concretamente, los sistemas que sustentan el equilibrio son: el sistema de la vista (ES), el sistema vestibular (VS) y el sistema propioceptivo (PS). Un episodio vertiginoso, mareo o episodios traumático puede afectar a estos sistemas en mayor o menor medida y, en consecuencia, al equilibrio del paciente (Derebery, 2000; Hickey, Ford, Buckley, & O'connor, 1990). Por lo tanto, es necesario tener métodos o indicadores para determinar cómo progresa cada BSS, especialmente, durante un tratamiento de una patología del equilibrio (Shumway-Cook et al., 2000). Los efectos de un tratamiento dado en un paciente a través de la realización de pruebas previas y posteriores pueden ayudar a tomar decisiones sobre ajustar, cambiar o suspender el tratamiento (Baker, 2006).

Cuando es difícil establecer una patología clara relacionada con cualquiera de los BSS, o cuando se encuentran múltiples orígenes de la afección, el diagnóstico clínico se complica (Derebery, 2000; Swanenburg et al., 2008), requiriéndose nuevas medidas y pruebas adicionales que proporcionen información relevante al facultativo sanitario.

En este sentido, se utilizan con frecuencia ciertas evaluaciones funcionales del equilibrio, como el test de Unterberger (Hickey et al., 1990), el test Up and Go (Martínez Carrasco, 2016; Shumway-Cook et al., 2000) y el test de apoyo unipodal (Vellas et al., 1997). Sin embargo, la fiabilidad y alcance de estos tests funcionales no es concluyente (Martínez Carrasco, 2016).

1.2.3. MEDICIÓN DEL EQUILIBRIO

Una evaluación del equilibrio implica medir el desplazamiento del centro de presiones (COP), y permite cuantificar el control postural durante la posición de pie, sentado y caminando. Se pueden utilizar diferentes pruebas de evaluación no invasivas que permiten objetivar el grado de control del equilibrio estático y dinámico (Duarte & Freitas, 2010). En los métodos de control de equilibrio estático, entre los que destaca el test de Romberg, los sujetos deben mantener su COP dentro de la base de apoyo durante todo el período de evaluación. Las evaluaciones del equilibrio postural dinámico, que es trascendental para el control motor, implican la medición de los límites de estabilidad (LOS), que corresponden al ángulo o distancia máxima voluntaria en la que un individuo puede regular su COP en una dirección determinada sin perder el equilibrio (Ku, Abu Osman, & Wan Abas, 2016). Estas son pruebas que pueden diagnosticar, medir la evolución de un tratamiento o incluso servir como un medio de terapia de reeducación postural (García et al., 2012). Además, la combinación de pruebas se realiza porque permite evaluar la consistencia y, en consecuencia, el grado de colaboración del paciente al completar las pruebas, lo cual es relevante en entornos médico - legales y forenses (Ramírez, Ordi, Fernández, & MI, 2014).

La investigación de los mecanismos involucrados en el equilibrio y el control postural, ha despertado el interés de profesionales de varios campos distintos a la medicina, como la fisioterapia, el deporte, la ingeniería, la física, la biomecánica, la psicología (Duarte & Freitas, 2010), el entretenimiento y el ocio (Weaver, Ma, & Laing, 2016). Sin embargo, estos profesionales utilizan una variedad de técnicas de medición y evaluación que a menudo generan resultados diferentes, con opiniones divergentes con respecto a cuestiones tales como: la duración de la prueba, la posición de los pies, qué variables analizar, o las condiciones del entorno (Cvecka, Cacek, & Ondracek, 2014; Kapteyn et al., 1983; Pinsault & Vuillerme, 2009; Scoppa, Capra, Gallamini, & Shiffer, 2013; Solovykh, Maksimovskaya, Bugrovetskaya, & Bugrovetskaya, 2011).

1.2.4. POSTUROGRAFÍA Y PROBLEMAS INHERENTES

Como alternativa o complemento a las pruebas funcionales, la posturografía se presenta como una prueba de uso habitual en clínica. Mediante el uso de una plataforma estabilométrica, la posturografía permite medir los movimientos del COP en posición de bipedestación. Constituye una prueba de diagnóstico con validez médico - legal que proporciona información objetiva sobre las patologías del equilibrio en la práctica clínica (De la Torre, Marin, Marin, Auria, & Sanchez-Valverde, 2017; Dounskaia, Peterson, & Bruhns, 2018; Lin et al., 2008). Aunque es un método validado y *Gold-Standard*, se encuentran dificultades para discernir el origen causado por el patrón de desequilibrio. Esto se debe a que, aunque los análisis sensoriales sugieren un patrón propioceptivo-visual-vestibular, esto no siempre es exacto (El-Kashlan, Shepard, Asher, Smith-Wheelock, & Telian, 1998; Stewart et al., 1999; Timothy C. Hain, May 5, 2019).

Las plataformas estabilométricas son útiles para evaluar el equilibrio porque obtienen numerosos parámetros del COP (López & Calidonio, 2009; Postolache & Postolache, 2017). Evalúan el equilibrio estático y dinámico a través de diferentes variables y métodos de aplicación; así como pueden obtener información objetiva relacionada con las patologías del equilibrio en la práctica clínica para mejorar la calidad de la atención médica y los tratamientos proporcionados (de la Torre et al., 2017; García et al., 2012; Lin et al., 2008; Meshkati, Namazizadeh, Salavati, & Mazaheri, 2011).

Por lo tanto, una prueba de equilibrio sería beneficiosa para los facultativos al proporcionarles un dispositivo de evaluación objetivo; sin embargo, la integración de este tipo de pruebas en la práctica clínica implica dificultades y limitaciones. Las plataformas estabilométricas se usan normalmente en laboratorio, por lo que su implementación en la clínica es compleja debido a la dificultad de uso y manejabilidad, lo que puede cuestionar su propia aplicabilidad (Swanenburg et al., 2008). El protocolo debe ser validado y adaptado para su uso continuo y asiduo en clínica (Bauer, Gröger, Rupprecht, & Gaßmann, 2008; Swanenburg et al., 2008). Por lo tanto, es relevante realizar estudios para medir el grado de variabilidad de una prueba de equilibrio para lograr una aplicación clínica exitosa.

La International *Society for Posture and Gait Research* (ISPGR), en la reunión de Bolonia (2009) (Scoppa et al., 2013), logró cierta unificación de criterios. La ISPGR abordó el problema de la estandarización estabilométrica clínica, estableciendo ciertos criterios para obtener la precisión y sensibilidad adecuadas en la medición del COP. Posteriormente, se han celebrado reuniones de la ISPGR (Vancouver en 2014 y Sevilla en 2015), así como diferentes propuestas y recomendaciones, como la estandarización japonesa de la evaluación de la estabilometría clínica. Sin embargo, todavía hay aspectos sin acuerdo, como la postura que debe adoptar el paciente o los parámetros clínicos más importantes.

En relación con lo anterior, aunque los datos clínicos de la posturografía tradicional son útiles, en algunos casos son insuficientes, ya que se requiere de dispositivos más inteligentes para el diagnóstico (Allum, Zamani, Adkin, & Ernst, 2002; Di Fabio, 1996). Los dispositivos de posturografía pueden proporcionar información útil en la toma de decisiones clínicas a nivel individual de cada paciente; sin embargo, se les

exige que sean prácticos y manejables. Por ello, no deberían requerir de técnicos o facultativos expertos adicionales para su uso, ya que esto implicaría altos costes, y la prueba perdería practicidad para el uso clínico (Visser et al., 2008). Con respecto a esta punto, la información obtenida de los informes de posturografía debe ser clara, concisa y fácil de interpretar por los médicos, así como estar suficientemente respaldada para lograr un valor diagnóstico adecuado. Sin embargo, actualmente, los informes de posturografía no ofrecen suficiente profundidad para alcanzar el valor diagnóstico que sí tienen otras pruebas (por ejemplo, pruebas de imagen) (Climent Barbera JM, 2003; Visser et al., 2008).

Otros autores también han profundizado en esta idea y consideraron que el informe de posturografía debería proporcionar información concisa sobre el estado de equilibrio del paciente y, si es posible, del BSS que tiene la mayor influencia en el desequilibrio del paciente (Derebery, 2000). Del mismo modo, el informe debe contar con suficiente automatización para que no requiera de largos períodos de procesamiento para su análisis e interpretación. Debe ser compatible con un lenguaje comprensible sin los tecnicismos que generalmente acompañan a ciertos equipos médicos (Visser et al., 2008; Von Lubitz & Wickramasinghe, 2006). Del mismo modo, se requiere tanto la validación como la estandarización de los protocolos de reproducibilidad, e incluso una posible comparación con estudios similares (Visser et al., 2008).

Por otra parte, aunque se han aplicado numerosas pruebas de evaluación del equilibrio en diferentes estudios mediante plataforma estabilométrica (Karlsson & Frykberg, 2000), las puntuaciones resultantes de estos test son diversos, muchas veces complejas y difíciles de interpretar (Peterson, Ferrara, Mrazik, Piland, & Elliott, 2003). Además, existen dificultades en el uso de puntuaciones subjetivas debido a la falta de estandarización e interpretación, lo que hace que sea difícil diagnosticar trastornos del equilibrio, así como identificar el BSS que origina el patrón de desequilibrio (Jacobs, Horak, Tran, & Nutt, 2006; Saxena & Prabhakar, 2013; Visser et al., 2008).

1.2.5. MONITORIZACIÓN DEL EQUILIBRIO DE PACIENTES

En el campo de la rehabilitación, es fundamental medir la progresión entre dos sesiones separadas, así como caracterizar objetivamente la respuesta a los tratamientos para mejorar la toma de decisiones médicas (Hamburg & Collins, 2010). Con este objetivo, para monitorizar al paciente, es clave determinar si se han producido cambios relevantes en un paciente a nivel individual entre dos medidas obtenidas en diferentes instantes temporales (Hopkins, 2017; Visser et al., 2008). Al respecto, podemos destacar la propuesta de Hopkins (2017) para evaluar el cambio entre dos mediciones en un individuo a través del método *magnitude-based decision* (MBD) (Hopkins, 2017).

Para poder aplicar el MBD, es necesario determinar previamente el mínimo cambio detectable (MDC). El MDC representa la variabilidad de las medidas de cada variable. Es decir, la consistencia de las variables, que incluye diferentes aspectos tales como: la variabilidad del instrumento, la variabilidad inherente a la persona y a su aprendizaje (duración de la prueba, condición visual o posición de los pies) y los procedimientos y protocolos aplicados para realizar la prueba. Por lo tanto, si se obtiene un cambio de una variable que es mayor que su valor de MDC, representaría un cambio real, que podría ser causado por el efecto del tratamiento y/o por las fluctuaciones de la enfermedad, y no por la variabilidad de la prueba (de Sá Ferreira & Baracat, 2014; Ruhe, Fejer, & Walker, 2010).

Determinar el valor del MDC de las diferentes variables involucradas en una prueba de equilibrio permitiría la evaluación de la variabilidad derivada del instrumento y / o del protocolo y la identificación de las variables más fiables. Todo esto permitiría reducir el número de variables evaluadas, facilitando la toma de decisiones médicas (Donoghue & Stokes, 2009; Kovacs et al., 2008; Lee et al., 2013; Marchetti, Lin, Alghadir, & Whitney, 2014; Schuck & Zwingmann, 2003; Steffen & Seney, 2008).

1.2.6. MODELOS PREDICTIVOS EN SALUD

En el campo de la salud, el aumento exponencial de los datos que los centros de salud producen y gestionan es cada vez más elevado. Por ello, ha surgido la necesidad de desarrollar procedimientos que faciliten estos procesos y que aprovechen todos los datos generados (Kayyali et al., 2013), detectando información desconocida y valiosa (Tomar & Agarwal, 2013). Por tanto, el volumen de datos generados es tal, que su procesamiento y análisis por métodos tradicionales resulta demasiado complejo y abrumador (Koh & Tan, 2011). Para afrontar este desafío, la Minería de Datos (DM) puede jugar un papel clave, ya que permite el descubrimiento de patrones y tendencias en grandes cantidades de datos complejos y la extracción de información oculta para ayudar en la toma de decisiones que pueden mejorar la calidad de los procesos de atención sanitaria (Clavel, Mahulea, Albareda, & Silva, 2020; Maity & Das, 2017; Sen & Khandelwal, 2018; Yoo et al., 2012). DM está estrechamente vinculada con la disciplina científica en el campo de la inteligencia artificial denominada *Machine Learning* (ML), que "emplea una variedad de técnicas estadísticas, probabilísticas y de optimización que permiten a las computadoras aprender de ejemplos pasados y detectar patrones difíciles de discernir de conjuntos de datos grandes, ruidosos o complejos" (Cruz & Wishart, 2006).

Consecuentemente, ML está generando un interés creciente en el campo de la salud, motivado por sus posibles aplicaciones tales como: evaluar la efectividad de los tratamientos, detectar fraudes y abusos en seguros médicos, gestión sanitaria, poner a disposición de los pacientes soluciones médicas a menor coste, detectar síntomas y enfermedades (Alotaibi, Mehmood, Katib, Rana, & Albeshri, 2020), descubrir patrones de tratamiento a partir de registros médicos electrónicos (Huang, Dong, Bath, Ji, & Duan, 2015), detectar grupos de incidentes (Bentham & Hand, 2012) e identificar métodos de tratamiento médico (Koh & Tan, 2011; Tomar & Agarwal, 2013), etc. Asimismo, ML también presenta otros beneficios para el campo de la salud: (1) una potencial reducción del tiempo y esfuerzo necesarios para el diagnóstico y los tratamientos; (2) la capacidad de examinar múltiples áreas simultáneamente; (3) una disminución del potencial error humano; y (4) una mayor accesibilidad a los datos en cualquier momento y lugar (Obenshain, 2004). Además, DM y ML son claves en el camino hacia la medicina personalizada, donde el objetivo es personalizar los tratamientos según las particularidades de cada individuo (Hamet & Tremblay, 2017; Joyner & Paneth, 2015; Weiss, Natarajan, Peissig, McCarty, & Page, 2012; Zhang, Wang, Hu, & Sorrentino, 2014).

Sin embargo, para aprovechar al máximo los beneficios que ofrece ML en el campo de la salud, es necesario remarcar varias consideraciones. Entre otros, se pueden destacar los siguientes aspectos:

- Los datos tienen que estar estructurados y organizados para procesarlos convenientemente y transformarlos en variables adecuadas, lo cual es fundamental en el desarrollo de cualquier software de reconocimiento de patrones y una tarea altamente dependiente del problema (Mannini & Sabatini, 2010; Moons et al., 2012).
- Además, un tratamiento y gestión de manera segura de los datos adquiere especial relevancia en el ámbito de la atención sanitaria, donde debe garantizarse la privacidad del paciente. La gestión de datos sensibles contrasta con otros campos de aplicación de ML (detección de fraudes, predicción de stock, etc.), donde los tratamientos de anonimización suelen ser innecesarios. Por tanto, se debe realizar un tratamiento específico que implique la anonimización y categorización de los datos, con el fin de garantizar la privacidad de los pacientes (Wilkowska & Ziefle, 2012). Por políticas de privacidad, en determinadas ocasiones, si no se realiza una adecuada anonimización de los datos y / o no se obtienen las autorizaciones necesarias para acceder a determinados datos, no se pueden realizar los estudios planificados. Por tanto, la disponibilidad de los datos también debe considerarse un factor clave (Wilkowska & Ziefle, 2012).

- Otro aspecto importante que se debe considerar es la existencia de diferentes costes de aplicar ML en el campo de la salud respecto otros ámbitos de aplicación. La presencia de un falso negativo (por ejemplo, no detectar que un paciente tiene una enfermedad específica) involucra un coste mayor que el de un falso positivo (p. ej., si inicialmente se considera que una persona tiene una enfermedad que realmente no tiene, se realizarán pruebas adicionales para descartarlo, que pueden ser costosas pero generalmente menos dañinas que no diagnosticar una enfermedad existente).
- También debe tenerse en cuenta que ML casi nunca sigue una secuencia lineal que termina en el primer intento; es más, se trata más bien de un proceso de retroalimentación iterativo donde las diferentes etapas interactúan entre sí. Además, en el campo de la salud, donde el flujo de datos es continuo y constante, es razonable asumir que el modelo predictivo puede diseñarse para ser un modelo de "aprendizaje" que debe actualizarse continuamente para mejorar sus predicciones en el tiempo. Por lo tanto, las diferentes etapas necesarias para generar un modelo predictivo fiable deben estar adecuadamente estructuradas, desde la adquisición de datos brutos hasta la toma de decisiones, lo cual es fundamental para lograr la efectividad del modelo (Dolley, 2015).

Todo ello motiva la necesidad de definir las particularidades de la aplicación de las técnicas de ML en el campo de la salud, donde las diferentes etapas del flujo de trabajo de ML deben estar correctamente definidas y estructuradas. La correcta aplicación de las técnicas de ML resulta beneficiosa para clínicos, investigadores, desarrolladores y diseñadores involucrados en el campo de la salud, donde la gestión de la información adquiere un papel trascendental. También se favorecería el diseño de nuevos productos y servicios para mejorar el acceso a la asistencia sanitaria (Andersen, Davidson, & Baumeister, 2013), creando soluciones tecnológicas verdaderamente accesibles (Marin, Blanco, & Marin, 2017) y mejorando la relación entre los sistemas sanitarios y las personas proporcionando servicios adecuados en el momento oportuno (Andersen et al., 2013).

2. OBJETIVOS E HIPÓTESIS

2.1. JUSTIFICACIÓN Y PLANTEAMIENTO DE LA INVESTIGACIÓN

La repercusión de los trastornos del equilibrio tiene un impacto trascendental en la población de los países desarrollados. Esto es debido a la relación directa entre los episodios vertiginosos, sintomatología cardinal de este tipo de trastornos, y las caídas, ya que son un factor de alto riesgo, especialmente, para la población mayor. Todo ello se traduce en un impacto notable en la economía (derivado de los gastos del sector de la salud).

Por ello, se muestra trascendental disponer de la instrumentación, los métodos, y los sistemas adecuados que permitan objetivar el estado de pacientes con trastornos del equilibrio, midiendo el estado de los BSS que lo sustentan, así como la progresión de ambos durante, por ejemplo, la administración de un tratamiento. Sin embargo, estos sistemas requieren ser adaptados a las necesidades y a la situación de uso en clínica de los facultativos sanitarios, ya que son estos profesionales los que van a hacer uso de dichos sistemas. La falta de conexión entre desarrolladores y tecnólogos con dichos profesionales puede derivar en un uso inadecuado de los mismos y, por tanto, en un desaprovechamiento de los recursos. Por ello, se requiere adaptar la información resultante de estos sistemas para hacerla más comprensible y de fácil interpretación para los clínicos.

Igualmente, en relación con la gestión del creciente volumen de datos generados por este tipo de tecnologías, así como por los propios centros de salud, las técnicas de ML permiten aprovechar dicha información creando sistemas que aprenden automáticamente con el continuo flujo de datos. Sin embargo, deben considerarse las particularidades inherentes al campo de la salud para poder hacer uso de estas técnicas; como la gestión de datos personales o la anonimización de pacientes para poder acceder y hacer uso de la información asociada a los mismos.

2.2. OBJETIVOS DE INVESTIGACIÓN

Derivados de los antecedentes expuestos, se exponen los objetivos principales planteados en esta tesis, acompañados de sus correspondientes objetivos específicos:

1. Generar y proporcionar información objetiva sobre el estado de pacientes que presentan trastornos del equilibrio. Para tal propósito, se propusieron los siguientes objetivos secundarios:
   1.1. Verificar que la plataforma estabilométrica diseñada para la investigación (MH-FCE-Balance ©UZ) cumple los estándares establecidos para este tipo de instrumentación en clínica.
   1.2. Diseñar y establecer un conjunto de pruebas de evaluación funcional del equilibrio, aplicando dichas pruebas en una muestra de sujetos asintomáticos, a fin de obtener datos de referencia.
   1.3. Comprobar la viabilidad de aplicar dicho conjunto de pruebas en entornos clínicos.
   1.4. Proporcionar información a los facultativos, de métodos y dispositivos de evaluación del equilibrio, adaptando los resultados y conclusiones para facilitar su interpretación clínica.
   1.5. Desarrollar una prueba que aporte información sobre la capacidad de percepción y ejecución de respuesta, atención e inhibición de los pacientes.
2. Monitorizar y analizar la progresión de este tipo de patologías en pacientes inmersos en tratamientos de rehabilitación, y, para ello:
   2.1. Detectar y proponer aquellas variables más útiles y significativas a efectos del diagnóstico de trastornos del equilibrio.
   2.2. Comprobar si es posible detectar cambios relevantes en el equilibrio entre dos instantes temporales, mediante el uso del MDC, a partir de un estudio de fiabilidad (test-retest) en una muestra de sujetos asintomáticos.
   2.3. Comprobar la aplicabilidad del objetivo anterior en pacientes con trastornos del equilibrio, durante la aplicación de un tratamiento.
   2.4. Desarrollar un método para monitorizar el equilibrio y los sistemas sensoriales que lo sustentan, proporcionando información objetiva e interpretable: midiendo, clasificando y calificando la progresión del equilibrio de pacientes.
   2.5. Aplicar el método propuesto en una muestra de pacientes con trastornos del equilibrio con vértigo como síntoma cardinal.
   2.6. Contrastar los resultados del método propuesto con la evaluación de un médico experto.
3. Generar modelos predictivos para mejorar el diagnóstico en el ámbito de la salud; para tal propósito:
   3.1. Establecer una metodología aplicable en el campo de la salud que permita generar modelos predictivos abarcando todas las etapas del proceso, desde la toma de datos sin procesar, hasta la toma de decisiones médicas.
   3.2. Mostrar y desarrollar las particularidades de aplicar técnicas de *Machine Learning* en el campo de la salud.
   3.3. Demostrar la aplicabilidad de las técnicas de ML como herramienta complementaria de diagnóstico a través de un caso de estudio con pacientes.

Encaje con las líneas de los programas Marco

Los objetivos propuestos en esta investigación están alineados con los objetivos que inspiran los Planes de Investigación de Aragón (Art. 11. Ley de Ciencia de Aragón) del II Plan Autonómico de Investigación, Desarrollo y Transferencia de conocimientos, en particular:

- El avance de la innovación y del desarrollo tecnológico, con la suficiente incidencia en la capacidad productiva de los diferentes sectores de la economía aragonesa.
- La mejora de la calidad de vida, tanto en lo relativo al bienestar social como a la salud y el acceso a la cultura.

El II Plan de Investigación de Aragón identifica cinco líneas estratégicas para el desarrollo regional, alineándose esta investigación con la línea *III*. *Seguridad y calidad de vida individual y colectiva*.

En la línea estratégica III, se establen ciertas líneas prioritarias, de las cuales podemos destacar la siguiente por estar claramente relacionada: *LÍNEAS CLÍNICAS EN CIENCIAS DE LA SALUD. Aquí podemos destacar: Nuevos equipos y materiales de uso clínico, así como herramientas informáticas para asistencia al diagnóstico.*

Por otra parte, la Estrategia Aragonesa de Investigación e Innovación para una Especialización Inteligente, RIS3 Aragón, establece una prioridad estratégica de *bienestar y calidad de vida* y en la cual identifica la siguiente línea estratégica: *Desarrollo de proyectos y servicios integrales, yendo más allá del concepto de envejecimiento saludable, incluyendo servicios para la población por el territorio, satisfaciendo las demandas de servicios de salud y servicios asistenciales de la población dispersa y envejecida.* Tal como se expone en el documento RIS3, la peculiaridad territorial de Aragón con una población envejecida y dispersa incorpora también la necesidad de: *desarrollo de actuaciones enfocadas al sector salud, como elemento de equilibrio territorial.* La tesis, cuyos beneficiarios más inmediatos son pacientes de edad avanzada, se identifica con las prioridades de RIS3

Finalmente, la investigación en salud era uno de los retos sociales en el Pilar 3 en el recién finalizado programa Horizon 2020 y lo seguirá siendo en el nuevo programa de investigación Horizon Europe que acaba de comenzar este año 2021, y se extenderá hasta 2027. En este nuevo programa, los retos en I+D en el ámbito de la salud pasan a formar parte del Cluster salud, dentro del Pilar 2 de retos mundiales y competitividad industrial europea. Al respecto, detecta nichos de desarrollo para lograr estos objetivos como son asegurar un envejecimiento activo, desarrollo de medicamentos y vacunas, tratamiento de las enfermedades crónicas e infecciosas, asegurando la disminución de desigualdades ante la salud.

Desde la Unión Europea se ha marcado como una de las principales líneas de especialización el Envejecimiento Saludable (*Healthy Ageing*), ya que uno de los retos de la sociedad europea viene derivado del aumento de la esperanza de vida y el consiguiente envejecimiento de la población.

La población objetivo de esta investigación se centran en los grupos especiales de población más sensibles: personas que han sufrido traumatismos, pacientes con patologías o enfermedades crónicas músculo esqueléticas o tercera edad. El objetivo de esta investigación es proporcionar al facultativo sanitario medidas objetivas que faciliten el seguimiento de tratamientos de rehabilitación, con el objetivo de mantener y mejorar su calidad de vida y autonomía personal.

2.3. HIPÓTESIS

Los objetivos expuestos pretenden responder a las siguientes hipótesis de investigación:

1. A partir de la plataforma estabilométrica utilizada en la investigación ¿es posible verificar que dicha instrumentación cumple con los estándares establecidos sobre plataformas estabilométricas para su aplicación en clínica, así como implementar y evaluar un conjunto de pruebas objetivas de evaluación del equilibrio?

2. ¿Es posible detectar cambios relevantes entre dos instantes temporales, a partir de la instrumentación y pruebas de evaluación desarrolladas, en pacientes con trastornos del equilibrio?

3. ¿Es viable desarrollar un método para medir, clasificar y calificar la progresión del equilibrio de pacientes, proporcionando información fácilmente interpretable por facultativos sanitarios?

4. ¿Es factible establecer una metodología de obtención, clasificación, comprensión y transformación de información en el campo de la salud, considerando las particularidades inherentes al mismo, que permita generar modelos predictivos complementarios al diagnóstico médico?

2.4. ACCIONES DE INVESTIGACIÓN

Durante el desarrollo de la tesis se han llevado a cabo diversas acciones de investigación y experimentaciones, asociadas a los objetivos expuestos. Se destacan las siguientes acciones:

En relación al Objetivo 1

- Diseño de un conjunto de pruebas para evaluación funcional del equilibrio basadas en la literatura, adaptadas la plataforma estabilométrica MH-FCE-Balance ©UZ y al software propio utilizados. Dichas pruebas han estado presentes en las distintas investigaciones llevadas a cabo relacionadas con el equilibrio.
- Estudio de las características metrológicas de la plataforma estabilométrica para verificación de los estándares establecidos por Scoppa (2013) et al. para el uso de plataformas estabilométricas en clínica.
- Estudio de evaluación del equilibrio con 30 sujetos sanos, con las pruebas previamente diseñadas, a fin de obtener datos de referencia y las variables más significativas según la literatura.
- Desarrollo de una prueba, basada en la literatura, que permita evaluar la capacidad de percepción y ejecución de respuesta, así como la atención e inhibición, involucrando procesos cognitivos.
- Estudio de aplicabilidad con 20 sujetos sanos, 11 jóvenes y 9 personas mayores, de la prueba desarrollada, a fin de obtener datos de referencia que permiten estudiar la viabilidad de uso en pacientes para prevención de caídas.

En relación al Objetivo 2

- Estudio de fiabilidad con 34 sujetos asintomáticos para obtener el MDC de las variables del equilibrio seleccionadas, para poder detectar cambios relevantes en el equilibrio entre dos instantes temporales.
- Desarrollo de un programa estadístico para obtención del MDC a partir de los datos obtenidos de la plataforma estabilométrica.
- Estudio de aplicación del MDC en pacientes (muestra de ocho pacientes) con trastornos del equilibrio para comprobar la aplicabilidad durante un tratamiento en pacientes.
- Desarrollo de un método para monitorización del equilibrio y de los sistemas sensoriales que lo sustentan, considerando las necesidades clínicas en cuanto a inmediatez y facilidad de interpretación de la información, que permita medir, clasificar y calificar la progresión del equilibrio de cada paciente.
- Estudio de aplicación del método desarrollado en pacientes con trastornos del equilibrio (muestra de 42 pacientes con diferentes trastornos del equilibrio con vértigo como síntoma cardinal).
- Estudio comparativo de los resultados obtenidos en el estudio anterior con la evaluación proporcionada por un médico especialista en cuanto a la progresión del equilibrio de pacientes.

En relación al Objetivo 3

- Desarrollo de una metodología para aplicar modelos predictivos en el campo de la salud, incluyendo dispositivos de Smart-Health, contemplando todas las etapas del proceso.
- Estudio de aplicación de la metodología desarrollada en un caso de evaluación de dolor cervical en una muestra de 151 pacientes y un total de 302 pruebas, al realizar dos test por paciente, afectados por latigazo cervical derivados de accidentes de tráfico y otras causas.
- Diseño de una base de datos de los 151 pacientes participantes en el estudio anterior.

Acciones de investigación complementarias

A continuación, se incluyen otras acciones de investigación llevadas a cabo por el doctorando que, si bien sus resultados no se han incluido en los artículos de investigación presentados, han reforzado las sinergias entre los distintos doctorandos involucrados en el grupo de investigación IDERGO, al trabajar en áreas de conocimiento similares.

- Estudio de monitorización de la marcha en pacientes espásticos que derivó en un artículo de investigación (acción realizada durante 2019 y parte de 2020).

  Marin, J.; Marin, J.J.; Blanco, T.; **De la Torre, J.**; Salcedo, I.; Martitegui, E.
  Is My Patient Improving? Individualized Gait Analysis in Rehabilitation. Appl. Sci. 2020, 10, 8558, doi: 10.3390/app10238558.

- Estudio de análisis de la marcha comparativo entre diferentes tecnologías de captura que derivó en un artículo de investigación (acción realizada durante 2019 y parte de 2020).

  Marín, J.; Blanco, T.; **De la Torre, J.**; Marín, J.J. **Gait Analysis in a Box: A System Based on Magnetometer-Free IMUs or Clusters of Optical Markers with Automatic Event Detection**. Sensors 2020, 20, 3338, doi: 10.3390/s20123338.

- Estudio de evaluación funcional en pacientes en el Hospital de la Mutua de Accidentes de Zaragoza MAZ (acción realizada entre abril y diciembre de 2019).

2.5. ARTÍCULOS DE INVESTIGACIÓN

La presente tesis se ha realizado como compendio de publicaciones. En concreto, de cinco trabajos científicos que han abordado investigaciones en distintas áreas de conocimiento conectadas: medición del equilibrio, trastornos del equilibrio, aplicación de modelos predictivos en salud y la generación de información objetiva para los facultativos sanitarios. Sus referencias son las siguientes:

ARTÍCULO 1

De la Torre, J.; Marin, J.; Marin, J. J.; Auria, J. M.; Sanchez-Valverde, M. B. **Balance study in asymptomatic subjects: Determination of significant variables and reference patterns to improve clinical application**. *J. Biomech*. **2017**, *65*, 161–168, doi:10.1016/j.jbiomech.2017.10.013.

ARTÍCULO 2

De la Torre, J.; Marin, J.; Polo, M.; Marín, J.J. **Applying the Minimal Detectable Change of a Static and Dynamic Balance Test Using a Portable Stabilometric Platform to Individually Assess Patients with Balance Disorders**. *Healthcare (S.I. Ageing Effects on Kinematics, Kinetics and Balance)* **2020**, *8*, 402, https://doi.org/10.3390/healthcare8040402.

ARTÍCULO 3

De la Torre J, Marin J, Polo M, Gómez-Trullén EM, Marin JJ. **2021. MCQ-Balance: a method to monitor patients with balance disorders and improve clinical interpretation of posturography.** PeerJ 9:e10916 https://doi.org/10.7717/peerj.10916

ARTÍCULO 4

De la Torre, J.; Marin, J.; Ilarri, S.; Marin, J.J. **Applying Machine Learning for Healthcare: A Case Study on Cervical Pain Assessment with Motion Capture**. *Appl. Sci (S.I. Medical Informatics and Data Analysis)* **2020**, *10*, 5942, doi: 10.3390/app10175942.

ARTÍCULO 5

De la Torre, J.; Bonnet, V.; Mauti, R.; Chabaud, P.; Robert, T. **Implementing a test to assess reaction, attention and inhibition capacity in elderly**. *Computer Methods in Biomechanics and Biomedical Engineering*, 2020, vol. 23 (sup1), S294-S296, doi: 10.1080/10255842.2020.1815313.

Se presenta en la Tabla 1 la relación entre los objetivos expuestos anteriormente, y los artículos donde se han llevado a cabo.

Tabla 1. Relación entre los objetivos y los artículos de investigación de la tesis doctoral.

Objetivos	Artículos				
	1	2	3	4	5
1. Proporcionar información objetiva sobre el estado de pacientes que presentan trastornos del equilibrio.	x	x	x	-	x
1.1. Verificación de estándares para uso en clínica de la plataforma.	x	-	-	-	-
1.2. Diseño pruebas de evaluación funcional del equilibrio.	x	-	-	-	x
1.3. Viabilidad de las pruebas en entornos clínicos.	x	x	-	-	x
1.4. Adaptación de la información al lenguaje médico.	x	-	x	-	-
1.5. Desarrollo prueba para evaluar la percepción, ejecución, atención e inhibición.	-	-	-	-	x
2. Monitorizar la progresión de patologías del equilibrio en pacientes inmersos en tratamiento.	-	x	x	-	-
2.1. Detección de las variables más significativas a efectos de diagnóstico.	x	x	-	-	-
2.2. Test-retest en sujetos asintomáticos.	-	x	-	-	-
2.3. Aplicabilidad en pacientes con trastornos del equilibrio.	-	x	x	-	-
2.4. Desarrollo de un método para monitorizar el equilibrio y los sistemas sensoriales.	-	-	x	-	-
2.5. Aplicación del método en pacientes con trastornos del equilibrio.	-	-	x	-	-
2.6. Contraste de los resultados del método con la evaluación de un facultativo.	-	x	x	-	-
3. Generar modelos predictivos para mejora del diagnóstico en el ámbito de la salud.	-	-	-	x	-
3.1. Desarrollo de metodología de ML aplicable en el campo de la salud.	-	-	-	x	-
3.2. Particularidades de aplicar técnicas de ML en el campo de la salud.	-	-	-	x	-
3.3. Aplicación de metodología en un caso de estudio con pacientes.	-	-	-	x	-

3. RESUMEN DE PUBLICACIONES

ARTÍCULO 1

De la Torre, J.; Marin, J.; Marin, J. J.; Auria, J. M.; Sanchez-Valverde, M. B. Balance study in asymptomatic subjects: Determination of significant variables and reference patterns to improve clinical application. *J. Biomech*. **2017**, *65*, 161–168, doi: 10.1016/j.jbiomech.2017.10.013.

Antecedentes y objetivos

El control postural resulta esencial para llevar a cabo las actividades cotidianas, y los posibles trastornos del mismo tienen un impacto muy significativo en la autonomía personal. La investigación de los mecanismos del equilibrio y del control postural ha despertado el interés de profesionales de diversos campos, sin embargo, dichos profesionales emplean diferentes técnicas de medición y evaluación que a menudo generan resultados desiguales, existiendo divergencia sobre diferentes cuestiones.

Dado el interés de proveer de medios que proporcionen información objetiva sobre la capacidad real del equilibrio de un individuo, así como, la necesidad de crear sistemas de evaluación útiles y sencillos de utilizar por parte de los profesionales de la salud, en este estudio se evaluó el equilibrio de una muestra de personas sanas a través de una nueva instrumentación, aplicando pruebas inspiradas en la literatura. Con ello, se perseguía el objetivo de generar conocimiento acerca del protocolo de uso, proveer de valores preliminares de referencia de personas sanas, y seleccionar la información más útil para evaluar el control postural, mejorando así la aplicabilidad clínica tanto en entornos asistenciales, como médico - legales o periciales.

Ante este objetivo, las etapas llevadas a cabo en el presente estudio fueron: 1) verificar que la instrumentación aplicada cumple los estándares establecidos de la *International Society for Posture and Gait Research* (ISPGR) sobre plataformas estabilométricas en relación a las características metrológicas y antropométricas, las condiciones ambientales y la adquisición de datos; 2) realizar un análisis y una valoración de los resultados de las pruebas de equilibrio a una muestra de sujetos asintomáticos, al objeto de obtener datos de referencia, y comprobar la idoneidad de los procedimientos y protocolos para su aplicación clínica; así como 3) proponer las variables más útiles y significativas a efectos del diagnóstico mediante plataforma estabilométrica.

Materiales y Métodos

El conjunto de pruebas de valoración del equilibrio fue aplicado en una muestra de 30 sujetos voluntarios (12 mujeres, 18 hombres) cuyas edades estaban comprendidas entre 18 y 30 años. En las pruebas se utilizó la plataforma estabilométrica MoveHuman-Dyna diseñada y fabricada por el grupo de investigación IDERGO. Dicha plataforma está constituida por dos placas de aluminio, con 4 cuatro células de carga tipo S de 100 Kg situadas en las esquinas. La información trasmitida por dichas células es transformada de voltios a Kg-fuerza acorde a los parámetros de calibración de cada célula de carga a una frecuencia de 60Hz. El procesado de dichos datos de fuerza, y acorde a la posición relativa de las células, permite calcular en tiempo real la posición de la trayectoria que describe la posición del COP, aplicando la fórmula de cálculo utilizada en diversos estudios.

Se comprobó que la plataforma estabilométrica utilizada cumple los estándares establecidos por el ISPGR para su aplicación clínica. Los parámetros medidos y contrastados fueron: características metrológicas (precisión, exactitud, resolución y linealidad), sistemas de referencia, rangos antropométricos de uso y muestreo.

Cada sujeto realizó un total de seis test de equilibrio: cuatro variantes del test estático de Romberg: 1) superficie rígida ojos abiertos, 2) superficie rígida ojos cerrados, 3) superficie blanda ojos abiertos y 4)

superficie blanda ojos cerrados; una prueba para valorar los límites dinámicos de estabilidad (LOS) y otra de control rítmico direccional (RWS).

Entre las variables inicialmente consideradas, se seleccionaron aquellas variables más significativas de acuerdo con la literatura.

Resultados

Se propusieron las siguientes variables como las más significativas a efectos de diagnóstico del equilibrio: velocidad media del COP, raíz de la media cuadrática (RMS), rango de desplazamiento del COP y área. El análisis de los resultados verifica que las variables evaluadas han arrojados valores similares a otros estudios. Se comprobó la consistencia de valores entre pruebas (ver Figura 1), y se obtuvieron unos valores preliminares de referencia para sujetos asintomáticos.

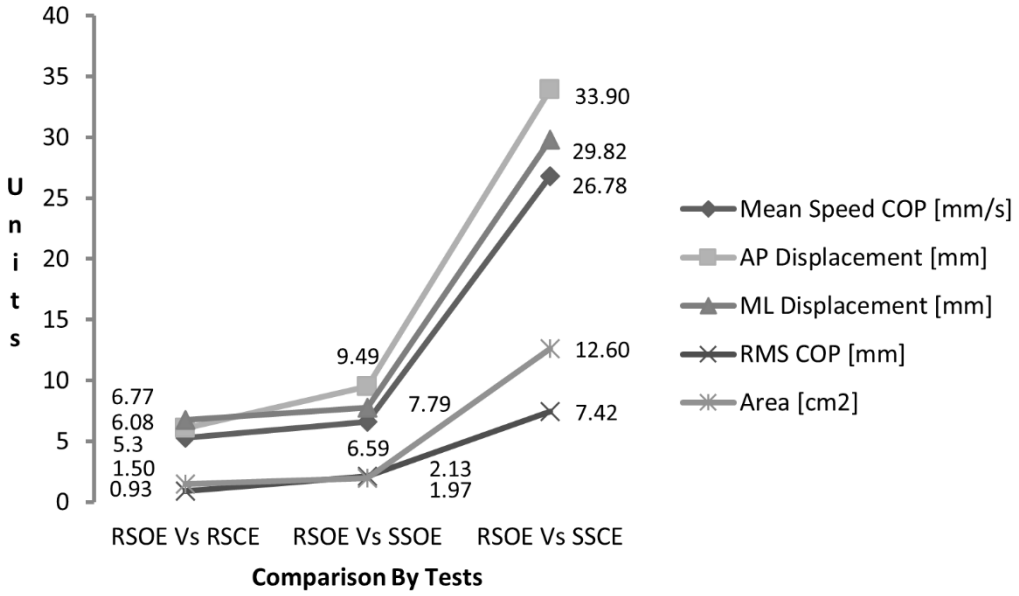


Figura 1. Comparación entre pruebas del test de Romberg.

Significancia y Conclusión

Como resultado de la investigación, se ha verificado que la instrumentación utilizada cumple los estándares establecidos para su aplicación clínica, así como se considera de interés tanto en el ámbito clínico-asistencial como forense-pericial, para valorar el control del equilibrio y el grado de colaboración durante las pruebas. La selección de variables propuesta y los datos de referencia obtenidos de los mismos en la muestra de sujetos asintomáticos participantes, puede favorecer la aplicación de esta tecnología al estudio del equilibrio en el ámbito de la salud.

ARTÍCULO 2

De la Torre, J.; Marin, J.; Polo, M.; Marín, J.J. Applying the Minimal Detectable Change of a Static and Dynamic Balance Test Using a Portable Stabilometric Platform to Individually Assess Patients with Balance Disorders. *Healthcare (S.I. Ageing Effects on Kinematics, Kinetics and Balance)* **2020**, *8*, 402, https://doi.org/10.3390/healthcare8040402

Antecedentes y objetivos

Los trastornos del equilibrio tienen cada vez más prevalencia entre la población mayor en los países desarrollados, y las caídas que resultan de los mismos implican altos costes al sector de la salud. Por ello, es necesario disponer de herramientas e indicadores para evaluar la respuesta a tratamientos de patologías del equilibrio. En relación con este punto, las plataformas estabilométricas pueden proporcionar información objetiva relativa a las patologías del equilibrio; sin embargo, para integrar los tests donde se utilizan estos dispositivos se deben solventar ciertas limitaciones y dificultades. Por otra parte, el mínimo cambio detectable (MDC) representa la variabilidad de las medidas de cada variable; es decir, si se obtiene que un cambio de una variable es mayor que su valor de MDC, sería un cambio relevante, que podría ser causado por el efecto del tratamiento y no por la variabilidad intrínseca de la prueba.

En relación a lo expuesto, el objetivo del presente estudio es detectar cambios relevantes a partir de los valores de MDC en pacientes con trastornos del equilibrio, con vértigo como síntoma cardinal. Para tal propósito, se propusieron las siguientes acciones/etapas de investigación: (1) realizar un test-retest de una prueba de evaluación del equilibrio estático y dinámico en una muestra de sujetos sanos para identificar las variables más sensibles con los valores de MDC más bajos. (2) Analizar si se detectan cambios relevantes entre los resultados de la prueba del equilibrio, antes y después de la aplicación de un tratamiento, en ocho pacientes con trastornos del equilibrio con vértigo de origen vestibular, contrastando los resultados con la evolución observada por un médico especialista.

Materiales y Métodos

Se realizó un test-retest de un conjunto de pruebas para evaluar el equilibrio estático y dinámico en 34 sujetos jóvenes voluntarios y sanos. Posteriormente, para mostrar la aplicabilidad del MDC, ocho pacientes diagnosticados con trastornos del equilibrio por un médico especialista del Departamento de Medicina Física y Rehabilitación del Hospital Alcañiz, realizaron las pruebas del equilibrio antes y después de un tratamiento (con un intervalo de 3 meses entre sesiones), contrastando los resultados con la evaluación realizada por un médico especialista.

El equipo utilizado fue la plataforma estabilométrica portátil MoveHuman-Dyna. La prueba de equilibrio consistió en cuatro tareas del test de Romberg para evaluación del equilibrio estático, mientras que el equilibrio dinámico fue evaluado a través de los límites de estabilidad.

El índice MDC se calculó para cada una de las variables resultantes del estudio test-retest. Para medir los efectos de un tratamiento en un paciente específico, es necesario evaluar el estado del tratamiento en diferentes momentos. Para ello, el enfoque a nivel del paciente propuesto por Hopkings et al. (2017) se aplicó para obtener resultados personalizados para cada paciente.

Resultados

Los resultados obtenidos en el test-retest muestran que la reproducibilidad del sistema es similar o mejor que los resultados encontrados en la literatura. Con respecto a las variables de equilibrio estático con el valor MDC más bajo, destacamos la velocidad promedio del COP en todas las tareas y el RMS, el área y el desplazamiento mediolateral en la tarea de superficie blanda con ojos cerrados. En el test de LOS, destacan todos los límites de COP, la velocidad promedio del COP y el RMS.

Tomando a un paciente como ejemplo, la mayoría de las variables estáticas disminuyeron y la mayoría de las variables dinámicas aumentaron. Según la evaluación del médico especialista, antes y después de los tres meses de la administración del tratamiento, se observó que la sintomatología clínica asociada disminuyó y no se produjeron episodios vertiginosos tras la administración del tratamiento. Se detectó una mejora objetiva en las pruebas funcionales clásicas realizadas por el médico especialista (test up and go, Unterberger test, y test de apoyo unipodal con ambas piernas). Asimismo, el paciente informó mejoría en la anamnesis post-tratamiento. La evaluación realizada por el médico especialista (que emitió evolución positiva del paciente) y los resultados objetivos de las pruebas previas y posteriores (disminución de las variables estáticas, y aumento de las dinámicas; lo cual indica mejoría) coincidieron.

De los ocho pacientes evaluados, existe concordancia entre la evaluación del médico y los resultados de la prueba del equilibrio en seis de ellos; en los dos pacientes restantes, el médico no informó evolución, mientras que la prueba del equilibrio mostró empeoramiento en el equilibrio de los pacientes.

Significancia y Conclusión

En el diseño de la experimentación del cálculo del MDC, se priorizó que este debería apuntar a estudiar y limitar cuantitativamente la variabilidad entre los resultados de las pruebas pre y post, específicamente, de aquellos factores intrínsecos al protocolo aplicado y a la instrumentación utilizada. El objetivo era que estos factores se aislaran del ruido asociado a otros factores como la patología, el tratamiento aplicado, la progresión de la propia enfermedad, etc. Así, se seleccionó una muestra de sujetos sanos y jóvenes, ya que permitió estudiar la variabilidad provocada por factores como la variabilidad intrínseca del ser humano, las condiciones de las pruebas (posición del cuerpo y del pie, duración de la prueba, condición visual y propiocepción, etc.), el factor de aprendizaje, incidentes normales durante el transcurso de las pruebas, o la propia comprensión del participante de cómo realizar la prueba.

Derivado del análisis de los resultados de los pacientes, se extrajeron ciertas pautas "lógicas" de interés para valorar el grado de cambio como positivo o negativo, a la hora de interpretar los resultados de estudios individuales como este. Una reducción de las variables relacionadas con las tareas de equilibrio estático se asoció con una mejora en el equilibrio. Un aumento en las variables de LOS refleja una mejora en el equilibrio dinámico y un mayor control del COP.

Los pacientes mostraron cambios en las variables analizadas que excedían los valores de MDC, siendo comparados con los resultados informados por un médico especialista. Se concluye que, al menos para estos ocho pacientes, ciertas variables fueron lo suficientemente sensibles como para detectar cambios relacionados con la evolución del equilibrio. Con ello se pretende mejorar la toma de decisiones y la monitorización individual de pacientes.

ARTÍCULO 3

De la Torre J, Marin J, Polo M, Gómez-Trullén EM, Marin JJ. **2021. MCQ-Balance: a method to monitor patients with balance disorders and improve clinical interpretation of posturography.** PeerJ 9:e10916 https://doi.org/10.7717/peerj.10916

Antecedentes y objetivos

A nivel mundial, aproximadamente del 20 al 30% de la población ha sufrido un episodio vertiginoso, y de este grupo, el 20% no recibió un diagnóstico claro. El diagnóstico del vértigo requiere medir el grado de alteración, de forma aislada o en combinación, de cada uno de los sistemas sensoriales que sustentan el equilibrio. Un episodio vertiginoso, los mareos o un episodio traumático pueden afectar a estos sistemas en mayor o menor medida y, en consecuencia, al equilibrio del paciente. Por lo tanto, es necesario tener indicadores y métodos para evaluar los BSS, especialmente, durante los tratamientos; por ello, los pacientes con trastornos del equilibrio deben ser monitorizados para detectar cambios a nivel individual, recopilando información objetiva. La posturografía es una herramienta que puede aportar información objetiva del equilibrio de un paciente, sin embargo, existen problemas que deben abordarse, tales como la dificultad de su uso asiduo en clínica, la falta de estandarización de puntuaciones que dificultan la interpretación de resultados, la falta de profundidad de los informes resultantes, etc.

Para abordar las inquietudes planteadas, en relación con la necesidad de proporcionar información objetiva y fácilmente interpretable sobre el equilibrio de los pacientes especificando el origen de la patología, se propuso el método de evaluación MCQ-Balance. Mediante una plataforma estabilométrica, este método detecta cambios relevantes entre dos pruebas de equilibrio consecutivas (monitorización) en pacientes con trastornos del equilibrio, proporcionando información objetiva sobre el origen del desequilibrio. El método MCQ-Balance comprende tres etapas separadas en las que se mide, luego se clasifica y finalmente se califica la progresión del equilibrio de un paciente. En este estudio, la evaluación MCQ-Balance se aplicó a pacientes con trastornos del equilibrio con vértigo como síntoma cardinal. Posteriormente, los resultados obtenidos se compararon con la evaluación de un médico especialista.

Materiales y Métodos

El método se aplicó a 42 pacientes (15 hombres y 27 mujeres) con trastornos del equilibrio de origen periférico o central. Los pacientes fueron incluidos en el estudio por el servicio de rehabilitación del hospital de Alcañiz tras haber sido evaluados previamente.

El equipo utilizado fue la plataforma estabilométrica portátil MoveHuman-Dyna con la que se realizó una prueba para medir el equilibrio estático y dinámico. La prueba de equilibrio consistió en cuatro tareas del test de Romberg para evaluación del equilibrio estático, mientras que el equilibrio dinámico fue evaluado a través de los límites de estabilidad (LOS).

El método de evaluación MCQ-Balance se ejemplifica a través del esquema de aplicación presentado en la Figura 2), el cual consta de tres etapas en las que se mide, clasifica y califica la progresión del equilibrio de un paciente. La entrada del método son las variables proporcionadas por el conjunto de pruebas de equilibrio en dos instantes temporales (variables en la sesión pre y post). El esquema de aplicación muestra las entradas y salidas de cada etapa, así como los procesos (P1-P5) aplicados a ellas. También incluye el tipo de información que se maneja y los cambios interpretativos durante el proceso.

Los resultados del método, para los 42 pacientes y en el transcurso de tres meses donde se les aplicó un tratamiento de rehabilitación de acuerdo a su patología del equilibrio, fueron comparados con la evaluación de un médico especialista. Asimismo, se realizó un estudio estadístico a modo de comparativa

para determinar el coeficiente estadístico Kappa de Cohen para medir la fiabilidad entre evaluadores de medidas categóricas.

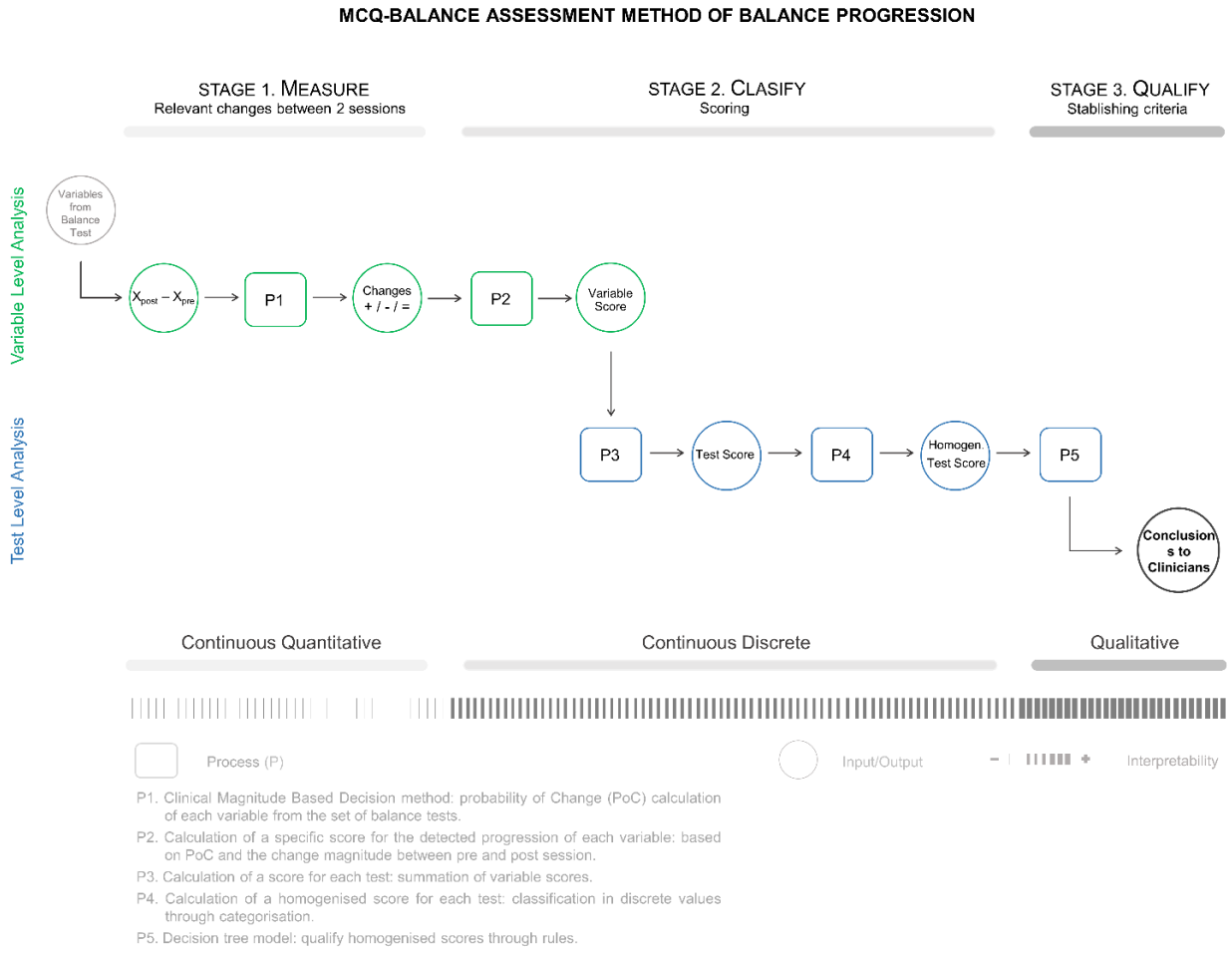


Figura 2. Método de evaluación MCQ-Balance: procesos, entradas y salidas de las diferentes etapas.

Resultados

Se presentaron los resultados para cada paciente y prueba como una escala gradual de progresión positiva, nula o negativa; así como el desglose de variables para su clasificación.

El método MCQ-Balance mostró una precisión del 83,4% y un coeficiente Kappa de Cohen de 0,752 en comparación con la evaluación de un médico especialista.

Significancia y Conclusión

El método MCQ-Balance facilita la monitorización del equilibrio del paciente y proporciona información objetiva que tiene el potencial de mejorar la toma de decisiones médicas y el ajuste de los tratamientos a nivel individual.

ARTÍCULO 4

De la Torre, J.; Marin, J.; Ilarri, S.; Marin, J.J. Applying Machine Learning for Healthcare: A Case Study on Cervical Pain Assessment with Motion Capture. *Appl. Sci (S.I. Medical Informatics and Data Analysis)* **2020**, *10*, 5942, doi: 10.3390/app10175942.

Antecedentes y objetivos

En el campo de la salud, el aumento exponencial de los datos que los centros de salud producen y gestionan es cada vez más significativo. Por ello, ha surgido la necesidad de desarrollar procedimientos que faciliten estos procesos y que aprovechen todos los datos generados, detectando información desconocida y valiosa. El volumen de datos generados es tal, que su procesamiento y análisis por métodos tradicionales resulta demasiado complejo y abrumador. Para afrontar este desafío, el DM y las técnicas de ML pueden jugar un papel clave, ya que permite el descubrimiento de patrones y tendencias a partir de conjuntos de datos grandes, ruidosos o complejos, así como la extracción de información oculta para ayudar en la toma de decisiones en el campo de la salud.

Consecuentemente, ML y DM están generando un interés creciente en el campo de la salud, motivado por sus posibles aplicaciones tales como evaluar la efectividad de los tratamientos, detectar fraudes y abusos en seguros médicos, descubrir patrones de tratamiento a partir de registros médicos electrónicos, etc. Sin embargo, para aprovechar al máximo los beneficios que ofrecen las técnicas de ML en el campo de la salud, son necesarias diversas consideraciones: la información debe estar correctamente estructurada y organizada, llevando a cabo el procesamiento y la transformación apropiados en variables; se debe considerar el tratamiento de datos sensibles de los pacientes, asegurando la privacidad de los mismos aplicando las técnicas necesarias; se debe considerar la existencia de diferentes costes de aplicar técnicas de ML en el campo de la salud respecto otros campos de aplicación en cuanto a los posibles fallos originados (presencia de falsos negativos), etc. Todo ello motiva la necesidad de definir las particularidades de la aplicación de técnicas de ML en el campo de la salud, donde las diferentes etapas en el flujo de trabajo de ML deben estar correctamente definidas y estructuradas.

En función de lo expuesto, los objetivos de este estudio fueron los siguientes: (1) mostrar y desarrollar las particularidades de la aplicación de técnicas de ML en el campo de la salud, detallando todas las fases que comprenden este proceso, desde la adquisición de datos en bruto, hasta toma de decisiones derivada de un modelo predictivo; y (2) demostrar y mostrar su aplicación práctica en un caso real. Específicamente, el proceso de ML se aplicó en un estudio de evaluación del dolor cervical con pacientes afectados por patologías de latigazo cervical derivadas de accidentes de tráfico y otras causas. Este caso de estudio muestra la metodología propuesta en acción para resolver un problema relevante específico. Además, el procedimiento aplicado ha pretendido constituir una guía de aplicación de ML para otros estudios similares en el campo de la salud.

Materiales y Métodos

El procedimiento para aplicar las técnicas de ML en el campo de la salud, considerando sus particularidades inherentes, se ejemplifica a través de una guía práctica. Además, para ilustrar las particularidades de uso de este tipo de técnicas en este campo, se presenta un caso de estudio con un conjunto de datos recolectado mediante la evaluación del movimiento de la columna cervical en 151 pacientes (60 sujetos asintomáticos, 42 con dolor cervical resultante de un accidente de tráfico, y 49 con molestias en el cuello debido a otras causas). El objetivo es tratar de estimar automáticamente la presencia de dolor cervical, lo que puede ayudar a objetivar un diagnóstico y aclarar problemas en caso de proceso judicial. Los pacientes eran sujetos colaboradores, evitando así alteraciones producidas por sujetos no colaboradores inmersos

en un proceso judicial con una compañía de seguros. La prueba de evaluación del movimiento cervical se realizó utilizando el sistema de captura de movimiento MH-Sensor. Los participantes realizaron una secuencia de pruebas de rango funcional cervical (ROM) de los siguientes movimientos: flexión-extensión, rotación y lateralización.

Se ha considerado el proceso completo de ML, con las particularidades de aplicación en un proyecto del campo de la salud. Dicho proceso se basa en siete etapas que van desde la definición del objetivo de diseño hasta el uso clínico del sistema, todo ello ejemplificado a través de un procedimiento de gestión de proyectos (Figura 3). En el estudio se han desarrollado todas las fases, proporcionando ejemplos y referencias para su replicación en otros estudios, así como se presenta su aplicación integra en el caso de estudio anteriormente citado.

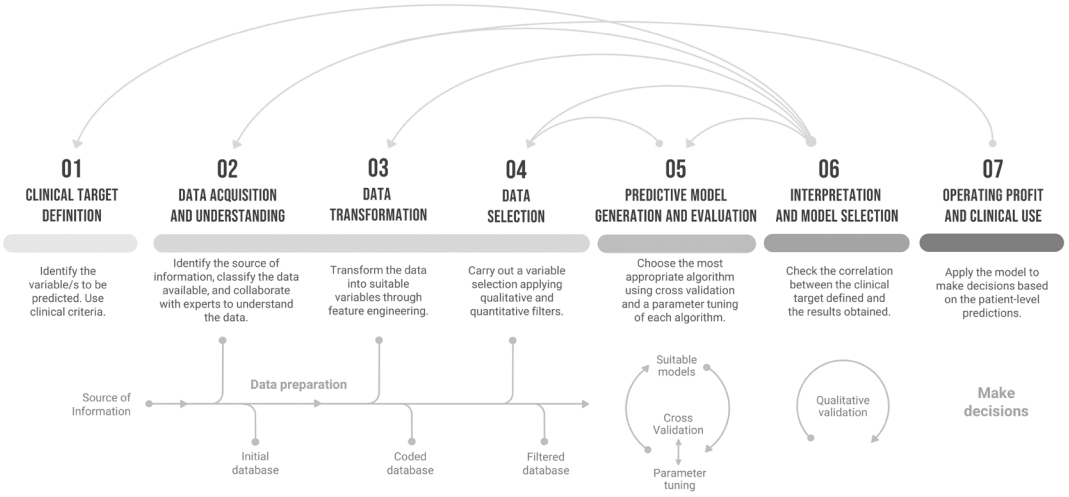


Figura 3. Propuesta de gestión de proyectos para la aplicación de Machine Learning en salud.

Resultados

Se presentan los resultados del caso de estudio para mostrar la aplicabilidad del estudio. Se han generado varios modelos predictivos, destacando el random forest, support vector machine, redes neuronales y gradient boost algorithm como aquellos que mejores resultados arrojaron en función de las métricas inicialmente establecidas (accuracy, precision y recall por encima del 85%) con el objetivo de detectar la presencia de dolor cervical. Asimismo, y de manera previa a la obtención de los resultados definitivos, se realizó un ajuste de parámetros de todos los algoritmos considerados inicialmente.

Significancia y Conclusión

Como resultado del estudio se ha realizado una tabla resumen de “aspectos claves” sobre la aplicación de DM y ML en el campo de la salud, categorizando cada aspecto clave en una clasificación general, seguido de una descripción, la situación real ejemplificada en nuestro estudio de caso, y algunas referencias importantes relacionadas. Asimismo, se han ampliado y destacado los siguientes aspectos por considerarse fundamentales en el estudio que nos ocupó: estructuración de datos, selección de variables, selección del modelo predictivo, variabilidad de las medidas adquiridas, recolección de datos en producción, multidisciplinariedad, explotación de los datos de los sensores y uso de recursos.

El enfoque y los resultados obtenidos pretende ayudar a objetivar los diagnósticos, mejorar la eficacia de tratamientos y ahorrar recursos en los sistemas de salud. El procedimiento, que se ha aplicado a los datos derivado de un estudio de evaluación cervical para verificación y evaluación, también es apropiado para cualquier ámbito sanitario independientemente del origen de los datos.

ARTÍCULO 5

De la Torre, J.; Bonnet, V.; Mauti, R.; Chabaud, P.; Robert, T. Implementing a test to assess reaction, attention and inhibition capacity in elderly. *Computer Methods in Biomechanics and Biomedical Engineering*, 2020, vol. 23 (sup1), S294-S296, doi: 10.1080/10255842.2020.1815313.

Antecedentes y objetivos

Las caídas y las lesiones relacionadas corresponden a problemas importantes de salud pública especialmente, entre las personas mayores. Para evitar una caída, se requiere la percepción de una amenaza a nivel postural, la selección de una respuesta correctora adecuada y la ejecución adecuada de dicha respuesta. La prueba *Choice Stepping Reaction Time* (CSRT) es una prueba integrada que es un buen indicador del riesgo de caída y se relaciona con funciones cognitivas deterioradas.

El objetivo de este estudio fue complementar el CSRT con tareas adicionales para evaluar mejor en las personas mayores la capacidad de percepción y ejecución de respuesta, atención e inhibición. Para ello, se propuso una prueba con diferentes tareas. Se probó en dos muestras diferentes, jóvenes y mayores, con el fin de determinar diferencias entre ellas y anticipar conclusiones relevantes.

Materiales y Métodos

Se realizó una prueba que consta de cuatro tareas diferentes para medir el tiempo de paso, en dos muestras de sujetos sanos. Un grupo joven, formado por once sujetos, y un grupo de nueve personas mayores. El dispositivo utilizado fue una plataforma de fuerza multimodal de análisis postural, la cual consta de seis sensores de presión de bajo coste, fabricados con una lámina de plástico cuya resistencia eléctrica varía con la presión. Hay cuatro áreas en las que se puede pisar como objetivos en las cuatro direcciones cardinales; los pies se colocan en las dos zonas restantes donde se detecta si se ha levantado un pie (Figura 4).

A) **B)** **C)**

Figura 4. A) Plataforma multimodal. B) Interfaz de las pruebas. C) Interfaz del test de Flanker.

Las cuatro tareas realizadas se han diseñado en base a la literatura. En todas las tareas, el sujeto debía dar un paso lo más rápido posible sobre un objetivo designado. Las tareas fueron: Tiempo de reacción simple (SRT: el sujeto era consciente de la dirección objetivo a pisar; esta prueba evalúa principalmente las capacidades de ejecución de percepción y respuesta), el CSRT (tarea similar a la prueba anterior, pero los objetivos son aleatorizados; requiere atención e involucra procesos de decisión), test Go / No Go (enfatiza el control inhibitorio) y tarea de Flanker (introduce estímulos congruentes y no congruentes; evalúa principalmente la atención selectiva y la inhibición perceptual).

Resultados

Se realizó un análisis descriptivo para cada tarea y grupo. Los resultados del análisis descriptivo y estadístico para cada grupo se muestran en la Figura 5. Como era de esperar, el grupo de personas mayores muestra valores de tiempo de paso significativamente más altos que el grupo de jóvenes en todas las tareas, debido a la disminución de las capacidades cognitivas y funcionales con la edad. En relación con este punto, ambos grupos fueron muy bien discriminados por el CSRT.

La comparación entre condiciones congruentes y no congruentes se ha llevado a cabo dando como resultado diferencias significativas en ambos grupos. Parece que la atención selectiva no se ve específicamente afectada en este grupo particular de personas mayores.

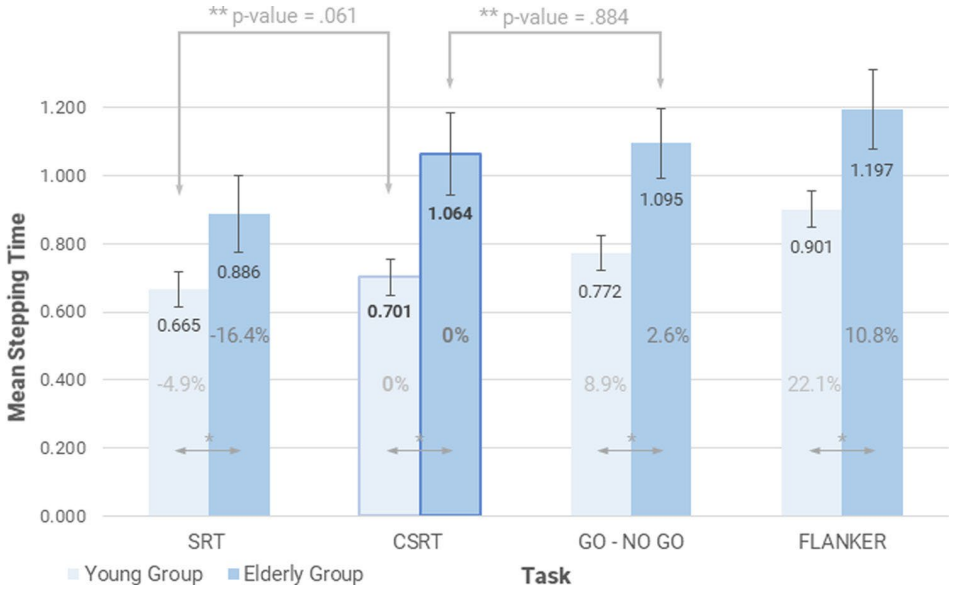


*Figura 5. Análisis descriptivo y estadístico***.**

Significancia y Conclusión

Los resultados confirmaron que el CSRT es muy discriminatorio para ambos grupos y parece particularmente relevante para evaluar el riesgo de caída. Sin embargo, también mostró que la prueba SRT agrega información relevante sobre la capacidad para realizar un paso rápido. Parece un complemento relevante al CSRT para evaluar mejor el riesgo de caída y resalta deficiencias específicas. Las tareas adicionales propuestas, en la forma en que fueron diseñadas, no aportaron mucho al CSRT o al SRT. Sin embargo, se han propuesto vías para mejorar estas tareas y deberían estudiarse en estudios futuros. La perspectiva consistiría en comparar los resultados de esta prueba con el estado clínico de los sujetos, con un seguimiento prospectivo de las posibles caídas. Así, el objetivo final sería obtener una prueba que pudiera complementar los datos clínicos dando como resultado un indicador fiable del riesgo de caída en personas mayores para llevar a cabo acciones preventivas que mejoren la toma de decisiones médicas.

4. ARTÍCULOS

4.1. ARTÍCULO 1

De la Torre, J.; Marin, J.; Marin, J. J.; Auria, J. M.; Sanchez-Valverde, M. B. Balance study in asymptomatic subjects: Determination of significant variables and reference patterns to improve clinical application. *J. Biomech*. **2017**, *65*, 161–168, doi:10.1016/j.jbiomech.2017.10.013.

JOURNAL OF BIOMECHANICS

ISSN:0021-9290

Factor de Impacto JCR (2019):

2.320

Áreas Temáticas:

Biomedical Engineering: Q3 (48/87)

Biophysics: Q3 (38/71)

Journal of Biomechanics 65 (2017) 161–168

Contents lists available at ScienceDirect

Journal of Biomechanics

journal homepage: www.elsevier.com/locate/jbiomech
www.JBiomech.com

Balance study in asymptomatic subjects: Determination of significant variables and reference patterns to improve clinical application

Juan de la Torre [a,*], Javier Marin [a], Jose J. Marin [a,b], Jose M. Auria [a,b], Maria B. Sanchez-Valverde [a,c]

[a] IDERGO – Research and Development in Ergonomics, Biomechanical Laboratory, I3A – University Institute of Research of Engineering of Aragon, University of Zaragoza, Zaragoza, Spain
[b] Department of Design and Manufacturing Engineering, University of Zaragoza, Zaragoza, Spain
[c] Department of Statistical Methods, University of Zaragoza, Zaragoza, Spain

ARTICLE INFO

Article history:
Accepted 15 October 2017

Keywords:
Force platform
Functional assessment
Posturography
Objective diagnosis

ABSTRACT

Postural control is essential when carrying out everyday activities and its possible disorders have a very significant impact on personal autonomy. To provide the means to accurately measure postural control in the clinical environment, this study checks and discusses the suitability of procedures for a new balance assessment system with a stabilometric platform (MoveHuman-Dyna © UZ-IDERGO), which meets the criteria of clinical stabilometric standardisation established by the International Society for Posture and Gait Research (ISPGR) at the Bologna meeting (2009). The study was applied to a sample of 30 healthy volunteers (12 women, 18 men) aged between 18 and 30 years. A total of six balance tests were performed: four variations of the Romberg test, one test for a study of the limits of stability (LoS) and one test for rhythmic weight shift (RWS). Analysis of the results confirms that the variables assessed yielded similar values to other studies, the consistency of values between tests was checked, and preliminary reference values were obtained for asymptomatic subjects. We propose the following variables as the most significant for balance diagnosis: *CoP mean speed, RMS, Range of CoP displacement and area.* As a result of the study, the system is considered of interest in the medical/legal and forensic settings to assess the balance control and degree of collaboration during the tests.

1. Introduction

Postural control is essential when carrying out everyday activities and its possible disorders will have a very significant effect on personal autonomy, being indicative of possible incipient physical or neuronal pathologies (García et al., 2012). These disorders are common dysfunctions encountered by both general practitioners and specialists alike (Baydal-Bertomeu et al., 2004). Included among these conditions are vertigo and dizziness which, together with gait disorders, represent the second leading cause of falls after accidents (Rubenstein and Josephson, 2002). Falls and fall-related injuries correspond to major, global, public health problems. Around 30% of people aged over 65 years and more than 50% of individuals in health centres or care homes will suffer one or more falls per year (Tinetti, 2003). Ageing, sensory changes, musculoskeletal or neurological disorders, cardiovascular diseases, infections or metabolic disorders (Ma et al., 2016), are closely related to balance disorders.

A balance assessment involves measuring centre of pressure (CoP) displacement and quantifies postural control during standing, sitting and walking. Different non-invasive evaluation tests can be used that allow objectifying the degree of control of static and dynamic balance (Duarte and Freitas, 2010). In static balance control methods (Romberg test), subjects must maintain their CoP within the support base throughout the assessment period. Assessments of dynamic postural balance, which is vital for motor control, involve measuring the limits of stability (LoS), corresponding to the maximum voluntary angle or distance in which an individual can regulate their CoP in a given direction without losing balance (Ku et al., 2016). These are tests that can diagnose, measure the evolution of a treatment, or even serve as a means of postural re-education therapy (García et al., 2012). Additionally, the combination of tests was conducted because it allows consistency to be assessed and, consequently, the patient's degree of collaboration when completing the tests, which is relevant in medical/legal and forensic settings (Ramírez et al., 2014).

* Corresponding author at: IDERGO – Research and Development in Ergonomics, Biomechanical Laboratory – I3A (University Institute of Research of Engineering of Aragon), University of Zaragoza, C/ Maria de Luna 3, 50018 Zaragoza, Spain.
E-mail address: 627471@unizar.es (J. de la Torre).

https://doi.org/10.1016/j.jbiomech.2017.10.013
0021-9290/

Research into mechanisms involved in balance and postural control has sparked the interest of professionals from several fields, including physical therapy, sport, engineering, physics, biomechanics, medicine, psychology (Duarte and Freitas, 2010) and entertainment and leisure (Weaver et al., 2016). However, these professionals also use a range of measurement and assessment techniques that often generate different results, with divergent opinions regarding questions such as: the duration of the test, position of the feet, which variables are included or surrounding conditions (Kapteyn et al., 1983; Scoppa et al., 2013; Pinsault and Vuillerme, 2009; Solovykh et al., 2011; Cvecka et al., 2014).

The International Society for Posture and Gait Research (ISPGR) at the Bologna meeting (2009) (Scoppa et al., 2013), has achieved some unification of criteria. The ISPGR addressed the problem of clinical stabilometric standardisation, establishing certain criteria to obtain appropriate accuracy and sensitivity in the measure of the CoP.

There have been subsequent meetings of the ISPGR (Vancouver in 2014 and Seville in 2015) as well as different proposals and recommendations, such as the Japanese standardisation of the evaluation of clinical stabilometry. However, there are still issues without agreement, such as the posture to be adopted by the patient or the most significant clinical parameters.

Given the interest in providing the means to supply objective information about the real balance ability of an individual as well as the need to create useful and simple user assessment systems for health professionals, this study evaluates the balance of a sample of healthy people through a new instrumentation, applying tests based on the literature. The objective is to generate knowledge about the protocol of use, to provide preliminary reference values in healthy people, and to select the most useful information to evaluate the postural control, thus improving the clinical applicability in health care, and forensic environments.

With this objective, the steps carried out in the present study were (1) to verify that the applied instrumentation complies with established standards (Scoppa et al., 2013) (ISPGR) on stabilometric platforms in relation to metrological and anthropometric characteristics, environmental test conditions, and data acquisition, (2) to perform an analysis and assessment of the results of the balance tests on a sample of asymptomatic subjects to obtain preliminary reference data and check the suitability of the procedures and protocols for its clinical application, and (3) to propose the most useful and significant variables for stabilometric diagnosis.

2. Methods

2.1. Ethics and participants

The series of balance assessment tests object of this study were applied to a sample of 30 volunteers (12 women, 18 men) aged over 18 years (mean age = 23.9 years (4.14); mean body weight: 67.52 kg (11.8); mean height: 172.13 cm (6.92)) (Table 1). Of all the volunteer participants, for the purposes of this study, the group whose ages ranged from 18 to 30 years was selected.

Table 1
Participant anthropometric characteristics, mean (SD).

Characteristics	Men (n = 18)	Women (n = 12)
Age (yr)	23.16 (4.02)	25.1 (4.23)
Weight (kg)	73.22 (9.17)	59.4 (10.52)
Height (cm)	174.84 (5.78)	167.83 (6.56)
Body fat index (%)	13.62 (5)	22.69 (7.33)
Foot length (cm)[a]	26.14 (1.05)	23.73 (0.93)

[a] Foot Length measurements were taken between the proximal and distal points on the foot outline (Pawar and Dadhich, 2012).

The following inclusion criteria were established for study participation: (i) aged between 18 and 30 years, (ii) no prior history of neurological, visual, vestibular or balance alterations, (iii) no record of musculoskeletal or neurological disorders in the last 12 months, (iv) no prior history of surgery on the lower limbs or which could affect balance, (v) ability to walk normally, (the subject's lifestyle does not represent an obstacle). The study was previously approved by the Government of Aragon's Human Research Ethics Committee. Before starting the tests, the volunteers signed a form that meant they consented to undergo the tests and understood the aim of the study.

Subjects completed a practice run of each test so that the tester could verify they understood how it worked, assumed the correct posture and executed the tests correctly. This also gave the subjects the chance to get used to the platform and environment, which are considered relevant factors in some balance studies (Taylor et al., 2015).

After finishing the tests, subjects were asked to complete an evaluation questionnaire regarding their performance: muscle aches (and their location), possible feelings of instability, dizziness or vertigo, degree of difficulty of the test, and a functional evaluation of the tests and the resources used, both software and hardware. The questionnaire was designed to detect any deficiencies while carrying out the tests and any points for improvement.

2.2. Description of the instrumentation

The tests made use of the force platform (*Move*Human-Dyna ©UZ) designed and manufactured by the IDERGO research group (the Aragon Institute for Engineering Research, *I3A*, at the University of Zaragoza). The platform consists of two aluminium plates (Al-6062). The upper plate is square-shaped, measuring 415×415 mm; the lower plate is a circular ring with an outer diameter of 555 mm; both plates are 15 mm thick. There are four 100 kg, S-type load cells located between the two plates; they are positioned in the corners, separated by 332 mm and connected to a *PhidgetBridge 4-Input* (Phidgets, 2017). The result is a monoaxial force platform (Postolache and Postolache, 2017) that can measure vertical forces (Fig. 1b and c).

The *PhidgetBridge* interface board was connected directly to a PC featuring the software that controls the device and gathers data from the load cells. The information transmitted by these cells is converted from volts to kg-force in accordance with each load cell's calibration parameters at a frequency of 60 Hz. Processing the force data in function of the cells' position means we can calculate the real-time position of the trajectory that describes the position of the CoP by applying the appropriate formula (Ma et al., 2016; López and Calidonio, 2009). A foam rubber "balance pad", with characteristics in line with those employed in the literature (Baydal-Bertomeu et al., 2004), was used in the tests that had to be conducted with a soft surface on top of the force platform (Fig. 1e).

2.3. Verification of instrumentation according to standards

It proved that the stabilometric platform (SP) meets the standards established by the ISPGR for clinical application (Scoppa et al., 2013). Regarding the metrological characteristics, the parameters were calculated according to the calibration data of each load cell supplied by the manufacturer (Phidgets, 2017), using the error propagation law. Based on the CoP calculation (López and Calidonio, 2009), the following values have been obtained:

- Accuracy: 0.0619 mm (better than 0.1 mm) (Scoppa et al., 2013). An error due to non-linearity and hysteresis in the most unfavourable situation has been considered in each cell.

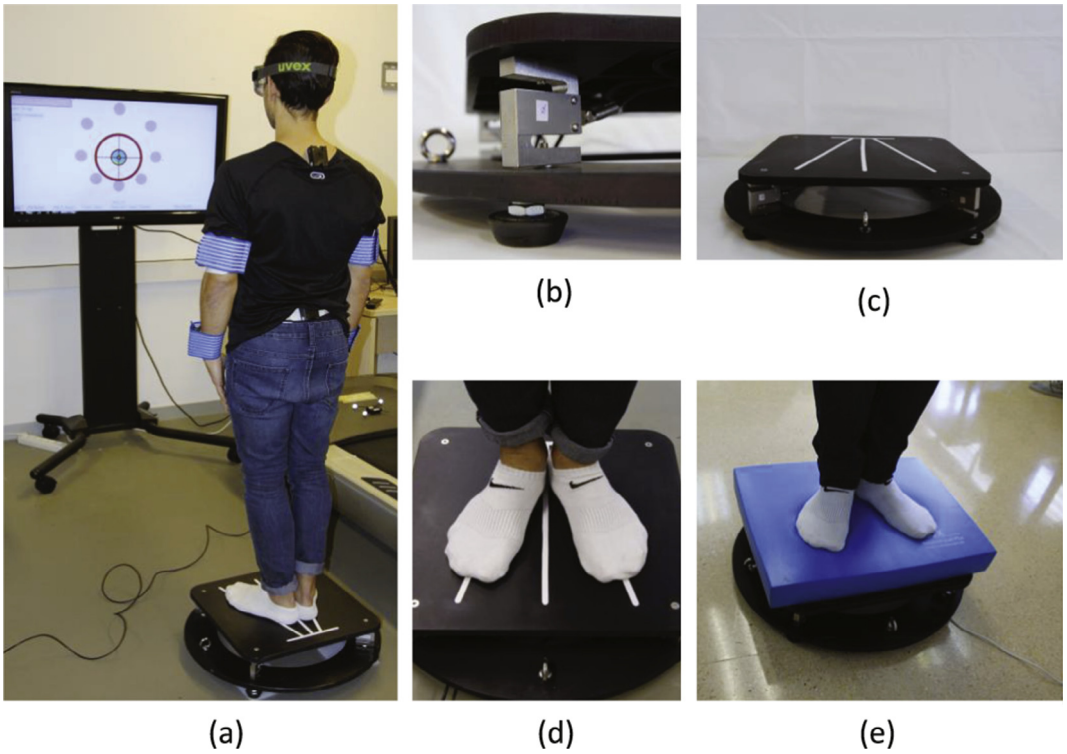


Fig. 1. Force platform and body position established for the tests. (a) General view of the body position, force platform and working environment. (b) Detailed view of one of the four load cells that make up the platform. (c) General view of the force platform. (d) Detailed view of the position of the feet at 30° and with heels together. (e) Detailed view of the foam rubber balance pad.

- Precision: 0.02878 mm (better than 0.05 mm) (Scoppa et al., 2013).
- Resolution: 0.005 mm (better than 0.05 mm) (Scoppa et al., 2013). Calculated from the registration data of the cells under conditions of maximum load.
- Linearity: 99.98% (more than 90% recommended) (Scoppa et al., 2013).

The origin of the (*X*, *Y*) coordinates of the CoP is normally in the left back corner of the platform with reference to the position of the subject during the test (Scoppa et al., 2013). In our SP, another reference system is applied (as will be justified later), specifically, the (*X*, *Y*) position of the CoP at the beginning of the test.

In relation to the anthropometric ranges, permissible weight is up to 300 kg working at 75% of the load cells (recommended: 20–200 kg) (Scoppa et al., 2013), subject height without limit (80–250 cm) (Scoppa et al., 2013), and foot length up to 40 cm (recommended not less than 35 cm) (Scoppa et al., 2013). The sampling rate was 60 Hz, which is higher than the minimum recommended rate of 50 Hz (Scoppa et al., 2013; Yamamoto et al., 2017).

2.4. Environmental test condition and body position

The environment conditions have been defined according to some papers (Kapteyn et al., 1983; Scoppa et al., 2013). The subject should remain in the same standing position during all the tests, erect body, arms held closely to the body (Yamamoto et al., 2017), neck in a neutral position, without any shoes (Terrier and Reynard, 2014) and with feet abducted at 30° and heels together (Fig. 1d). There are discrepancies between authors, but this feet position is the most employed and because of this, authors recommend it (Baydal-Bertomeu et al., 2004; Kapteyn et al., 1983; Pinsault and Vuillerme, 2009; Scoppa et al., 2013; Yamamoto et al., 2017).

2.5. Experimental protocol

Each study subject completed a total of six balance tests: four variations of the static Romberg test, one to assess the dynamic LoS and another for the rhythmic weight shift test (RWS).

2.5.1. Romberg test

The Romberg test (Khasnis and Gokula, 2003), during which subjects had to remain as still as possible, was devised specifically for assessing static balance control and studying the influence of the three sensory systems involved in balance: vestibular, proprioceptive and visual (Patrícia Paludette et al., 2015; Toledo and Olivera, 2012; Vanmeerhaeghe et al., 2009; Wolf et al., 2008). This was done by completing four consecutive tests to evaluate CoP trajectory (Fig. 2a); subjects had to stay still in an upright, neutral position (Fig. 1a) for as long as possible and according to the following situations:

- On the platform's rigid surface: (i) "rigid surface, eyes open" (RSEO) and (ii) "rigid surface, eyes closed" (RSEC). In (ii), subjects were deprived of visual input.
- On an unstable surface, created by placing a foam rubber pad on the platform, which altered the proprioceptive input: (iii) "soft surface, eyes open" (SSEO) and (iv) "soft surface, eyes closed" (SSEC) (Baydal-Bertomeu et al., 2004; Peydro de Moya et al., 2005; Gil-Agudo et al., 2006). In (iv), subjects were also deprived of visual input and so they only received vestibular input.

We established a 40-s period in which subjects must remain motionless for each of the cited situations; while there is some controversy in the literature regarding this period, we chose an intermediate period from among those found in different articles (Kapteyn et al., 1983; Scoppa et al., 2013; Caña-Pino et al., 2015;

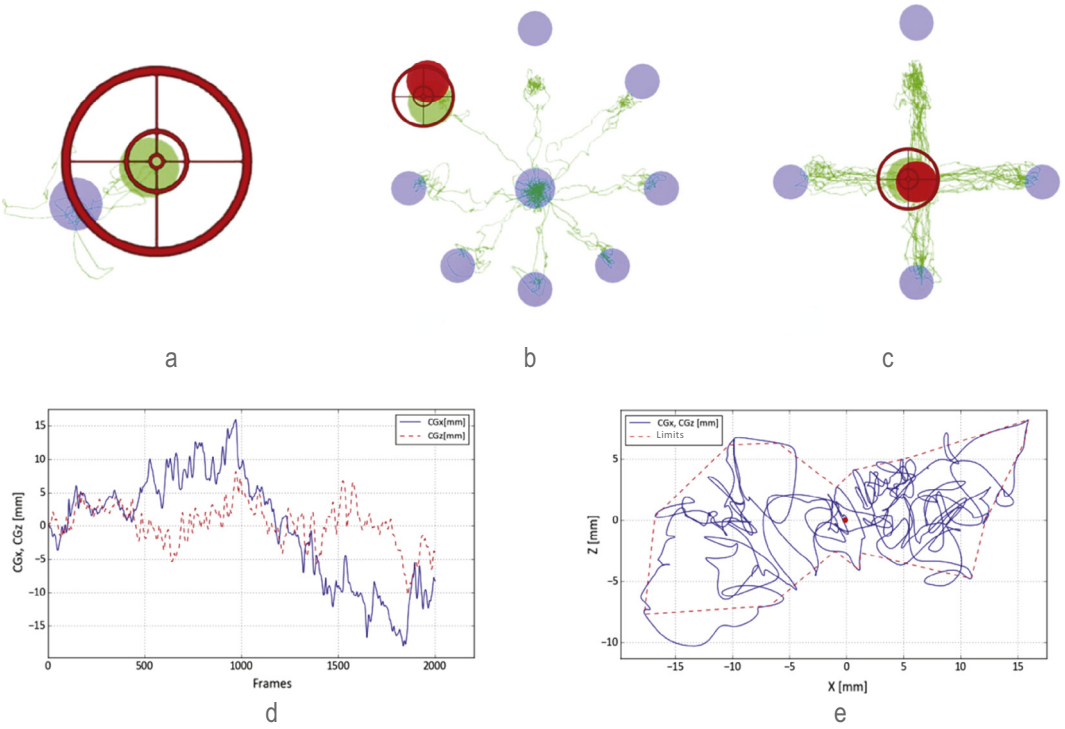


Fig. 2. (a) Centre of pressure (CoP) trajectory in Romberg test. (b) Limits of stability directions defined as 1: forward; 2: forward–rightward; 3: rightward; 4: backward–rightward; 5: backward; 6: backward–leftward; 7: leftward; 8: forward–leftward. (c) The rhythmic weight shift test. (d) Stabilogram. (e) Statokinesiogram.

Peydro de Moya et al., 2005; de Sá Ferreira and Baracat, 2014; Taylor et al., 2015; Parreira et al., 2013; Yamamoto et al., 2017). At the start of the tests, the tester instructed the subjects how to stimulate abdominal toning since it has an influence on stability and balance (Ayllón and Fernández, 2006). The Romberg tests were designed and executed so that there was no biofeedback; during the tests performed with the eyes open, subjects were asked to focus on a fixed point thereby promoting postural control (Taylor et al., 2015).

2.5.2. Limits of stability (LoS)

The LoS assessment test, as well as the RSW test, described below, was based on protocols found in the literature (Peydro de Moya et al., 2005). The purpose of these tests was to complete a dynamic balance assessment and, in contrast to the Romberg test, they were designed around a biofeedback approach. These tests were also carried out in the position indicated in Fig. 1a.

LoS aims to study each subject's maximum reach when moving their CoP along the radii of an octagon. The octagon's first radius corresponds to north (in front of the subject when they are looking towards the monitor) and the others occur every 45° in a clockwise direction (Fig. 2b).

In Fig. 2b the CoP is represented by a bullseye that moves around according to the body's movements. The red ball, called the target ball, indicates the position the subject must try to reach with the bullseye (CoP) by controlling their body movements.

This red target ball moved from the central position to the external positions (1–8, indicated with blue reference marks) whose locations were predefined according to typical reach distances and then adjusted to each subject's height. The red target ball moved at a moderate, constant speed and travelled consecutively between positions (from 1 to 2, 2 to 3, etc.), having previously passed through the central position. Every time the red target ball reached the end of its trajectory along one axis, it paused for 8 seconds while the subjects had to maintain the position they had reached for their CoP.

2.5.3. Rhythmic weight shift (RWS)

The RWS test was designed to analyse each subject's control in terms of managing CoP displacement, but rhythmically and while following linear trajectories. In this case a red target ball moved around while its speed increased progressively, first in an anteroposterior direction and then in a mediolateral direction (Fig. 2c).

During the test, subjects also had to try to follow the red target ball with the bullseye, ensuring they were superimposed at all times, but this time with the added difficulty of a ball travelling at an ever-increasing speed each time it passed along an axis.

2.6. Statistical analysis

We used the statistical software *Minitab Version 17* (*Minitab Inc NYST, Pennsylvania*) for the statistical analysis of the data. We applied methods used in the literature within the same context as the current area of interest (González et al., 2014).

To analyse the results statistically, first, it is necessary to consider which variables are most significant. In this way, in the Romberg tests, the group of variables given in Table 2 was initially calculated in accordance with the literature. Nevertheless, after the analysis performed in the present study and the recommendations proposed by the ISPGR (Scoppa et al., 2013), we selected the following variables as the most significant (marked with an asterisk in the table): Range of CoP displacement in anteroposterior (AP) and mediolateral (ML) directions (Baydal-Bertomeu et al., 2004; Pinsault and Vuillerme, 2009; Solovykh et al., 2011; Maze et al., 2016; Corazza et al., 2016; Bruniera et al., 2015; Sample et al., 2016; Patrícia Paludette et al., 2015; Borges et al., 2016; Duarte and Freitas, 2010, Cuozzo Lemos et al., 2016; García et al., 2012; Dorneles et al., 2014), area (Pinsault and Vuillerme, 2009; López

Table 2
Variables initially considered in the balance tests.

Variable	Applicable in tests	Description
Range of CoP displacement (mm)[a]	Romberg/ LoS/RWS	The distance between the max and min CoP displacement for each direction. The greater the values, the worse the postural stability. Represents a global measure that allows to estimate overall postural performance. It provides the maximum displacements in the Z-axis (anteroposterior) and in the X-axis (mediolateral)
Area (cm^2)[a]	Romberg/ LoS/RWS	The surface area covered by the trajectory of the CoP with 90% confidence interval. Area is a measure of the COP spatial variability (García et al., 2012)
CoP mean speed (mm/s)[a]	Romberg/ LoS/RWS	The total distance covered by CoP divided by the duration of the sampled period and constitute a good index of the amount of activity required to maintain stability (García et al., 2012)
RMS (mm)[a]	Romberg/ LoS/RWS	It is the root of the quotient of the sum of the squares of the distances of the samples to the mean value of these samples and the number of samples (Dorneles et al., 2014)
Total CoP displacement (mm)	Romberg/ LoS/RWS	The total distance travelled by the CoP. Calculated by multiplying the mean speed by the capture time
Mean CoP_X (mm)/Mean CoP_z (mm)	Romberg/ LoS/RWS	Mean value for the point cloud of CoP coordinates along the X and Z axes respectively. Represents the mean position of the CoP during the test
Coefficient distance-area (mm/cm^2)	Romberg/ LoS/RWS	Ratio between the total distance travelled and the area. A measurement of the energy required to maintain balance. A larger value represents less efficiency and implies worse energy management by the patient
CoG height (mm)	Romberg/ LoS/RWS	Height of the patient's CoG. Calculated by multiplying subject height by coefficient 0.55
CoG angle axis Z (°)/CoG angle axis X (°)	Romberg/ LoS/RWS	Taking into account the height of the patient's CoG, this corresponds to the angle, in both the anteroposterior and mediolateral planes, whose arc is the maximum CoG displacement in this plane
CoP foot axis Z (%)	Romberg/ LoS/RWS	Percentage of the CoG movement with respect to foot size
CoP limits (mm)[a]	LoS/RWS	Maximum displacement reached along each axis of the octagon radii
"Success" (%)[a]	LoS/RWS	Percentage of points registered in which the CoP coordinates were situated within the red target ball (N_{OK}) in relation to the total number of points for each axis (N), i.e., all data points considered to represent a success. "*Success*" (%) = $N_{OK}/N \cdot 100$
Directional rhythm ability (DRA) (%)[a]	RWS	This quantifies the average of the distances between the red target ball and the bullseye (CoP) at each instant recorded (50 Hz). It is a dimensionless value, wherein a lower value denotes greater control along this axis
Control and effectiveness of movement (CEM) (mm)[a]	RWS	This quantifies the subject's involuntary movement in a direction perpendicular to the red target ball's axis of movement. A lower value denotes better control along this axis

[*] CG: Centre of Gravity. RWS: Rhythmic Weight Shift. CoP: Centre of Pressure.
[a] Variables selected in the present study.

and Calidonio, 2009; Solovykh et al., 2011; Bruniera et al., 2015; Duarte and Freitas, 2010; García et al., 2012; Dorneles et al., 2014), CoP mean speed (Pinsault and Vuillerme, 2009; Cvecka et al., 2014; Solovykh et al., 2011; Bruniera et al., 2015; Sample et al., 2016; Borges et al., 2016; Duarte and Freitas, 2010; García et al., 2012; Dorneles et al., 2014; Taylor et al., 2015) and root mean square (RMS) (Kapteyn et al., 1983; López and Calidonio, 2009; Duarte and Freitas, 2010; Caña-Pino et al., 2015; Weaver et al., 2016).

The variables selected in both the LoS and RWS tests were based on those proposed in the literature (Baydal-Bertomeu et al., 2004), as shown in Table 2.

A descriptive analysis of the results was performed in different tests. In the Romberg test, the dependence of the results on the factors of gender, height, and foot length has been verified. To perform the test by gender in each of the five variables in the four situations of the Romberg test, a comparison of means with independent samples was conducted. After the division of the sample by gender, the size of the subgroups was less than 25, so a normality test has been performed in each group. In cases in which both groups are normal, a contrast is performed using the student's *t*-test (with equal or different variances depending on the variance test results). When absence of normality was detected, a Mann-Whitney test was performed. To check whether the results are related to height or foot length, the significance of the correlation coefficient was obtained. The possible differences between the RSOE test and the other three situations of the Romberg test were analysed using a repeated-measure analysis of variance (ANOVA).

In relation to the LoS test, the symmetry of the axes has been analysed in terms of reach and success obtained with the ANOVA with repeated measurements, Tukey paired comparison test, and paired data contrasts using student's *t*-test. The alpha value used is the usual 0.05.

3. Results

3.1. Romberg test

The results of the Romberg test are presented in Table 3. We observed a growing trend in the values recorded for each of the variables as difficulty of the tests grew.

To verify whether there was a relationship between the selected variables and gender, the corresponding comparisons were made using student's *t*-test. The absence of normality of one of the groups (variables were area and RMS in the RSCE test) required the use of Mann-Whitney's non-parametric contrast in only two cases. No difference was detected by gender since a *p*-value of less than .05 was obtained in only two of the variables (in the SSOE test: RMS, with *p*-value = .014 and anteroposterior displacement, with *p*-value = .025).

We used the Pearson correlation coefficient to determine whether there was any correlation between height/foot length and the selected variables. As expected, in all the variations of the Romberg test, we did not observe any correlation between variables (*p*-value > .05), with the exception of certain variables in the SSEO test:

- Comparing height: RMS (*p*-value = .001) and area (*p*-value = .004).
- Comparing foot length: RMS (*p*-value = .000) and anteroposterior displacement (*p*-value = .003).

Once the independence of the results was verified with the factors of gender, height, and foot length, we evaluated the mean of the differences for the other three Romberg tests with respect to the RSEO test for all five variables in the study. Fig. 3 shows the relations between tests, highlighting a growing trend of the

Table 3
Results of the main variables selected in the four variations of the Romberg test, mean (SD).

Variable	Rigid surface, eyes open	Rigid surface, eyes close	Soft surface, eyes open	Soft surface, eyes close
CoP mean speed (mm/s)	7.788 (2.111)	12.413 (3.648)	14.379 (3.139)	33.281 (7.639)
RMS (mm)	4.842 (1.230)	5.771 (1.581)	6.975 (1.604)	12.266 (2.222)
Area (cm^2)	1.435 (0.629)	2.938 (1.624)	3.261 (1.163)	14.034 (5.258)
Anteroposterior displacement (mm)	17.653 (4.349)	23.729 (6.965)	27.146 (6.149)	51.552 (10.348)
Mediolateral displacement (mm)	14.003 (4.102)	20.777 (5.401)	20.996 (5.023)	43.825 (8.268)

resulting values that followed the increasing difficulty of the tests. The plots for all the variables display comparably increasing gradients.

3.2. Limits of stability (LoS)

The data obtained for the CoP limits and the "success" in the different directions are presented in Table 4. We can see that the maximum values reached for the LoS were in the anterior zone, specifically in the left-anterior axis. The minimum values recorded correspond to the posterior zone, the lowest value was observed in the right-posterior axis.

The analysis of variance (repeated measures ANOVA) revealed significant differences in a collective comparison of the limits reached along the eight axes (p-value = .000). The Tukey pairwise comparison pooled the data according to the values obtained for each axis. The axes were grouped as follows: 1-2-8/7-3/3-6/6-5-4. Symmetry analysis, studied by other authors (Korhonen et al., 2010), was carried out with respect to the anteroposterior axis and did not reveal any significant differences, apart from between axes 4 and 6 (p-value = .023).

Results for the LoS success obtained after analysing the variance (repeated measures ANOVA) are noteworthy as they do not demonstrate any significant differences following a simultaneous comparison along the eight axes (p-value = .110). The Tukey pairwise comparison supports these results considering that all the values belong to the same "pool". Thus, the average total success is 79.5% (12.3).

3.3. Rhythmic weight shift (RWS)

Table 5 presents some values for the directional rhythmic ability (DRA) parameter that were notably higher in the posterior direction than in the anterior one; there was also symmetry between the mediolateral directions. This lack of uniformity in the anteroposterior direction compared to the symmetry in the mediolateral direction translated into a greater reach capacity in the mediolateral movement. Regarding the RWS test, anteroposterior sway was notably greater than mediolateral sway, which is consistent with other, similar studies (Baydal-Bertomeu et al., 2004; Kapteyn and Wit, 1972).

In contrast, higher values for the control and effectiveness of movement (CEM) parameter contained in the mediolateral direction compared to the anteroposterior direction indicated a higher level of control over mediolateral movements.

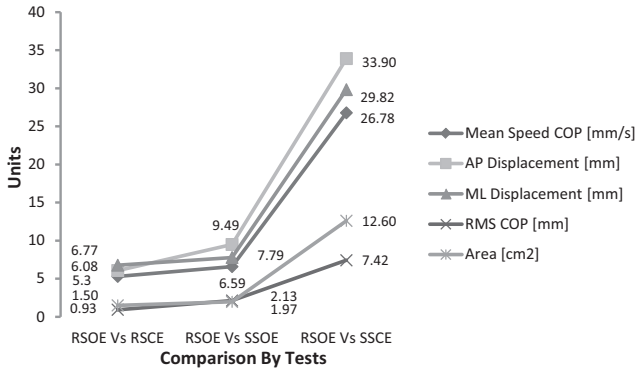


Fig. 3. Comparison of the differences in the Romberg tests.

4. Discussion

In this study, we have analysed a sample of healthy people, with a new balance platform using tests based on the literature. Following the results, we discuss the instrumentation, selected variables, relationship between the results with the characteristics of the sample (height, age, and foot length), consistency of the results, and study limitations.

The platform used in the study fulfils the criteria established by the ISPGR; however, in relation to the position of the coordinate origin, another criterion has been applied. The ISPGR states that the origin of the (X, Y) coordinates of the CoP is in the left back corner of the platform (Scoppa et al., 2013). However, if that position were chosen, the resulting coordinates would depend on the dimensions of the platform and the initial position of the subject at the beginning of the test, which makes it difficult to compare different tests, especially in the stabilogram and statokinesiogram graphs. The comparison between tests is especially necessary with the Romberg RSOE test of the same subject performed in the same session and is useful among different subjects.

To facilitate such a comparison, it was proposed that the reference system of the (X, Y) coordinates of the CoP corresponds to its position at the beginning of the test. That is, the (X, Y) coordinates of the CoP are initialised to zero at the beginning of the capture and immediately after the 'adaptation phase', which should not be less than 5 s (Taylor et al., 2015), and once the operator has checked that the subject is in a stable position. Adopting this approach provides the following advantages:

- The (X, Y) coordinates of the CoP are normalised with the subject itself (from the initial position of stability), regardless of the dimensions of the platform and its initial position on it. In this way, the graphs will show the patient's evolution during the test from the initial position and thus facilitate the clinical interpretation.
- The stabilogram (Fig. 2d) shows positive and negative values. This is especially useful for visualising possible asymmetries.
- In the same way, the statokinesiogram (Fig. 2e) shows an origin point with positive and negative values, which makes it easy to distinguish movements of the CoP in the four quadrants of the area described by the same from the starting position of the test.
- It does not affect classical parameters (Scoppa et al., 2013) since these do not depend on the established origin.

However, in the initial instant where this coordinate adjustment is performed at zero, it is recommended, as in our experimentation, to save the displacement of the CoP with respect to

Table 4
Centre of pressure (CoP) Limits and "success" in limits of stability, mean (SD).

Variable	Forward	Forward–rightward	Rightward	Backward–rightward	Backward	Backward–leftward	Leftward	Forward–leftward
CoP limits (mm)	85.23 (13.03)	89.58 (12.27)	71.23 (7.49)	60.52 (10.69)	61.64 (11.97)	64.11 (10.84)	74.09 (8.70)	91.61 (13.65)
Success (%)	77.05 (12.11)	77.98 (13.53)	77.32 (12.50)	78.65 (10.52)	83.30 (12.51)	80.39 (12.52)	81.95 (11.72)	77.89 (13.25)

Table 5
DRA and CEM in the rhythmic weight shift test, mean (SD).

Variable	Forward	Backward	Rightward	Leftward
DRA (%)	17.39 (3.39)	26.18 (6.97)	23.35 (4.85)	22.61 (5.61)
CEM (mm)	5.23 (1.26)	5.33 (0.97)	8.07 (2.55)	6.90 (1.70)

DRA: Directional Rhythm Ability. CEM: Control and Effectiveness of Movement.

the central point of the platform. This allows undoing the reference change and using the criterion established by the ISPGR.

The proposed change of CoP coordinate origin is considered to facilitate the clinical interpretation of the stabilometric graphs. With this same objective, identifying the variables that produce the most significant diagnostic data is considered vital in order to help practitioners interpret and evaluate balance test results. A review of the literature revealed a wide range of variables (Table 2) that correspond with those implemented in the series of balance tests used in the present study. An analysis of the results of these tests, and in accordance with relevant references (Baydal-Bertomeu et al., 2004), allowed us to identify the most representative variables (Table 2), which simplifies and facilitates the assessment process.

Regarding the variables suggested for the diagnosis, note that the CoP mean speed and total CoP displacement are of course mathematically related. Although the displacement variable is easier to interpret and used on a frequent basis, we preferred to select the CoP mean speed to help compare the results of the Romberg test (40 s) against those of authors who set a different immobility period, since the CoP mean speed does not depend on the duration of this period.

As expected with the sample of subjects selected between the ages of 18 and 30, including both genders and similar anthropometric measures, no correlation was found between foot length and height with the selected variables. At this point, the result of our work differs from other authors with a sample of greater variability in height. These authors found slight correlation as a function of height (Alonso et al., 2012).

To evaluate the consistency of the obtained results, we used patterns consulted in the literature related to the forensic field. As a diagnostic test, the proposed series of balance assessment tests can provide information about the degree of patient collaboration by detecting possible exaggerations (Ramírez et al., 2014) of the pathology (inconsistent physiological patterns), whether due to fear, anxiety or other reasons; a relevant point of interest in expert and forensic settings. Regarding this point, other studies (Peydro de Moya et al., 2005; Goebel et al., 1997) have identified non-physiological patterns that potentially indicate a lack of collaboration by observing CoP displacement in different tests. Specifically, according to the literature, it would be inconsistent if better results were observed in the more difficult tests (Baydal-Bertomeu et al., 2004; Peydro de Moya et al., 2005), i.e., better results in the SSEC Romberg test than in the RSEO Romberg test.

Therefore, these patterns and the coherence analysis have served to measure the consistency of the tests performed on the sample of asymptomatic subjects and consequently check the suitability of the established protocols for their clinical application.

In fact, as observed in other studies (Baydal-Bertomeu et al., 2004), we did actually observe an increase in postural sway as the difficulty of the static Romberg tests increased, both when vision was deprived and proprioception was distorted with a foam rubber pad; when these two impediments were combined, the postural sway was even greater (e.g., the area defined by the CoP is eight times greater in the SSEC test than in the RSEO test).

In the LoS test for dynamic assessment, another consistent pattern is obtaining a greater reach in the anterior zone compared to the posterior one. This is because the posterior zone is more critical, since the area supporting the feet is smaller (Baydal-Bertomeu et al., 2004). In the cases studied here, there is a difference of 25–30 mm between the anterior and posterior zones. Furthermore, we did not detect any significant differences in the "*sucess*" variable for the eight different directions (mean of 79.3%). Therefore, we can state that the results for the success variable in healthy subjects are independent of the direction of the stability axes.

Analysis of the LoS test results reveals significant differences (p-value = .023) in the reaches achieved along the left and right posterior axes (64.1 and 60.5 mm, respectively). We believe this is due to the "experience factor" detected during data analysis. This occurred because subjects adapted to the test as they were performing it, thus resulting in greater CoP reaches along the axes examined towards the end of the test, which corresponded to those on the left-hand side of the octagon. We suggest future studies randomise the order in which the red target ball travels along the axes to avoid this "experience factor".

Note that certain limitations have been found in the study. A larger evaluated population sample (although the sample was sufficient) would be advisable. With this population increase, the accuracy of the study could be improved and would allow extending the studied age range as well as allow consolidation of the data.

In view of the results of the study, the adequacy of the instrumentation used and the procedures and protocols for clinical application are highlighted, emphasising the importance of generating information about the postural behaviour of healthy individuals. Additionally, participants evaluated the study positively, as shown by information collected during questionnaires; all subjects understood the test instructions correctly.

5. Conclusion

It has been verified that the instrumentation meets the standards established (Scoppa et al., 2013) for its clinical application. As a result of this study, we have managed to elaborate and formalise a series of tests to assess balance in a standing position, based on literature, which is of interest in healthcare and forensic/-

expert settings. These tests measure the level of balance control and the limits of stability of the subjects as well as the degree of consistency and collaboration implemented during the tests. The selection of proposed variables and the reference data obtained from them in the sample of asymptomatic participants may favour the application of this technology to the study of balance in the health field.

Acknowledgments

The authors wish to thank *Alejandro Moreno, physiotherapist*, for his support carrying out the tests.

We also thank the I3A – University Institute of Research of Engineering of Aragon, University of Zaragoza, Zaragoza, Spain, for the materials they provided and the time in their Biomechanics Laboratory where the tests were carried out.

This project was co-funded by the Government of Aragon, the European Regional Development Fund and the University of Zaragoza.

Conflict of interest statement

The authors declare that they have not conflicts of interest.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.jbiomech.2017.10.013.

References

Alonso, A.C., Luna, N.M.S., Mochizuki, L., Barbieri, F., Santos, S., Greve, J.M.D., 2012. The influence of anthropometric factors on postural balance: The relationship between body composition and posturographic measurements in young adults. Clinics 67 (12), 1433–1441.

Ayllón, F.N., Fernández, D.F., 2006. Función y entrenamiento de la musculatura abdominal. una visión científica. J. Human Sport Exercise.

Baydal-Bertomeu, J.M., Barberà, I., Guillem, R., Soler-Gracia, C., Peydro De Moya, M. F., Prat, J.M., Barona De Guzmán, R., 2004. Determinación de los patrones de comportamiento postural en población sana española. Acta Otorrinolaringológica Española. 55 (6), 260–269. https://doi.org/10.1016/S0001-6519(04)78520-9.

Borges, A.P.O., Carneiro, J.A.O., Zaia, J.E., Carneiro, A.A.O., Takayanagui, O.M., 2016. Evaluation of postural balance in mild cognitive impairment through a three-dimensional electromagnetic system. Braz. J. Otorhi. 82 (4), 433–441.

Bruniera, J.R.Z., Camiloti, J.F., Penha, O.d.M., Franco, P.P.R., Junior, S., Marchiori, Luciana Lozza de Moraes, 2015. Comparative analysis of postural balance by posturography in patients with isolated vertigo or associated with hearing loss. Audiol.-Commun. Res. 20 (4), 321–326.

Caña-Pino, A., Apolo-Arenas, M.D., Moral-Blanco, J., Álvaro-de Diego, J., Fernández, Gutiérrez C., 2015. Valoración del equilibrio postural en bipedestación-sedestación en sujetos sanos mediante acelerometría. estudio piloto. Fisioterapia 37 (6), 271–278.

Corazza, S.T., Laux, R.C., Foesch, M.L.S., et al., 2016. Benefits of functional training for the balance and propioception in visually impaired persons. Revista Brasileira de Medicina do Esporte 22 (6), 471–475.

Cuozzo Lemos, L.F., de David, A.C., Bolli, Mota C., 2016. Development of postural balance in brazilian children aged 4–10 years compared to young adults. Braz. J. Kineanthropometry Human Perform. 18 (4), 419.

Cvecka, J., Cacek, J., Ondracek, J., 2014. Agings and the parameters in static postural way. J. Human Sport Exercise 9(Proc1), S533–S537. doi: 10.14198/jhse.2014.9.Proc1.43.

de Sá Ferreira, A., Baracat, P.J.F., 2014. Test–retest reliability for assessment of postural stability using center of pressure spatial patterns of three-dimensional statokinesigrams in young health participants. J. Biomech. 47 (12), 2919–2924.

Dorneles, P.P., Meereis, E.C.W., Pranke, G.I., Mota, C.B., RELAÇÃO DO ÍNDICE DO ARCO PLANTAR COM O EQUILÍBRIO POSTURAL-DOI: http://dx. doi. org/10.18511/0103-1716/rbcm. v22n2p114-120. Revista Brasileira de Ciência e Movimento 22(2), 114–120.

Duarte, M., Freitas, S.M., 2010. Revision of posturography based on force plate for balance evaluation. Braz. J. Phys. Ther. 14 (3), 183–192.

García, R.B., Corresa, S.P., Bertomeu, J.M.B., Suárez-Varela, M.M.M., 2012. Posturografía estática con pruebas dinámicas. utilidad de los parámetros biomecánicos en la valoración del paciente vestibular. Acta Otorrinolaringológica Española 63 (5), 332–338.

Gil-Agudo, A., Baydal-Bertomeu, J., Fernández-Bravo, C., et al., 2006. Determinación de parámetros cinéticos en las pruebas de equilibrio y marcha de pacientes con latigazo cervical. Rehabilitación 40 (3), 141–149.

Goebel, J.A., Sataloff, R.T., Hanson, J.M., Nashner, L.M., Hirshout, D.S., Sokolow, C.C., 1997. Posturographic evidence of nonorganic sway patterns in normal subjects, patients, and suspected malingerers. Otolaryngol.-Head Neck Surg. 117 (4), 293–302.

González, M.Á.M., Villegas, A.S., Atucha, E.T., Fajardo, F.J.F., 2014. Bioestadística Amigable. Elsevier, España.

Kapteyn, T., Wit, G.d., 1972. Posturography as an auxiliary in vestibular investigation. Acta Otolaryngol. 73 (2–6), 104–111.

Kapteyn, T., Bles, W., Njiokiktjien, C., Kodde, L., Massen, C., Mol, J., 1983. Standardization in platform stabilometry being a part of posturography. Agressologie 24(7), 321–326.

Khasnis, A., Gokula, R.M., 2003. Romberg's test. J. Postgrad. Med. 49 (2), 169–172.

Korhonen, M.T., Suominen, H., Viitasalo, J.T., Liikavainio, T., Alen, M., Mero, A.A., 2010. Variability and symmetry of force platform variables in maximum-speed running in young and older athletes. J. Appl. Biomech. 26 (3), 357–366.

Ku, P.X., Abu Osman, N.A., Wan Abas, W.A.B., 2016. The limits of stability and muscle activity in middle-aged adults during static and dynamic stance. J. Biomech. 49 (16), 3943–3948.

López, L.F.C., Calidonio, J.D.L., 2009. "CgMed" design and construction of a platform to determine the position of the center of gravity in standing. Revista Ingeniería Biomédica 3 (6), 26–36.

Ma, C.Z., Wong, D.W., Lam, W.K., Wan, A.H., Lee, W.C., 2016. Balance improvement effects of biofeedback systems with state-of-the-art wearable sensors: a systematic review. Sensors 16 (4), 434.

Maze, F., Blázquez-Tejada, M.T., Rojas, F.J., 2016. Exploring body sway to disclose changes in postural control strategy associated with proprioceptive training. Eur. J. Human Movement 37, 1–20.

Parreira, R.B., Boer, M.C., Rabello, L., Costa, Viviane de Souza P., de Oliveira Jr., E., da Silva, R.A., 2013. Age-related differences in center of pressure measures during one-leg stance are time dependent. J. Appl. Biomech. 29 (3), 312–316.

Patrícia Paludette, D., Fabrício Santana da, S., Carlos Bolli, M., 2015. Comparação do equilíbrio postural entre grupos de mulheres com diferentes faixas etárias/comparison of postural balance among groups of women with different age ranges/comparación del balance postural en un grupo de mujeres de franjas etarias distintas. Fisioterapia e Pesquisa 22 (4), 392.

Pawar, P.K., Dadhich, A., 2012. Study of correlation between human height and hand length in residents of mumbai. Int. J. Biol. Med. Res. 3 (3), 2072–2075.

Peydro de Moya, M.F., Baydal Bertomeu, J.M., Vivas Broseta, M.J., 2005. Evaluación y rehabilitación del equilibrio mediante posturografía. Rehabilitación 39 (6), 315–323. https://doi.org/10.1016/S0048-7120(05)74365-6.

phidgets. Load cells. 2017 (17 May).

Pinsault, N., Vuillerme, N., 2009. Test–retest reliability of centre of foot pressure measures to assess postural control during unperturbed stance. Med. Eng. Phys. 31 (2), 276–286.

Postolache, O.A., Postolache, G.B., 2017. Development and selection of balance sensing devices. IEEE Instrum. Meas. Mag. 20 (1), 38–48.

Ramírez, C., Ordi, G., Fernández, S., Mi, C.M., 2014. Detección de exageración de síntomas en esguince cervical: pacientes clínicos versus sujetos análogos. Trauma 25 (1), 4–10.

Rubenstein, L.Z., Josephson, K.R., 2002. The epidemiology of falls and syncope. Clin. Geriatr. Med. 18 (2), 141–158.

Sample, R.B., Jackson, K., Kinney, A.L., Diestelkamp, W.S., Reinert, S.S., Bigelow, K.E., 2016. Manual and cognitive dual tasks contribute to fall-risk differentiation in posturography measures. J. Appl. Biomech. 32 (6), 541–547.

Scoppa, F., Capra, R., Gallamini, M., Shiffer, R., 2013. Clinical stabilometry standardization: basic definitions–acquisition interval–sampling frequency. Gait Posture 37 (2), 290–292.

Solovykh, E.A., Maksimovskaya, L.N., Bugrovetskaya, O.G., Bugrovetskaya, E.A., 2011. Comparative analysis of methods for evaluation of stabilometry parameters. Bull. Exp. Biol. Med. 152 (2), 266–272.

Taylor, M.R., Sutton, E.E., Diestelkamp, W.S., Bigelow, K.E., 2015. Subtle differences during posturography testing can influence postural sway results: the effects of talking, time before data acquisition, and visual fixation. J. Appl. Biomech. 31 (5), 324–329.

Terrier, P., Reynard, F., 2014. To what extent does not wearing shoes affect the local dynamic stability of walking?: Effect size and intrasession repeatability. J. Appl. Biomech. 30 (2), 305–309.

Tinetti, M.E., 2003. Preventing falls in elderly persons. N. Engl. J. Med. 348 (1), 42–49.

Toledo, E., Olivera, R., 2012. Desarrollo de una plataforma dinamométrica para el cálculo del centro de presiones utilizando galgas extensiométricas.

Vanmeerhaeghe, A.F., Rodriguez, D.R., Tutusaus, L.C., Calafat, C.B., Riera, M.L., Vidal, A.M., 2009. Diferencias en la estabilidad postural estática y dinámica según sexo y pierna dominante. Apunts. Medicina de l'Esport 44 (162), 74–81.

Weaver, T.B., Ma, C., Laing, A.C., 2016. Use of the nintendo wii balance board for studying standing static balance control: technical considerations, force-plate congruency, and the effect of battery life. J. Appl. Biomech. 1–22.

Wolf, F., Krebs, R.J., Detânico, R.C., Keulen, G., Braga, R., 2008. Estudo do equilíbrio plantar do iniciante de tiro com arco recurvo. Rev. Educ. Fís/UEM 19 (1), 1–9.

Yamamoto, M., Ishikawa, K., Aoki, M., et al., 2017. Japanese standard for clinical stabilometry assessment: current status and future directions. Auris Nasus Larynx.

4.2. ARTÍCULO 2

De la Torre, J.; Marin, J.; Polo, M.; Marín, J.J. Applying the Minimal Detectable Change of a Static and Dynamic Balance Test Using a Portable Stabilometric Platform to Individually Assess Patients with Balance Disorders. *Healthcare (S.I. Ageing Effects on Kinematics, Kinetics and Balance)* **2020**, *8*, 402, https://doi.org/10.3390/healthcare8040402

HEALTHCARE

ISSN: 2227-9032

Factor de Impacto JCR (2019):

1.916

Áreas Temáticas:

Health Care Sciences & Services: Q3 (62/102)

healthcare

MDPI

Article

Applying the Minimal Detectable Change of a Static and Dynamic Balance Test Using a Portable Stabilometric Platform to Individually Assess Patients with Balance Disorders

Juan De la Torre [1,*], Javier Marin [1], Marco Polo [2] and José J. Marín [1,3]

[1] IDERGO-Research and Development in Ergonomics, Biomechanical Laboratory, I3A-University Institute of Research of Engineering of Aragon, University of Zaragoza, 50009 Zaragoza, Spain; 647473@unizar.es (J.M.); jjmarin@unizar.es (J.J.M.)
[2] Physical Medicine and Rehabilitation, Hospital of Alcañiz, 44600 Teruel, Spain; marcopoloroyo@gmail.com
[3] Department of Design and Manufacturing Engineering, University of Zaragoza, 50009 Zaragoza, Spain
* Correspondence: 627471@unizar.es

Received: 22 July 2020; Accepted: 12 October 2020; Published: 14 October 2020

Abstract: Balance disorders have a high prevalence among elderly people in developed countries, and falls resulting from balance disorders involve high healthcare costs. Therefore, tools and indicators are necessary to assess the response to treatments. Therefore, the aim of this study is to detect relevant changes through minimal detectable change (MDC) values in patients with balance disorders, specifically with vertigo. A test-retest of a static and dynamic balance test was conducted on 34 healthy young volunteer subjects using a portable stabilometric platform. Afterwards, in order to show the MDC applicability, eight patients diagnosed with balance disorders characterized by vertigo of vestibular origin performed the balance test before and after a treatment, contrasting the results with the assessment by a specialist physician. The balance test consisted of four tasks from the Romberg test for static balance control, assessing dynamic postural balance through the limits of stability (LOS). The results obtained in the test-retest show the reproducibility of the system as being similar to or better than those found in the literature. Regarding the static balance variables with the lowest MDC value, we highlight the average velocity of the center of pressure (COP) in all tasks and the root mean square (RMS), the area, and the mediolateral displacement in soft surface, with eyes closed. In LOS, all COP limits and the average speed of the COP and RMS were highlighted. Of the eight patients assessed, an agreement between the specialist physician and the balance test results exists in six of them, and for two of the patients, the specialist physician reported no progression, whereas the balance test showed worsening. Patients showed changes that exceeded the MDC values, and these changes were correlated with the results reported by the specialist physician. We conclude that (at least for these eight patients) certain variables were sufficiently sensitive to detect changes linked to balance progression. This is intended to improve decision making and individualized patient monitoring.

Keywords: posturography; vertigo; vestibular disorders; test-retest; patient-level analysis

1. Introduction

Balance disorders have a high prevalence among elderly people in developed countries [1]. Connected with the current trend of increasing age of the population [2,3], this has led to a rise in pathologies that affect balance, causing an increased risk of falls in the elderly population. Around 30% of people aged over 65 years, and more than 50% of individuals in health centers or care homes suffer

one or more falls per year, and between approximately 20% and 30% of the global population has had or will have a vertiginous episode in their lives [4–10].

Therefore, it is necessary to have tools and indicators to assess the response to treatments of pathologies that affect the sensory systems involved in balance: visual, vestibular, and proprioceptive [11]. The effects of a given treatment on a patient through the performance of previous and subsequent tests can help in making decisions about adjusting, changing, or stopping the treatment [12].

Stabilometric platforms are useful in assessing balance because they obtain numerous parameters of the centers or pressure (COP) [13,14]. Stabilometric platforms can assess static balance control and dynamic postural balance through different variables and application methods [15,16]. On the one hand, in static balance control methods (such as the Romberg test), subjects must maintain their COP within the support base throughout the assessment period of time. On the other hand, the assessment of dynamic postural balance, which is vital for motor control, involves measuring the limits of stability (LOS), corresponding to the maximum voluntary angle or distance in which an individual can regulate their COP in a given direction without losing balance [17]. Stabilometric platforms can obtain objective information related to balance pathologies in clinical practice to improve the quality of healthcare and the provided treatments [4,18–20].

Therefore, a balance test would be beneficial for clinicians by providing them with an objective assessment device; however, the integration of this type of test in clinical practice involves various difficulties and constraints. Although stabilometric platforms are normally used in the lab, their implementation in the clinic is complex due to difficulty of use and manageability, as well as size, which may call into question clinical applicability [21]. The protocol must be validated and adapted for continuous and assiduous use in the clinic [21,22]. Therefore, it is relevant to conduct studies to measure the degree of variability of a balance test to achieve a successful clinical application.

The minimal detectable change (MDC) index represents the variability of the measures of each variable, i.e., the consistency of the variables, which includes different aspects: the variability of the instrument, the inherent variability of the person and their learning (test duration, visual condition, or position of the feet) and the procedures and protocols applied to perform the test. Therefore, if we obtain a change of one variable that is higher than its MDC value, this would be a relevant change, which could be caused by the treatment effect and not by the test variability [23,24].

No balance assessment studies, based on the MDC results of the test-retest, to detect changes in the results associated with balance progression have been conducted for patients with balance disorders. Numerous test-retest studies have been found in the COP measures but have not included patients with balance disorders in the application [4,19,24–32].

Therefore, the aim of this study is to detect relevant changes through MDC values in patients with balance disorders, specifically with vertigo. To address this, the following research actions/steps are proposed: (1) to perform a test-retest of a static and dynamic balance test in a sample of healthy subjects identifying the most sensible variables with the lower MDC values; (2) to analyze whether relevant changes are detected in the results of the balance test before and after the application of a treatment in eight patients diagnosed with balance disorders with vertigo of vestibular origin, contrasting the results with the progression observed by a specialist physician.

2. Materials and Methods

2.1. Participants and Ethics Statement

A test-retest study, which consists of repeating the test at two different times with a homogeneous sample of participants under the same conditions, of a static and dynamic balance test was conducted on a sample of 34 healthy young volunteer subjects. Participating in this study were 20 males and 14 females, (age 22.89 ± 3.51 years, height 172.51 ± 9.01 cm, weight 67.38 ± 11.82 kg, body fat index 20.07 ± 9.12 %, foot length [33] 25.39 ± 1.81 cm, abdominal perimeter 79.76 ± 9.77 cm).

The calculation for the choice of 34 subjects was based on the research of Bujang et al. [34], an article in which the relationship between the interclass correlation coefficient (ICC), statistical power, and the number of subjects is established. Taking into account the necessity for two measurements to be made per subject, in order to satisfy the requirement for repetition of the test at two different times, to set a statistical power of 80% and to establish a minimum ICC of 0.5, the sample for analysis needed to number at least 22. Therefore, the choice of 34 subjects is considered appropriate.

The inclusion criteria established for participation in the study were (i) no history of neurological, visual, or vestibular alterations, (ii) no record of musculoskeletal or neurological diseases in the last year, and (iii) no history of limb surgery that may affect the patient's balance.

It has been demonstrated in several studies that the variability of variables related to the COP are higher in young subjects, which has motivated the selection of the sample [4,28].

As shown in Section 2.5, eight patients with balance disorders characterized by vertigo of vestibular origin were also analysed to apply and assess the test-retest results. The patients were analysed before and after treatment to obtain the individual results of their progression.

The choice of eight patients was based on the sample size calculation for quantitative variables of Charan et al. [35]. The formula used was the following:

$$N = (Z_{1-\alpha/2})^2 \times \mathrm{SD}^2 / d^2 \quad (1)$$

where $Z_{1-\alpha/2}$ is the standard normal variate with $p_v < 0.05$ (error type 1), corresponding to value of 1.96 (in the majority of studies p_v values are considered significant below 0.05, hence 1.96 is used in the formula); SD is the standard deviation of the variable (standard deviation value can be taken from previously studies); and d is the absolute error or precision (decided by researcher).

The SD has been selected from the study conducted by Balaguer et al. [20], for using variables and tests similar to those used in this study. Specifically, the test in which all the sensory systems are available has been considered (rigid surface with eyes open), selecting one of the most used variables in balance studies [29,36]: the average speed of the COP. The SD value, for this variable and for this test is 6 mm/s; selecting a d value of 4.1 mm/s, and maintaining a $Z_{1-\alpha/2}$ value of 1.96, we obtained an N value of 8.22 in the eight selected patients.

The selected patients met the following inclusion criteria: (i) between 65 and 75 years old and (ii) having suffered a vertiginous episode in the last year. The following were the exclusion criteria: (i) presented acute osteo-muscular pathology in the lower limbs or lumbar spine, which may alter the outcome of the stabilometric platform, (ii) presented any amputation in the lower limbs, (iii) presented oncological pathology or was in active treatment with chemotherapy, radiotherapy, or hormonal therapy, (iv) presented degenerative diseases (which can cause loss of capacity during the assessment period of time), (v) having been influenced by work or family events (or other kind of events) that have detrimentally affected the patient's mood.

The study was approved by the Government of Aragon's Human Research Ethics Committee (CEICA) (16 January 2019). Prior to the beginning of the tests, the participants signed a form, consenting to undergo them and indicating that they understood the aim of the study.

2.2. Instrumentation

The device used was the stabilometric platform MoveHuman-Dyna UZ, which was designed and manufactured by the IDERGO (Research and Development in Ergonomics, University of Zaragoza, Spain) research group (Figure 1). This is a static posturography device designed for research, which comprises four load cells and a lightweight aluminum structure, whose dimensions and characteristics are detailed in the study of [18]. The findings of this device can be replicated in a straightforward manner by other researchers, which enhances the applicability of this study. The acquisition and processing of the platform data, as well as the format and method of exporting them, have been carried out according to the procedure used by [18]. Processing the force

data as a function of the cells' position means we can calculate the real-time position of the trajectory that describes the position of the CoP by applying the appropriate formula [7,13].

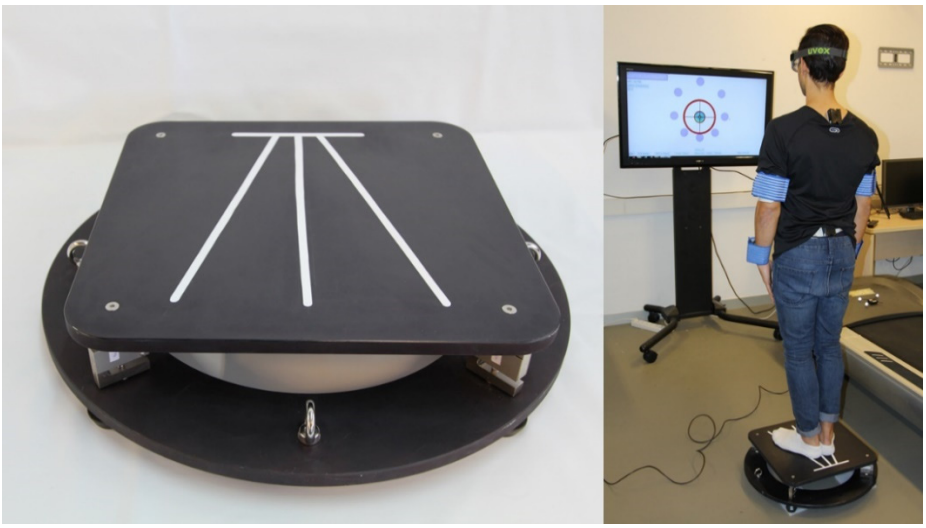


Figure 1. Stabilometric platform and environmental test condition.

Likewise, in accordance with the aforementioned study, the stabilometric platform "meets the standards established by the International Society for Posture and Gait Research (ISPGR) for its clinical application" [37] in relation to various parameters, such as accuracy, precision, linearity, dimensions, resolution, sampling, etc. The precision parameters (accuracy, precision, linearity, dimensions and resolution) were obtained through an experiment in which the metrological characteristics of the platform were tested with a gold standard force platform, along with the error of measurement [17]. Specifically, the gold standard force platform with which it was compared was the AMTI 0R6-7-1000, U.S. Patent Number 4.493.220. The two platforms were positioned one above the other, in order to enable their results to be directly compared and examined for any variability, according to the study of Huurnink et al. (2013) [38]. The stabilometric platform has since been used in several research projects with patients in different hospitals, both public and private; all these research projects have been approved by the CEICA Committee.

The use of this stabilometric platform in this study was justified by its portability characteristics, dimensions, and weight that enable its use by physicians in an examination room, often with limited space [13,18,37,39,40]. A portable device such this platform allows its use in different medical centers due to the quicker installation process and smaller space required, which improves its accessibility and applicability [41–43].

2.3. Protocol

The static and dynamic balance were both assessed with a set of tests previously applied in other studies [18]. The static balance control was assessed with a test based on the Romberg test and the Modified Clinical Test of Sensory Interaction in Balance (CTSIB-M), with consideration given to four different situations: (1) rigid surface with eyes open (RSEO), (2) rigid surface with eyes closed (RSEC), (3) soft surface with eyes open (SSEO), and (4) soft surface with eyes closed (SSEC).

On the other hand, the dynamic postural balance was assessed measuring the LOS that a subject is able to reach and, with it, the management capacity of COP [17]. The dynamic LOS test was based on protocols found in the literature [44]. The inclusion of the LOS, complementary to the assessment of the static balance control, provides additional value to the balance assessment protocol [4,8,45]

The protocol applied in the tests (the position of the body, arms and feet during the test [18], environmental conditions (e.g., noise, space, etc.) and the additional instrumentation used as a

foam rubber for soft surface, instruments for anthropometric data collection, etc.) is that used by Delatorre et al (2017) for this stabilometric platform (Figure 1). This protocol fulfils certain clinical conditions [21,46–48]; it must be fast and should not require multiple repetitions to issue a definite, consistent result [21]. It must also be effective, with clear, specific stages defined by the instructions given by the operator that are understandable by the patient [28].

The test-retest in the sample of healthy subjects was performed on the same day with an interval of 6 hours between the two trials. In this interval, the participants did not perform physical activities that required extra muscular activation and/or caused fatigue that could affect the results of the second trial. The tests of all study participants were coordinated by the same operator. The time interval between the tests performed by the patients was 3 months (process detailed in Section 2.5).

2.4. Statistical Analysis: Test-Retest Study

2.4.1. Selection of Variables

The variables selected for the present test-retest study were those determined by [18] to be more significant in balance assessment studies, whose details and methods are also explained in the same study. The variables selected for the assessment of static and dynamic balance were the range of displacement in the anteroposterior and mediolateral directions, area (surface area covered by the trajectory of the COP), average speed of the COP, and root mean square (RMS) position. Additionally, in the LOS test, two more variables were assessed: the COP limits (maximum displacement reached along each axis of the octagon radii), and the "success" variable (quantification of the management and coordination of the COP along each axis of the octagon radii), both defined in a previous study [18].

2.4.2. Minimal Detectable Change Calculation

The MDC index represents the variability of the measures of each variable. If a change of one variable that is lower than its MDC value is detected, it would not be considered relevant, since it is lower than the variability of the test. Therefore, to narrow the intrinsic variability of the test, we considered that the best choice would be to perform the test-retest with healthy and young people (this topic will be justified in the discussion section).

The MDC index was calculated for each of the variables resulting from the test-retest study. The value of the MDC was calculated from the following Equations (2) and (3) [27,30,49–53]:

$$\text{MDC95\%} = 1.96 \times \sqrt{2}\,\text{SEM} \tag{2}$$

$$\text{SEM} = \text{SD-pooled} \times \sqrt{(1 - r)} \tag{3}$$

where r is the interclass correlation coefficient (ICC), SD-pooled is the pooled average of the standard deviation of the test and retest, and the SEM is the standard deviation error of measurement.

The dimensionless value of the effect size (MDC.es 95%) was also calculated using Equation (4), which indicates the number of standard deviations that the experiment is capable of detecting [54]:

$$\text{MDC.es 95\%} = \text{MDC95\%} / \text{SD test} \tag{4}$$

where SD test is the standard deviation of the test (initial test).

The ICC results can be classified according to Cicchetti (1994) [55], who provided the following intervals to characterize the ICC inter-rater agreement measures: below 0.40: poor; between 0.40 and 0.59: fair; between 0.60 and 0.74: good; between 0.75 and 1.00: excellent.

2.5. Clinical Application Foundations: Assessment of Patients' Progression

To verify whether the variables in the balance test could detect changes that exceed the MDC index before and after treatment, eight patients diagnosed with balance disorders characterized by

vertigo of vestibular origin performed the aforementioned test. This allowed us to verify whether the analyses of the results of the pre-tests and post-tests coincided with the subjective patient progression (positive, null or negative progression) perceived by a specialist physician.

2.5.1. Patients' Initial Diagnosis

The patients were referred by the Otorhinolaryngology Service of the Alcañiz Hospital (Teruel, Spain) after being diagnosed with a balance disorder. The methods for determining the vestibular deficit were medical history, magnetic resonance imaging, videonystagmography, and tests such as the Dix-Hallpike maneuver (Table 1).

Table 1. Data of patients assessed.

Patient Code	Age	Gender	Anamnesis and Examination
01	72	Male	BPPV with associated signs of bilateral hearing loss (detected by audiometry) and nystagmus.
02	73	Female	Ménière syndrome.
03	70	Male	BPPV with associated signs of bilateral hearing loss (detected by audiometry).
04	73	Female	Osteosclerosis. Bilateral hearing loss.
05	68	Male	Ménière syndrome. Unilateral tinnitus and pathological nystagmus.
06	74	Male	Ménière syndrome. Unilateral tinnitus.
07	68	Female	Vestibular hypofunction. Pathological nystagmus.
08	75	Male	BPPV with associated signs of bilateral hearing loss (detected by audiometry).

BPPV: Benign Paroxysmal Peripheral Vertigo.

2.5.2. Clinician 1 Assessment: History and Physical Examination

A doctor (clinician 1) from the Physical Medicine and Rehabilitation Department of the Alcañiz Hospital (Teruel, Spain) evaluated the patients using medical history and a physical examination, as well as functional balance assessment tests such as the Unterberger test [56,57], the up and go test [58,59], and unipodal support test [60]. Clinician 1 prescribed the rehabilitation treatment, which consists of a set of vestibular rehabilitation exercises (to be performed by the patients), which is commonly used in the clinic, establishing these over a period of three months [61,62]. Clinician 1 evaluated the patients at two different times, once before starting the treatment (pre-data), and once three months after starting the treatment (post-data).

2.5.3. Clinician 2 Assessment: Patient Progression Evaluation

The pre- and post- data collected by clinician 1 were assessed by a specialist physician (clinician 2), which allowed an assessment of the balance progression of each of the eight patients. To avoid the results being influenced or contaminated by the interaction between the clinicians, there was no contact between them during the research. The assessment of clinician 2 established three possible categories to evaluate patient progression: positive, null or negative progression.

2.5.4. Magnitude-Based Decision (MBD) to Monitor Patients with Balance Disorders

To measure the effects of a treatment on a specific patient, it is necessary to evaluate the treatment status at different times. Thus, the patient-level approach proposed by Hopkings (2017) [63] was applied to obtain personalized results for each patient. This approach allows the assessment the change between two measurements in an individual through the magnitude-based decision (MBD) method (formerly known as magnitude-based inferences) [64].

Hopkings (2017) [63] introduced the patient-level approach providing "A Spreadsheet for Monitoring an Individual's Changes". Based on this spreadsheet, we developed a script that applies the MBD method using as input the measurements taken from the pre- and post-balance tests of one specific patient, and the threshold MDC95% previously calculated in the test-retest study. The script was developed using WorldViz-Vizard 6.2 (based on Python 2.7), and the Pandas and Matplotlib libraries.

The patients in the study performed the Romberg and LOS tests in the rehabilitation clinic of the Alcañiz Hospital on two occasions, once before starting the treatment and another three months after starting the treatment. Using the script that we have developed, it is possible to measure how much each variable has changed and whether the change detected is significant. According to the MBD method, inputs are required to analyze each variable:

- Xdif: difference between the measures taken at two temporal points: pre-value and post-value (Equation (5)). In this case, the pre-value is the measure of each variable just before starting the treatment; and the post-value is the measure three months after starting the treatment.

$$Xdif = X_{post} - X_{pre} \quad (5)$$

- MBD threshold: for this method, a threshold (numerical value) must be defined from which a change is considered relevant. In our case, we selected the MDC previously calculated.

In this approach, the statistical analysis detects whether the changes exceed a particular threshold, in our case the MDC value. Specifically, we determined where the confidence interval of the difference was located (between pre- and post-tests) in relation to the thresholds of the MDC [54,65]. The intervals of the differences of each variable between the pre-test and post-test results were determined to be on the negative side of the MDC threshold (% negative differences), within the threshold (% trivial differences), or on the positive side (% positive differences), respectively [63]. We deal with the selection of the MDC as threshold in the discussion section.

Firstly, the value and sign (positive or negative) of Xdif is obtained through the difference between the pre-value and post-value. Subsequently, following the calculation method set forth by [63], the probability of change is obtained, which can be defined as the probability that the difference between the two values is relevant. This probability corresponds to the percentage of the confidence interval of the difference (calculated using the Xdif) that is outside of the range (+MDC, −MDC). Finally, the probability that the change is relevant was qualitatively classified (as proposed by Batterham and Hopkins (2006)) [54]. The qualitative classification of the significance of the changes is: most unlikely (<1%), very unlikely (1% to 5%), unlikely (5% to 25%), possibly (25% to 75%), probably (75% to 95%), very likely (95% to 99%) and most likely (>99%).

For the calculations and different graphs presented in this study, Python 2.7 and the numpy, scipy and pandas modules were used.

3. Results

3.1. Test-Retest Results from Balance Analysis. Minimal Detectable Changes Index

Table 3 shows the results of the test-retest of the variables selected. The values of the means (μ) and standard deviations (SD) of each of the analyzed variables are shown by task. Likewise, the results of the variability through $ICC_{3,k}$ (similar to $ICC_{2,1}$) [24], absolute value of the MDC (95%), and dimensionless value of the effect size (MDC.es 95%) are included.

Table 2. Test-retest results from balance tests. Minimal detectable changes index.

Balance Tasks	Variables	Test μ (SD)	Retest μ (SD)	ICC	MDC95_es	MDC95
RSEO	COP mean speed [mm/s] *	8.1 (2.3)	8.4 (2.2)	0.87	0.9	2.3
	RMS [mm]	5.1 (1.2)	5.3 (1.7)	0.61	2.1	2.6
	Area [cm^2]	1.7 (0.9)	1.7 (0.9)	0.71	1.5	1.4
	AP disp. [mm]	19 (4.7)	19.5 (6.1)	0.42	2.4	11.6
	ML disp. [mm]	15.8 (4.4)	15.8 (4.8)	0.70	1.5	7.1
RSEC	COP mean speed [mm/s] *	13.9 (4.2)	13.6 (3.9)	0.92	0.7	3.3
	RMS [mm] *	6.2 (1.9)	6.5 (1.8)	0.87	0.9	1.9
	Area [cm^2] *	3.5 (1.9)	3.4 (1.8)	0.85	1.0	2.0
	AP disp. [mm]	24.3 (6.1)	25.9 (7.7)	0.75	1.5	9.7
	ML disp. [mm] *	22.4 (8.3)	22.7 (7.6)	0.85	1.0	8.7

Table 2. *Cont.*

Balance Tasks	Variables	Test μ (SD)	Retest μ (SD)	ICC	MDC95_es	MDC95
SSEO	COP mean speed [mm/s] *	13.7 (2.7)	13.5 (2.3)	0.88	0.9	2.5
	RMS [mm]	6.9 (1.3)	6.9 (1.5)	0.74	1.4	2.0
	Area [cm^2]	3.7 (1.0)	3.4 (1.1)	0.76	0.4	1.5
	AP disp. [mm]	27.5 (5.1)	27.2 (6.7)	0.72	1.7	8.8
	ML disp. [mm]	23.4 (4.3)	22.2 (4)	0.56	1.7	7.7
SSEC	COP mean speed [mm/s] *	36.2 (7.9)	34 (6.6)	0.83	1.0	8.5
	RMS [mm]	13.6 (2.2)	13.3 (2.3)	0.77	1.3	3.0
	Area [cm^2]	17.2 (5.3)	15.7 (4.7)	0.74	1.3	7.2
	AP disp. mm]	54.8 (12.4)	52.7 (10.9)	0.61	1.6	20.4
	ML disp. [mm]	51.6 (10.6)	50.4 (9.5)	0.42	2.0	21.4
LOS	COP mean speed [mm/s] *	15.9 (3.1)	15.4 (3.0)	0.93	0.7	2.3
	RMS [mm] *	5.8 (0.3)	5.9 (0.3)	0.96	0.5	0.2
	Area [cm^2] *	179.5 (43.7)	183.2 (45.2)	0.97	0.4	21.7
	AP disp. [mm]	153.8 (20.4)	157 (19.7)	0.94	0.7	14.4
	ML disp. [mm]	157.1 (20)	158 (23.1)	0.94	0.7	14.8
	Lim.COP.Forward [mm] *	84.6 (17.5)	87.5 (15.3)	0.9	0.8	15.0
	Lim.COP.Forward-rightward [mm] *	90.4 (14.4)	91.8 (11.8)	0.93	0.7	10.2
	Lim.COP.Rightward [mm]	77.3 (12.2)	75.5 (12.2)	0.83	1.1	14.0
	Lim.COP.Backward-rightward [mm]	68.2 (12.7)	68 (10.6)	0.88	0.8	11.4
	Lim.COP.Backward [mm]	68.1 (14.3)	68.5 (13.6)	0.84	1.0	15.6
	Lim.COP.Backward-leftward [mm]	71.2 (12.2)	71 (14.1)	0.85	1.1	14.4
	Lim.COP.Leftward [mm]	77.6 (10.5)	81.4 (13.8)	0.86	1.2	13.1
	Lim.COP.Forward-leftward [mm] *	91 (14.8)	92.6 (12.4)	0.89	0.8	12.5
	Success.Forward [%]	75.7 (14.4)	76.8 (11.7)	0.5	1.7	25.8
	Success.Forward-rightward [%]	79.4 (12.2)	79.1 (13.9)	0.87	1.0	13.2
	Success.Rightward [%]	78.3 (12.3)	80.4 (10.7)	0.68	1.4	18.1
	Success.Backward-rightward [%]	79.1 (11.6)	78 (9.9)	0.42	1.9	23.0
	Success.Backward [%]	83.5 (11.)	85.6 (8.2)	0.57	1.5	18.5
	Success.Backward-leftward [%]	80.4 (10.5)	82.8 (11.7)	0.2	2.6	27.7
	Success.Leftward [%]	80.8 (11.7)	81.6 (11.1)	0.8	1.2	14.4
	Success.Forward-leftward [%]	78.8 (13.3)	80.5 (11.4)	0.77	1.2	16.5

μ: mean; SD: standard deviation; ICC: intraclass correlation coefficient; MDC95_es: minimal detectable change in dimensionless value effect size at 95%.; MDC95: minimal detectable change in absolute value at 95%; RSEO: rigid surface, open eyes; RSEC: rigid surface, eyes close; SSEO: soft surface, eyes open; SSEC: soft surface, eyes close; LOS: limits of stability; COP: center of pressure; RMS: root mean square; AP: anteroposterior; ML: mediolateral. * Balance variables selected with higher ICC.

3.2. Results of the Patients-Level Study

Due to the large amount of information, the results of the patient with Code 01 are presented in this section as an example, and the results of the remaining patients are presented in the Supplementary Material. Thus, for each of the eight patients, the same information as shown in Table 3 for Patient 1 has been calculated, which includes the values of the (a) variables for the pre- and post-tests, (b) differences, (c) MDC, and (d) percentages of negative, trivial, and positive differences (-/0/+). In addition, Figures 2 and 3 show the produced changes (black lines) in relation to the MDC (light grey rectangle) in each variable for each of the tests.

Table 3. Results of the patient 01.

Balance Tasks	Variables	Value Pre	Value Post	Difference	MDC	% (− / 0 / +)
RSEO	COP mean speed [mm/s]	13.3	11.2	−2.0	2.3	41/59/0
	RMS [mm]	6.5	5.0	−1.5	2.6	22/78/0
	Area [cm^2]	3.4	2.1	−1.2	1.4	43/57/0
	AP disp. [mm]	20.0	17.0	−3.0	11.6	8/91/1
	ML disp. [mm]	22.4	17.7	−4.7	7.1	26/74/0
RSEC	COP mean speed [mm/s]	20.3	16.9	−3.4	3.3	53/47/0
	RMS [mm]	7.6	7.5	−0.1	1.9	4/94/2
	Area [cm^2]	5.1	3.3	−1.8	2.0	40/60/0
	AP disp. [mm]	26.9	31.6	4.7	9.7	0/84/16
	ML disp. [mm]	30.4	20.0	−10.4	8.7	65/35/0
SSEO	COP mean speed [mm/s]	18.5	18.0	−0.6	2.5	7/92/1
	RMS [mm]	7.3	6.8	−0.5	2.0	7/92/1
	Area [cm^2]	4.9	3.7	−1.1	1.5	33/67/0
	AP disp. [mm]	23.3	23.3	0.0	8.8	3/94/3
	ML disp. [mm]	44.3	28.4	−15.9	7.7	98/2/0

Table 3. *Cont.*

Balance Tasks	Variables	Value Pre	Value Post	Difference	MDC	% (− / 0 / +)
SSEC	COP mean speed [mm/s]	62.6	49.4	−13.2	8.5	86/14/0
	RMS [mm]	19.2	15.0	−4.2	3.0	78/22/0
	Area [cm^2]	41.0	25.4	−15.6	7.2	99/1/0
	AP disp. [mm]	102.4	62.2	−40.2	20.4	97/3/0
	ML disp. [mm]	66.5	72.5	6.0	21.4	1/91/8
LOS	COP mean speed [mm/s]	19.0	21.2	2.2	2.3	0/52/48
	RMS [mm]	5.5	6.2	0.7	0.2	0/0/100
	Area [cm^2]	170.0	234.0	64.0	21.7	0/0/100
	AP disp. [mm]	154.9	176.8	21.9	14.4	0/16/84
	ML disp. [mm]	150.3	179.5	29.2	14.8	0/3/97
	Lim.COP.Forward [mm]	86.8	108.2	21.4	15.0	0/21/79
	Lim.COP.Forward-rightward [mm]	9.5	108.3	15.8	10.2	0/14/86
	Lim.COP.Rightward [mm]	0.0	92.8	92.8	14.0	0/0/100
	Lim.COP.Backward-rightward [mm]	60.4	72.5	12.1	11.4	0/45/55
	Lim.COP.Backward [mm]	32.7	69.0	36.4	15.6	0/1/99
	Lim.COP.Backward-leftward [mm]	60.0	82.1	22.1	14.4	0/15/85
	Lim.COP.Leftward [mm]	72.7	87.0	14.3	13.1	0/43/57
	Lim.COP.Forward-leftward [mm]	106.6	104.7	−1.9	12.5	5/93/2
	Success.Forward [%]	50.7	68.8	18.2	25.8	0/72/28
	Success.Forward-rightward [%]	69.0	73.6	4.6	13.2	1/89/11
	Success.Rightward [%]	43.9	88.2	44.2	18.1	0/0/100
	Success.Backward-rightward [%]	45.3	81.5	36.2	23.0	0/13/87
	Success.Backward [%]	53.9	91.9	37.9	18.5	0/2/98
	Success.Backward-leftward [%]	43.6	70.2	26.6	27.7	0/53/47
	Success.Leftward [%]	55.9	75.1	19.1	14.4	0/26/74
	Success.Forward-leftward [%]	58.5	79.2	20.8	16.5	0/31/69

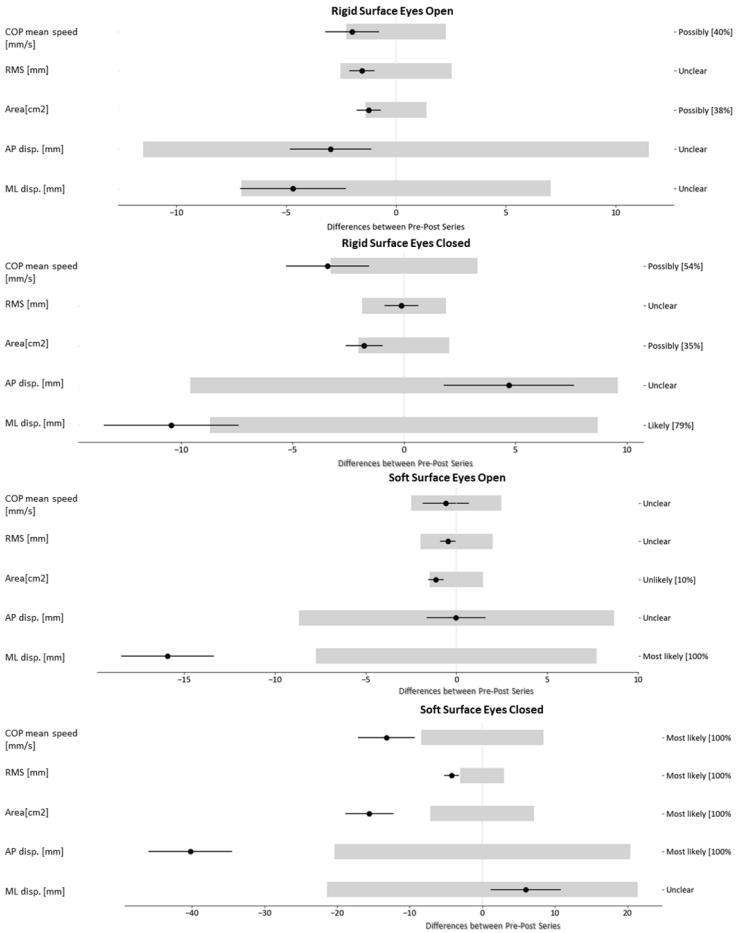


Figure 2. Patient level analysis. Romberg Test. Minimal detectable change (MDC): light grey bars; Differences: thin Black bars.

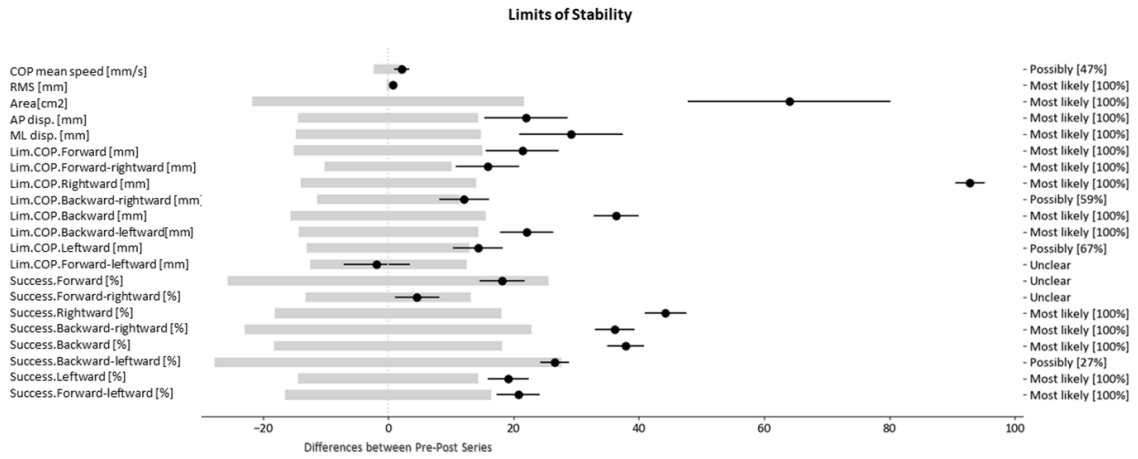


Figure 3. Patient level analysis. Limits of Stability. Minimal detectable change: light grey bars; Differences: Black thin bars.

As shown in Figures 2 and 3, the grey thresholds represent the MDC—or, what is the same—and the change in a variable must exceed the value of this interval in order for us to consider that the change is relevant. So as to quantify each change, the probability that it is relevant is shown in the right margin beside each of the variables.

Analyzing patient 01, most static variables decreased, highlighting the positive or negative differences according to the tests: RSEO: −2 mm/s of COP mean speed; RSEC: −3.4 mm/s of COP mean speed, −10.4 mm of ML displacement; SSEO: −15.9 mm of ML displacement; SSEC: −8.4 of COP mean speed, −4.2 mm in RMS, −15.6 cm^2 in area and −40.2 mm in AP displacement. In contrast, most dynamic variables increased: +15 mm in Lim.COP.Forward, +15.6 mm in Lim.COP.Backward, +14.3 mm in Lim.COP.Backward-leftward, +25.6% in Success.Forward, and +22.9% in Success.Backward-rightward.

According to the assessment of the specialist physician prior to and after three months from the administration of the treatment, it was observed that the associated clinical symptomatology decreased, and no vertiginous episodes occurred. The physician performed the following functional tests to assess the patient's balance before and after the treatment was applied: test up and go (15″–12″), Unterberger test (from positive to negative), and unipodal standing time on the right leg (4″–7″) and left leg (3″–10″). Objective improvement in all performed functional tests was demonstrated. Finally, the patient reported improvement in the anamnesis post-treatment, affirming progress in the performance of daily life activities, such as getting out of bed, walking without help, and showering.

The assessment by the specialist physician (positive progression of the patient) and the objective results of the pre- and post-tests (decreased static variables and increased dynamic variables, indicating improvement) coincide.

In relation to the remaining patients, as seen in the Supplementary Material, the physician evaluated two other patients (Codes 02 and 03) with positive progression who experienced objective improvement in the balance test variables. The physician assessed two patients (05 and 06) with negative progression who exhibited objective worsening in the balance test variables. The physician evaluated one patient (04) with no progression without a change in the balance test variables. In addition, the physician assessed another two patients (07 and 08) with no progression who obtained an objective worsening in certain balance test variables. In summary, of the eight patients, an agreement between the physician and the balance test results exists in six, and for two the physician reported no progression, whereas the balance test showed worsening in certain variables.

4. Discussion

In this study, a test-retest of the static and dynamic balance test was conducted on a sample of healthy young subjects to identify variables with a lower MDC value. Likewise, eight patients diagnosed with balance disorders characterized by vertigo of vestibular origin performed the test before and after treatment to verify whether the comparative analyses, pre- and post-tests, corroborated the subjective progression perceived by a specialist physician.

The test-retest results are satisfactory in comparison to similar studies found in the literature; specifically, they are improved compared to most of the studies reviewed by [24]. We highlighted the ICC values, which exceeded 0.7, of the COP mean speed and area of the four tasks of static balance, agreeing with the most recent studies in the review (we compare ICC values because they are the most used metrics in the studies consulted). Likewise, taking the study by Salvati et al. (2009) [66] as reference because it is the most recent and is similar to ours, our study shows improvement of the ICC of the area on the soft surface both with eyes open (0.66 vs. 0.33) and closed (0.77 vs. 0.64). Regarding the LOS results, we highlight the COP mean speed and RMS, which slightly improve the ICC values obtained by [21]. Likewise, comparing the ICC values in COP limits and success, greater values are obtained than reported by [25].

The usefulness of a test depends on the MDC value of the variables highlighting in our study the variables with lower MDC in Table 3 (marked with an *). Thus, the lower the MDC of the variables, the more useful the test will be and the greater capacity it will have to detect relevant changes. So the variability of the test is shown, and it is possible to detect variations in balance between different instances of time.

The presented results and the analysis of the patients have advantages of application; however, certain questions and decisions must be discussed, which are set out below.

The MDC prevents us from mistaking changes due to the 'noise' inherent in repeating a test over time. We understand that the existence of 'noise' is inseparable from this type of test because the active involvement of the patients is required.

It is necessary to justify and discuss the motivations and methodological approach chosen in this research. The motivation for the study was to determine the MDC in order to subsequently apply it to patients with balance disorders characterised by vertigo of vestibular origin. Deciding on the set-up for our experimentation, as has been described for the calculation of the MDC, was a matter that led discussions to arise among the multidisciplinary team that participated in this study.

In the design of the MDC calculation experimentation, it was prioritised that this should aim at studying—and quantitatively limiting—the variation (variability) between the results of pre- and post-testing, specifically, only from those factors intrinsic to the protocol applied and the instrumentation used. The intention was for these factors to be isolated from the noise associated with other factors such as the pathology, the treatment applied between pre- and post-testing, the disease progression, etc. Thus, a sample of healthy and young subjects was selected, since they would enable to study to determine the variability caused by factors such as the intrinsic variability of human beings, the conditions of the tests (body and foot position, duration of the test, visual condition and proprioception, etc.), the learning factor, normal incidents during the course of the tests, or the participant's own understanding of how to perform the test. Furthermore, it was considered that choosing a sample of healthy subjects would be beneficial for other similar medical tests/research, because in general the recruitment of healthy people is more reliable than accessing a homogeneous sample of patients with a specific pathology.

Likewise, in the experimentation, a period of six hours was established between the two tests for the calculation of the MDC, in order to exclude variability relating to time and, instead, focus on limiting the variables to the factors detailed above. Six hours was also intended to limit variability relating to certain physical activities that healthy subjects could perform between test and retest, (training for weight loss, to gain muscle mass, preparation for an important sporting event, etc.). The six-hour period was also considered to be pertinent as any fatigue factor that could have existed between two

closely followed tests was eliminated, as well as the learning factor. The subjects were also controlled in terms of food and drink intake that could affect the test (stimulant drinks, specific diets, etc.), as well as physical activities.

If we would have obtained an MDC from a group of patients with the same pathology and similar age to the presented patients, the influences of the specific treatment could be assessed and that is of interest. However, this restricts the MDC application framework to these patients and does not allow the assessment of changes beyond these patients due to the studied treatment.

According to the above, any change detected between pre- and post-testing in a patient, which exceeds the threshold of the MDC value, is known with certainty not to have been caused by the intrinsic factors already mentioned. Among the extrinsic factors of the test that may affect pre- and post-testing, the effect of the treatment and/or fluctuations of the pathology should be assessed by the clinician in order to analyse the results of the proposed method. However, it must be considered that other factors may also have affected the patient during the assessment period and blur the effects of the treatment; specifically, other diseases that may affect balance (e.g., a degenerative disease) or even personal events (e.g., the death of family members, physical accidents, etc.) that could detrimentally affect the patient's mood. Regarding this point, we proposed that these considerations should be included in the exclusion criteria included in this study, which could be used in studies of similar theme, in order to limit the extrinsic factors to be assessed from those relative to the treatment administered by the clinician and/or to the fluctuations of the pathology. The first three exclusion criteria were based on [67–70]; the criteria regarding the presentation of degenerative diseases was based on [71]; and that concerning the influence of personal events was based on [72–74]. Therefore, the clinician receives useful information since he can assess the change detected in a patient, excluding the intrinsic factors of the test, allowing him to clearly discern the progression that the patient has followed. This may be relevant to the clinician, who understands the situation of each patient and can inquire about the reasons for these changes.

It is of interest to note that, although a relevant change in a specific variable is obtained in a patient (high probability of change), it is the clinician who must assess whether this change is significant for the pathology that is being evaluated. This is important since, although this type of metric is intended to provide objective information to support a diagnosis, the MDC itself is not a direct diagnosis and should be used only as a reference. If the change detected for a variable is located "outside" the grey area that represents the MDC, it is the clinician who must assess whether this change represents a positive or negative progression, which may be the result of the treatment administered.

Following this line of argumentation, parallel to the study of the MDC for each variable, it would be of interest to inquire about the concept of the minimal important difference [52] (MID) applied in this field. The MID study involves a complex qualitative interpretation process, although its complexity should not detract from its development because the MID index should complement MDC values in the future. In this sense, [75] argued that both values are related, and if it is possible to define a MID value for a specific test, the experiment related to that test must have sufficient accuracy, which is defined by the value of the MDC. Thus, the MID value should preferably be applied, unless the value is lower than that of the MDC, which limits the accuracy of the system.

On the other hand, the use of the MoveHuman-Dyna UZ stabilometric platform does not intend to establish a direct relationship between the findings of the research and the instrumentation used, since it is a platform that can be replicated with easily available materials and devices (e.g., two aluminium plates and four load cells). It is considered that the results obtained and the method followed are easily replicable and can be extrapolated by other researchers; these results are intended to transcend the instrumentation used. The platform is piquing the interest of clinicians as its practical application in the clinic begins to be demonstrated. In relation to this point, it is considered that the platform used provides solutions to the problems of manageability and required space (recurrent in the use of stabilometric platforms), since it is a device with portable characteristics in terms of the little space it requires, its ease of use, and its simple installation.

Regarding the results presented in this study, we highlight the graphs shown in Figures 2 and 3. These graphs were designed to visually and intuitively show the changes detected in the set of variables, thus facilitating the clinician's quick identification of those variables that require special attention and analysis. Likewise, the opinions of the clinicians, collected in the follow-up meetings of the study, can be summarized in the following sentences: "The numerical results may not be very easy to understand, but with the graphs it is easier identify whether there has been a positive or negative change . . . "; "The use of this type of graph can help us when we are doubting whether to maintain or change treatment . . . ".

In the comparative analysis of the results of the patient with Code 01, the pre- and post-treatment balance tests demonstrated the positive progression of the patient, motivated by a decrease in numerous static variables and an increase in dynamics, possibly derived from recovery in the vestibular system of the patient. This improvement coincides with the evaluation of the positive progression of the patient by clinician 2.

Likewise, the balance test results and the specialist physician evaluations were compared for all patients (see Supplementary Material). Of the eight evaluated patients, agreement occurs between those established by clinician 2 and the balance test results in six out of eight cases. In the two mismatches (Patients 07 and 08), the specialist did not establish any changes in the evaluation of the patient, whereas the balance test detected changes that caused worsening in the balance (motivated by a latent process not detected by the doctor). Therefore, as future work, the sensitivity of the balance test should be deepened by analyzing the causes of the discrepancies between the results from the specialist physician and the balance test.

The sample size for comparison of the results of the balance test with the assessment of clinician 2 is perhaps not large. The goal is to exemplify the application of the results at the individual patient level; therefore, one patient would have been enough. However, a sample size of eight patients was obtained through a calculation based on the formula of Charan et al. [35], taking the research of Balaguer et al. [20] as reference; thus illustrating different cases (improvement, worsening, or maintenance).

The variables of each patient were individually analysed in an effort to discern which showed a greater relevant change in relation to the evaluation of clinician 2. On the one hand, among the variables related to the control of static balance, we highlight the following for each of the tests: RSEO, i.e., COP mean speed and ML displacement; RSEC, i.e., COP mean speed, AP displacement, and ML displacement; SSEO, i.e., ML displacement; and SSEC, i.e., COP mean speed, RMS, and area. On the other hand, the variables that presented the greatest relevant change for dynamic postural balance were area, AP displacement and ML displacement, Lim. COP Forward-rightward, Lim. COP Leftward, Success Rightward, Success Leftward, and Success Forward-rightward. Considering the statistical application of the MBD method in the sample of eight patients, these are the variables that are most reliable for assessing change at the individual level and, therefore, for assessing the efficacy of an intervention/treatment. The results have been compared with the study by Tsukamoto et al. (2015) [70] of patients with vestibular complaints, and the following variables coincide as those in which the greatest change was detected: RSEO, i.e., ML displacement; RSEC, i.e., COP mean speed; and SSEC, i.e., COP mean speed. The application of different tests, and the coincidence of only some variables in both studies, implies that there is no greater coincidence between the results.

Derived from the analysis of the patients' results with agreement from the specialist physician, certain 'logical' guidelines of interest were extracted to assess the degree of change as positive or negative, when interpreting the results of individual studies such as this one. A reduction in the variables related to the static balance tasks was associated with an improvement in balance, allowing us to discern whether the improvement comes from the visual, vestibular, or proprioceptive system, depending on in which task the improvement occurs. An increase in the LOS variables reflects an improvement in the dynamic balance and greater control of the COP.

Personalized medicine is a recurrent goal for health professionals [76]. Individual patient assessment, such as that proposed in this study, objectively characterizes the response to treatments such as balance

disorders. The conclusions and assessments at the individual level can help recommend each patient to continue, change, or adjust the treatment. The conclusions of the study at the individual level are useful in the clinic, facilitate improved decision-making, and show the benefits of the results.

Regarding the limitations of the study, we consider that the proposed 'logical' guidelines can serve as the basis for future studies with patients for the evaluation of treatments; however, it is necessary to deepen the development of rules that allow more efficient qualification of the change detected in the variables, providing the method with greater intelligence for the evaluation of treatments. Likewise, a simple scoring system could be developed (where positive values mean improvement and negative values mean patient deterioration), adding it this to the information already provided by the proposed graphics. On the other hand, we consider that the comparison between the results of the method and the specialist physician assessment should be deepened, establishing a more detailed protocol to more efficiently contrast both evaluations.

Likewise, we assume that the sample of young and healthy subjects could be widened in order to cover other specific age ranges, although we consider it to be adequate for the purpose of this study. On the other hand, future research should include a greater number of patients with different origins of the balance disorder, in order to assess the method's effectiveness with other kind of pathologies. Likewise, future research should focus on including the MID in conjunction with the MDC; in this way, by determining the MID thresholds in combination with the MDC, research would provide a more effective method with which to assess treatments and support diagnosis.

5. Conclusions

A test-retest of the balance test was conducted on a sample of healthy young subjects, identifying variables that are more reproducible with a lower MDC value: the COP mean speed in the RSEO, RSEC, SSEO tasks, the LOS, RMS, and area in RSEC and LOS, and all the COP limits in the LOS. Regarding the assessment of the eight patients with balance disorders characterized by vertigo of vestibular origin, the evaluations reported by the specialist physician mostly coincide with the objective pre- and post-test results. In this way, the applicability of the results to the assessed patient was shown because certain variables were sufficiently sensitive to detect important changes linked to a progression (improvement/worsening/no changes) in balance. The results of this study aim is to improve medical decision-making and the individualized follow-up of patients.

Supplementary Materials: The following is available online at http://www.mdpi.com/2227-9032/8/4/402/s1, S1: Supplementary Material. Patient Level Study.

Author Contributions: Conceptualization: J.D.l.T., J.M., M.P. and J.J.M.; data curation: J.D.l.T., J.M., M.P. and J.J.M.; formal analysis: J.D.l.T., J.M., M.P. and J.J.M.; funding acquisition: J.J.M.; investigation: J.D.l.T. and M.P.; methodology: J.D.l.T., J.M. and J.J.M.; project administration: J.J.M.; resources: M.P. and J.J.M.; software: J.D.l.T., J.M. and J.J.M.; supervision: J.J.M.; validation: J.D.l.T.; visualization: J.D.l.T.; writing—Original Draft Preparation: J.D.l.T., J.M. and J.J.M.; writing—Review & Editing: J.D.l.T., J.M., M.P. and J.J.M. All authors have read and agreed to the published version of the manuscript.

Funding: This project was co-funded by the Government of Aragon, the European Regional Development Fund and the University of Zaragoza.

Acknowledgments: We thank the I3A—University Institute of Research of Engineering of Aragon, University of Zaragoza, Zaragoza, Spain, for the materials they provided and the time in their Biomechanics Laboratory where the tests were carried out. We also thank Alcañiz Hospital (Teruel, Spain) for allowing us to use their facilities.

Conflicts of Interest: The authors declare that they have not conflicts of interest.

References

1. Penger, M.; Strobl, R.; Grill, E. Country-specific and individual determinants of dizziness in Europe: Results from the survey of health ageing and retirement in Europe (SHARE). *Public Health* **2017**, *149*, 1–10. [CrossRef] [PubMed]
2. Gavrilov, L.A.; Heuveline, P. Aging of population. *Encycl. Popul.* **2003**, *1*, 32–37.

3. Vaupel, J.W.; Loichinger, E. Redistributing work in aging Europe. *Science* **2006**, *312*, 1911–1913. [CrossRef] [PubMed]
4. Lin, D.; Seol, H.; Nussbaum, M.A.; Madigan, M.L. Reliability of COP-based postural sway measures and age-related differences. *Gait Posture* **2008**, *28*, 337–342. [CrossRef]
5. Baydal-Bertomeu, J.M.; Barberà I Guillem, R.; Soler-Gracia, C.; Peydro De Moya, M.F.; Prat, J.M.; Barona De Guzmán, R. Determinación de los patrones de comportamiento postural en población sana Española. *Acta Otorrinolaringológica Española* **2004**, *55*, 260–269. [CrossRef]
6. Tinetti, M.E. Preventing falls in elderly persons. *N. Engl. J. Med.* **2003**, *348*, 42–49. [CrossRef]
7. Ma, C.Z.; Wong, D.W.; Lam, W.K.; Wan, A.H.; Lee, W.C. Balance improvement effects of biofeedback systems with state-of-the-art wearable sensors: A systematic review. *Sensors* **2016**, *16*, 434. [CrossRef]
8. Salehi, R.; Ebrahimi-Takamjani, I.; Esteki, A.; Maroufi, N.; Parnianpour, M. Test-retest reliability and minimal detectable change for center of pressure measures of postural stability in elderly subjects. *Med. J. Islamic Repub. Iran* **2010**, *23*, 224–232.
9. Wolf, S.L.; Barnhart, H.X.; Kutner, N.G.; McNeely, E.; Coogler, C.; Xu, T.; Atlanta FICSIT Group. Reducing frailty and falls in older persons: An investigation of Tai Chi and computerized balance training. *J. Am. Geriatr. Soc.* **1996**, *44*, 489–497. [CrossRef]
10. Sánchez-Sánchez, M.L.; Belda-Lois, J.; Mena-del Horno, S.; Viosca-Herrero, E.; Igual-Camacho, C.; Gisbert-Morant, B. A new methodology based on functional principal component analysis to study postural stability post-stroke. *Clin. Biomech.* **2018**, *56*, 18–26. [CrossRef]
11. Patrícia Paludette, D.; Fabrício Santana da, S.; Carlos Bolli, M. Comparação do equilíbrio postural entre grupos de mulheres com diferentes faixas etárias/Comparison of postural balance among groups of women with different age ranges. *Fisioterapia e Pesquisa* **2015**, *22*, 392.
12. Baker, R. Gait analysis methods in rehabilitation. *J. Neuroeng. Rehabil.* **2006**, *3*, 4. [CrossRef] [PubMed]
13. López, L.F.C.; Calidonio, J.D.L. "CgMed" Design and construction of a platform to determine the position of the center of gravity in standing. *Revista Ingeniería Biomédica* **2009**, *3*, 26–36.
14. Postolache, O.A.; Postolache, G.B. Development and selection of balance sensing devices. *IEEE Instrum. Meas. Mag.* **2017**, *20*, 38–48.
15. Ito, T.; Sakai, Y.; Ito, Y.; Yamazaki, K.; Morita, Y. In Association between back muscle strength and proprioception or mechanoreceptor control strategy in postural balance in elderly adults with lumbar spondylosis. *Healthcare* **2020**, *8*, 58. [CrossRef] [PubMed]
16. Choi, H.; Lee, J. In immediate effect of balance taping using kinesiology tape on dynamic and static balance after ankle muscle fatigue. *Healthcare* **2020**, *8*, 162. [CrossRef] [PubMed]
17. Ku, P.X.; Abu Osman, N.A.; Wan Abas, W.A.B. The limits of stability and muscle activity in middle-aged adults during static and dynamic stance. *J. Biomech.* **2016**, *49*, 3943–3948. [CrossRef] [PubMed]
18. De la Torre, J.; Marin, J.; Marin, J.J.; Auria, J.M.; Sanchez-Valverde, M.B. Balance study in asymptomatic subjects: Determination of significant variables and reference patterns to improve clinical application. *J. Biomech.* **2017**, *65*, 161–168. [CrossRef] [PubMed]
19. Meshkati, Z.; Namazizadeh, M.; Salavati, M.; Mazaheri, M. Reliability of force-platform measures of postural sway and expertise-related differences. *J. Sport Rehab.* **2011**, *20*, 442–456. [CrossRef]
20. García, R.B.; Corresa, S.P.; Bertomeu, J.M.B.; Suárez-Varela, M.M.M. Posturografía estática con pruebas dinámicas. Utilidad de los parámetros biomecánicos en la valoración del paciente vestibular. *Acta Otorrinolaringológica Española* **2012**, *63*, 332–338. [CrossRef]
21. Swanenburg, J.; de Bruin, E.D.; Favero, K.; Uebelhart, D.; Mulder, T. The reliability of postural balance measures in single and dual tasking in elderly fallers and non-fallers. *BMC Musculoskelet. Disord.* **2008**, *9*, 162. [CrossRef] [PubMed]
22. Bauer, C.; Gröger, I.; Rupprecht, R.; Gaßmann, K.G. Intrasession Reliability of Force Platform Parameters in Community-Dwelling Older Adults. *Arch. Phys. Med. Rehabil.* **2008**, *89*, 1977–1982. [CrossRef] [PubMed]
23. De Sá Ferreira, A.; Baracat, P.J.F. Test–retest reliability for assessment of postural stability using center of pressure spatial patterns of three-dimensional statokinesigrams in young health participants. *J. Biomech.* **2014**, *47*, 2919–2924. [CrossRef] [PubMed]
24. Ruhe, A.; Fejer, R.; Walker, B. The test–retest reliability of centre of pressure measures in bipedal static task conditions—A systematic review of the literature. *Gait Posture* **2010**, *32*, 436–445. [CrossRef]

25. Geldhof, E.; Cardon, G.; De Bourdeaudhuij, I.; Danneels, L.; Coorevits, P.; Vanderstraeten, G.; De Clercq, D. Static and dynamic standing balance: Test-retest reliability and reference values in 9 to 10 year old children. *Eur. J. Pediatr.* **2006**, *165*, 779–786. [CrossRef]
26. Lafond, D.; Corriveau, H.; Hébert, R.; Prince, F. Intrasession reliability of center of pressure measures of postural steadiness in healthy elderly people. *Arch. Phys. Med. Rehabil.* **2004**, *85*, 896–901. [CrossRef]
27. Lee, P.; Liu, C.; Fan, C.; Lu, C.; Lu, W.; Hsieh, C. The test–retest reliability and the minimal detectable change of the Purdue Pegboard Test in schizophrenia. *J. Formosan Med. Assoc.* **2013**, *112*, 332–337. [CrossRef]
28. Pagnacco, G.; Carrick, F.R.; Wright, C.H.G.; Oggero, E. Between-subjects differences of within-subject variability in repeated balance measures: Consequences on the minimum detectable change. *Gait Posture* **2015**, *41*, 136–140. [CrossRef]
29. Pinsault, N.; Vuillerme, N. Test–retest reliability of centre of foot pressure measures to assess postural control during unperturbed stance. *Med. Eng. Phys.* **2009**, *31*, 276–286. [CrossRef]
30. 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]
31. Lee, H.; Granata, K.P. Process stationarity and reliability of trunk postural stability. *Clin. Biomech.* **2008**, *23*, 735–742. [CrossRef] [PubMed]
32. De Oliveira Silva, D.; Briani, R.V.; Pazzinatto, M.F.; Ferrari, D.; Aragão, F.A.; de Albuquerque, C.E.; Alves, N.; de Azevedo, F.M. Reliability and differentiation capability of dynamic and static kinematic measurements of rearfoot eversion in patellofemoral pain. *Clin. Biomech.* **2015**, *30*, 144–148. [CrossRef]
33. Pawar, P.K.; Dadhich, A. Study of correlation between human height and hand length in residents of Mumbai. *Int. J. Biol. Med. Res.* **2012**, *3*, 2072–2075.
34. Bujang, M.A.; Baharum, N. A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: A review. *Arch. Orofac. Sci.* **2017**, *12*, 1–11.
35. Charan, J.; Biswas, T. How to calculate sample size for different study designs in medical research? *Indian J. Psychol. Med.* **2013**, *35*, 121–126. [CrossRef] [PubMed]
36. Sample, R.B.; Jackson, K.; Kinney, A.L.; Diestelkamp, W.S.; Reinert, S.S.; Bigelow, K.E. Manual and cognitive dual tasks contribute to fall-risk differentiation in posturography measures. *J. Appl. Biomech.* **2016**, *32*, 541–547. [CrossRef] [PubMed]
37. Scoppa, F.; Capra, R.; Gallamini, M.; Shiffer, R. Clinical stabilometry standardization: Basic definitions–acquisition interval–sampling frequency. *Gait Posture* **2013**, *37*, 290–292. [CrossRef]
38. Huurnink, A.; Fransz, D.P.; Kingma, I.; van Dieën, J.H. Comparison of a laboratory grade force platform with a Nintendo Wii Balance Board on measurement of postural control in single-leg stance balance tasks. *J. Biomech.* **2013**, *46*, 1392–1395. [CrossRef]
39. Duarte, M.; Freitas, S.M. Revision of posturography based on force plate for balance evaluation. *Braz. J. Phys. Ther.* **2010**, *14*, 183–192. [CrossRef]
40. Bonnechere, B.; Jansen, B.; Jan, S.V.S. Cost-effective (gaming) motion and balance devices for functional assessment: Need or hype? *J. Biomech.* **2016**, *49*, 2561–2565. [CrossRef]
41. Zhu, Y. Design and validation of a low-cost portable device to quantify postural stability. *Sensors* **2017**, *17*, 619. [CrossRef] [PubMed]
42. Uebbing, T.J. User Experience in Smart Environments: Design and Prototyping. Master's Thesis, University of Twente, Twente, The Netherlands, 2016.
43. Williams, Q.I.; Gunn, A.H.; Beaulieu, J.E.; Benas, B.C.; Buley, B.; Callahan, L.F.; Cantrell, J.; Genova, A.P.; Golightly, Y.M.; Goode, A.P.; et al. Physical therapy vs. internet-based exercise training (PATH-IN) for patients with knee osteoarthritis: Study protocol of a randomized controlled trial. *BMC Musculoskelet. Disord.* **2015**, *16*, 264. [CrossRef] [PubMed]
44. Peydro de Moya, M.F.; Baydal Bertomeu, J.M.; Vivas Broseta, M.J. Evaluación y rehabilitación del equilibrio mediante posturografía. *Rehabilitación* **2005**, *39*, 315–323. [CrossRef]
45. Tesio, L.; Rota, V.; Longo, S.; Grzeda, M.T. Measuring standing balance in adults: Reliability and minimal real difference of 14 instrumental measures. *Int. J. Rehabil. Res.* **2013**, *36*, 362–374. [CrossRef]
46. Hoving, J.L.; Pool, J.J.; van Mameren, H.; Devillé, W.J.; Assendelft, W.J.; de Vet, H.C.; de Winter, A.F.; Koes, B.W.; Bouter, L.M. Reproducibility of cervical range of motion in patients with neck pain. *BMC Musculoskelet. Disord.* **2005**, *6*, 59. [CrossRef]

47. Benvenuti, F.; Mecacci, R.; Gineprari, I.; Bandinelli, S.; Benvenuti, E.; Ferrucci, L.; Baroni, A.; Rabuffetti, M.; Hallett, M.; Dambrosia, J.M. Kinematic characteristics of standing disequilibrium: Reliability and validity of a posturographic protocol. *Arch. Phys. Med. Rehabil.* **1999**, *80*, 278–287. [CrossRef]
48. Doyle, T.L.; Newton, R.U.; Burnett, A.F. Reliability of traditional and fractal dimension measures of quiet stance center of pressure in young, healthy people. *Arch. Phys. Med. Rehabil.* **2005**, *86*, 2034–2040. [CrossRef]
49. Schuck, P.; Zwingmann, C. The 'smallest real difference's a measure of sensitivity to change: A critical analysis. *Int. J. Rehabil. Res.* **2003**, *26*, 85–91.
50. Tao, W.; Liu, T.; Zheng, R.; Feng, H. Gait analysis using wearable sensors. *Sensors* **2012**, *12*, 2255–2283. [CrossRef]
51. Donoghue, D.; Stokes, E.K. How much change is true change? The minimum detectable change of the Berg Balance Scale in elderly people. *J. Rehabil. Med.* **2009**, *41*, 343–346. [CrossRef]
52. Kovacs, F.M.; Abraira, V.; Royuela, A.; Corcoll, J.; Alegre, L.; Tomás, M.; Mir, M.A.; Cano, A.; Muriel, A.; Zamora, J. Minimum detectable and minimal clinically important changes for pain in patients with nonspecific neck pain. *BMC Musculoskelet. Disord.* **2008**, *9*, 43. [CrossRef] [PubMed]
53. Marchetti, G.F.; Lin, C.C.; Alghadir, A.; Whitney, S.L. Responsiveness and minimal detectable change of the dynamic gait index and functional gait index in persons with balance and vestibular disorders. *J. Neurol. Phys. Ther.* **2014**, *38*, 119–124. [CrossRef] [PubMed]
54. Batterham, A.M.; Hopkins, W.G. Making meaningful inferences about magnitudes. *Int. J. Sports Physiol. Perform.* **2006**, *1*, 50–57. [CrossRef]
55. Cicchetti, D.V. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychol. Assess.* **1994**, *6*, 284. [CrossRef]
56. Hickey, S.; Ford, G.; Buckley, J.; O'connor, A.F. Unterberger stepping test: A useful indicator of peripheral vestibular dysfunction? *J. Laryngol. Otol.* **1990**, *104*, 599–602. [CrossRef]
57. Bartual, J.; Pérez, N. El sistema vestibular y sus alteraciones. *Tomo* **1998**, *1*, 21–22.
58. Shumway-Cook, A.; Brauer, S.; Woollacott, M. Predicting the probability for falls in community-dwelling older adults using the Timed Up & Go Test. *Phys. Ther.* **2000**, *80*, 896–903.
59. Martínez Carrasco, Á. Análisis del riesgo de caídas en ancianos institucionalizados mediante escalas de marcha y equilibrio. Ph.D. Thesis, Universidad de Murcia, Murcia, Spain, 28 January 2016.
60. Vellas, B.J.; Wayne, S.J.; Romero, L.; Baumgartner, R.N.; Rubenstein, L.Z.; Garry, P.J. One-leg balance is an important predictor of injurious falls in older persons. *J. Am. Geriatr. Soc.* **1997**, *45*, 735–738. [CrossRef]
61. Hansson, E.E.; Persson, L.; Malmström, E.M. Influence of vestibular rehabilitation on neck pain and cervical range of motion among patients with whiplash-associated disorder: A randomized controlled trial. *J. Rehabil. Med.* **2013**, *45*, 906–910. [CrossRef]
62. Orejas, J.I.B.; Varea, J.A.; Rodrigo, J.V.; Navas, A.C. Resultados y seguimiento de la rehabilitación vestibular. *Revista ORL* **2020**, *11*, 107–114. [CrossRef]
63. Hopkins, W.G. A spreadsheet for monitoring an individual's changes and trend. *Sportscience* **2017**, *21*, 10.
64. Hopkins, W.G. Rebranding MBI as magnitude-based decisions (MBD). *Sportscience* **2019**, *i-iii*, 23.
65. Hopkins, W.G.; Marshall, S.W.; Batterham, A.M.; Hanin, J. Progressive statistics for studies in sports medicine and exercise science. *Med. Sci. Sports Exerc.* **2009**, *41*, 3–13. [CrossRef] [PubMed]
66. Salavati, M.; Hadian, M.R.; Mazaheri, M.; Negahban, H.; Ebrahimi, I.; Talebian, S.; Jafari, A.H.; Sanjari, M.A.; Sohani, S.M.; Parnianpour, M. Test–retest reliabty of center of pressure measures of postural stability during quiet standing in a group with musculoskeletal disorders consisting of low back pain, anterior cruciate ligament injury and functional ankle instability. *Gait Posture* **2009**, *29*, 460–464. [CrossRef] [PubMed]
67. Degani, A.M.; Leonard, C.T.; Danna-dos-Santos, A. The effects of early stages of aging on postural sway: A multiple domain balance assessment using a force platform. *J. Biomech.* **2017**, *64*, 8–15. [CrossRef]
68. Hanes, D.A.; McCollum, G. Cognitive-vestibular interactions: A review of patient difficulties and possible mechanisms. *J. Vestib. Res.* **2006**, *16*, 75–91.
69. Swanenburg, J.; de Bruin, E.D.; Stauffacher, M.; Mulder, T.; Uebelhart, D. Effects of exercise and nutrition on postural balance and risk of falling in elderly people with decreased bone mineral density: Randomized controlled trial pilot study. *Clin. Rehabil.* **2007**, *21*, 523–534. [CrossRef]
70. Tsukamoto, H.F.; Costa, V.d.S.P.; Silva Junior, R.A.d.; Pelosi, G.; Marchiori, L.L.d.M.; Vaz, C.R.S.; Fernandes, K.B.P. Effectiveness of a vestibular rehabilitation protocol to improve the health-related quality of life and postural balance in patients with vertigo. *Int. Arch. Otorhinolaryngol.* **2015**, *19*, 238–247. [CrossRef]

71. Baloh, R.W.; Jacobson, K.M.; Enrietto, J.A.; Corona, S.; Honrubia, V. Balance disorders in older persons: Quantification with posturography. *Otolaryngol. Head Neck Surg.* **1998**, *119*, 89–92. [CrossRef]
72. Kim, S.K.; Kim, Y.B.; Park, I.S.; Hong, S.J.; Kim, H.; Hong, S.M. Clinical analysis of dizzy patients with high levels of depression and anxiety. *J. Audiol. Otol.* **2016**, *20*, 174–178. [CrossRef]
73. Yuan, Q.; Yu, L.; Shi, D.; Ke, X.; Zhang, H. Anxiety and depression among patients with different types of vestibular peripheral vertigo. *Medicine* **2015**, *94*. [CrossRef] [PubMed]
74. Yardley, L.; Redfern, M.S. Psychological factors influencing recovery from balance disorders. *J. Anxiety Disord.* **2001**, *15*, 107–119. [CrossRef]
75. De Vet, H.C.; Terwee, C.B. The minimal detectable change should not replace the minimal important difference. *J. Clin. Epidemiol.* **2010**, *63*, 804. [CrossRef] [PubMed]
76. Hamburg, M.A.; Collins, F.S. The path to personalized medicine. *N. Engl. J. Med.* **2010**, *363*, 301–304. [CrossRef] [PubMed]

4.3. ARTÍCULO 3

De la Torre J, Marin J, Polo M, Gómez-Trullén EM, Marin JJ. **2021. MCQ-Balance: a method to monitor patients with balance disorders and improve clinical interpretation of posturography.** PeerJ 9:e10916 https://doi.org/10.7717/peerj.10916

PEERJ

ISSN: 2167-8359

Factor de Impacto JCR (2019):

2.379

Áreas Temáticas:

Multidisciplinary Sciences: Q2 (32/71)

MCQ-Balance: a method to monitor patients with balance disorders and improve clinical interpretation of posturography

Juan De la Torre[1,2], Javier Marin[1,3], Marco Polo[4], Eva M. Gómez-Trullén[1,5] and Jose J. Marin[1,3]

[1] IDERGO (Research and Development in Ergonomics) Research Group, I3A (Aragon Institute of Engineering Research), Zaragoza, Spain
[2] Department of Biomedical Engineering, Universidad de Zaragoza, Zaragoza, Spain
[3] Department of Design and Manufacturing Engineering, Universidad de Zaragoza, Zaragoza, Spain
[4] MD Physical Medicine and Rehabilitation, Hospital of Alcañiz, Teruel, Spain
[5] Department of Physical Medicine and Rehabilitation and Nursing, Health Sciences Faculty, Universidad de Zaragoza, Zaragoza, Spain

ABSTRACT

Background. An estimated 20% to 30% of the global population has suffered a vertiginous episode. Among them, 20% do not receive a clear diagnosis. Improved methods, indicators and metrics are necessary to assess the sensory systems related to balance, especially when patients are undergoing treatment for vertiginous episodes. Patients with balance disorders should be monitored for changes at the individual level to gather objective information. In this study, we evaluate the use of the MCQ-Balance (Measure, Classify and Qualify) assessment for examining a patient's balance progression using tests to measure static balance control and dynamic postural balance with a stabilometric platform.

Materials and Methods. The MCQ-Balance assessment comprises three stages: (i) measuring the progression of each variable between two separate and consecutive days (called sessions) using the Magnitude-Based Decision analysis; (ii) classifying the progression of the patient's balance with a score; and (iii) qualifying the progression of the patient's balance from the resulting scores using a set of rules. This method was applied to 42 patients with balance disorders of peripheral or central origin characterised by vertigo as the cardinal symptom. Balance progression was measured using the MCQ-Balance assessment over the course of three months, and these results were compared with the assessment of a clinical expert.

Results. The MCQ-Balance assessment showed an accuracy of 83.4% and a Cohen's Kappa coefficient of 0.752 compared to the assessment of a clinical expert.

Conclusion. The MCQ-Balance assessment facilitates the monitoring of patient balance and provides objective information that has the potential to improve medical decision making and the adjustment of individual treatment.

Submitted 28 July 2020
Accepted 18 January 2021
Published 23 February 2021

Corresponding author
Juan De la Torre, 627471@unizar.es

Academic editor
Joao Rocha

Additional Information and Declarations can be found on page 21

DOI **10.7717/peerj.10916**

Subjects Anatomy and Physiology, Evidence Based Medicine, Kinesiology, Neurology, Statistics

Keywords Balance sensory systems, Objective information, Patient-level analysis, Stabilometric platform, Monitoring, Vertigo, Personalized medicine, Methodology

INTRODUCTION

Vertigo is an illusion of movement, either of the external world revolving around the individual or of the individual revolving in space (*Medical Subject Headings, 2020*). It is the cardinal symptom of balance, which leads to a significant reduction in the quality of life and an increase in disability, anxiety and depression (*Neuhauser, 2016*). There is a high prevalence of balance disorders among elderly people in developed countries (*Penger, Strobl & Grill, 2017*; *Rubenstein & Josephson, 2002*). In combination with a gradual increase in the ageing index of the population (*Vaupel & Loichinger, 2006*; *Muir et al., 2013*), it has resulted in an increase in the risk of falls of elderly people (*Aftab, Robert & Wieber, 2016*; *Khalaj et al., 2014*). Globally, approximately 20 to 30% of the population has a vertiginous episode of various origins and severity over a lifetime (*Da Costa Barbosa & Vieira, 2017*; *Lin et al., 2008*; *Wolf et al., 1996*; *Tinetti, 2003*). Moreover, 20% of them do not receive a clear diagnosis (*Swanenburg et al., 2008*).

Vertigo is most often caused by dysfunction resulting from a peripheral or central lesion (*Stanton, 2020*); therefore, depending on the origin, it can be classified as vertigo of peripheral or central origin (*Wipperman, 2014*; *Baumgartner, 2019*; *Strupp, Dieterich & Brandt, 2013*). Vertigo can also have a cervical origin in cases where the central nervous system is unaffected (except spinal cord syndrome), with a demonstrated relationship between cervicalgia and vertigo (*Dieterich & Eckhardt-Henn, 2004*; *Reiley et al., 2017*). Vertigo of cervical origin is identified most frequently in rehabilitation consultations , occurring most often in patients between the ages of 30 and 50 (*Jull, 2004*; *Solomon, 2000*).

To analyse the causes of vertigo, the degree of alteration must be measured in isolation or in combination of each balance sensory system (BSS), including the vestibular (VS), visual (ES; eye-sight), and proprioceptive (PS) systems. A vertiginous episode or trauma can affect these systems to a greater or lesser extent, and consequently, the patient's balance (*Hanes & McCollum, 2006*; *Shumway-Cook et al., 2001*). It is therefore necessary to have methods or indicators to determine how the BSS progresses and to standardize the initial evaluation of patients' balance and its progression, especially during the treatment of a balance disorder (*Patrícia Paludette, Fabrício Santana da & Carlos Bolli, 2015*).

When it is challenging to establish a clear pathology related to any of the BSSs, or when multiple origins of the condition are found, the clinical diagnosis becomes complicated (*Derebery, 2000*), and additional measures and tests are required to provide important information to the clinician (*Hickey et al., 1990*; *Martínez Carrasco, 2016*; *Vellas et al., 1997*).

As an alternative or complement to the functional tests, using a stabilometric platform, posturography allows movements of the centre of pressure (COP) in the standing position to be measured. Stabilometric platforms can assess static balance control and dynamic postural balance through different variables and application methods (*Ito et al., 2020*; *Choi & Lee, 2020*). It constitutes a functional assessment with medical-legal validity that provides objective information regarding balance disorders in clinical practice (*De la Torre et al., 2017*; *Dounskaia, Peterson & Bruhns, 2018*; *Lin et al., 2008*). Although posturography is a validated assessment, difficulties are encountered with regard to discerning the origin

caused by the imbalance pattern. This is because, although sensory analyses suggest a proprioceptive-visual-vestibular pattern, this is not always accurate (*El-Kashlan et al., 1998*; *Stewart et al., 1999*; *Timothy & Hain, 2019*). Related to the above, although the clinical results from traditional posturography are useful, they are insufficient in certain cases, requiring smarter devices (*Allum et al., 2002*; *Di Fabio, 1996*).

Posturography devices can provide information on patients' balance that is useful for clinical decision-making, as a functional assessment value, measuring data related to the patients' balance; however, in order for such devices to be practical, they must be easy for clinicians to use without consulting external experts (*Visser et al., 2008*).

Although several balance assessment tests have been applied through a stabilometric platform (*Karlsson & Frykberg, 2000*), their resulting scores are sometimes complex and difficult to interpret (*Peterson et al., 2003*). To this end, subjective scoring lacks standardization and can be difficult to interpret, thus making difficult classify the patient balance status, which resulting in difficulties diagnosing balance disorders and identifying the BSS from which the imbalance pattern originates (*Jacobs et al., 2006*; *Visser et al., 2008*; *Saxena & Prabhakar, 2013*).

Posturography reports should involve easily understandable, non-technical language, qualifying the patient's balance status in an understandable way for clinicians (*Von Lubitz & Wickramasinghe, 2006*; *Visser et al., 2008*). Likewise, both validation and standardization of the protocols for reproducibility and a possible comparison with similar studies are required (*Visser et al., 2008*).

In rehabilitation, it is critical to measure the progression of a patient's balance between two separate sessions in order to objectively characterize the response to treatments (*Hamburg & Collins, 2010*); this helps determine whether relevant changes have occurred in the patient at the individual level, thus informing future treatment decision making (*Visser et al., 2008*; *Hopkins, 2017*). Regarding this, we can highlight the proposal of (*Hopkins, 2017*) to assess the change between two measurements in an individual through the magnitude-based decision (MBD) method (*Hopkins, 2019*), which is used in this work.

The development of the MCQ-Balance assessment method was motivated by these issues, in relation to the necessity of providing objective, easily-interpretable information about patients' balance that specifies the origin of the pathology. Using a stabilometric platform, this method detects relevant changes between two consecutive balance tests (monitor) in patients with balance disorders, providing objective information about the origin of the imbalance. The MCQ-Balance assessment comprises three separate stages in which the progression of a patient's balance is measured, then classified, and finally qualified. In this study, the MCQ-Balance assessment was applied to balance disorder patients with vertigo as the cardinal symptom. Subsequently, the results obtained were compared with the evaluation of a specialist clinician.

MATERIALS & METHODS

Participants

A total of 42 patients with balance disorders characterised by vertigo (of peripheral or central origin) as the cardinal symptom were monitored via balance tests with a stabilometric platform.

The patients were referred by the Primary Care, Otorhinolaryngology and Neurology Services of the Alcañiz Hospital (Teruel, Spain) after being diagnosed with a balance disorder. The methods for determining the deficit, varied according to the service where the diagnosis was made: (i) in Primary Care, medical history was considered; (ii) in Otorhinolaryngology, in addition to the medical history, magnetic resonance imaging, videonystagmography, and tests such as the Dix-Hallpike manoeuvre were used; and (iii) in Neurology, in addition to medical history and magnetic resonance imaging, computerized axial tomography and neurophysiology tests, were used. Table 1 presents the main diagnoses of the patients with respect to peripheral or central deficits. Table 2 shows general and anthropometric characteristics of the patient population. The study included 27 females (64%) and 15 males (36%), and there were no statistically significant differences in any baseline characteristics according to sex (*Yin et al., 2009*).

From these diagnostic services, patients from the Teruel region (Spain) were referred to the Physical Medicine and Rehabilitation Service (PM&R) of the Alcañiz Hospital, having been identified with vertigo as the cardinal symptom. A doctor (clinician 1) of the PM&R service then evaluated the patients, considering (i) their medical history and physical examination, (ii) previous diagnosis and (iii) the results of the functional balance assessments carried out, such as the Unterberger test (*Bartual & Pérez, 1998*; *Hickey et al., 1990*), up and go test (*Martínez Carrasco, 2016*; *Shumway-Cook, Brauer & Woollacott, 2000*), and unipodal support test (*Vellas et al., 1997*).

The selected patients met the following inclusion criteria: (i) between 35 and 70 years old and (ii) having suffered a vertiginous episode of peripheral or central origin in the last year. The following were the exclusion criteria, which were based on (*Degani, Leonard & Danna-dos Santos, 2017*; *Hanes & McCollum, 2006*; *Swanenburg et al., 2007*; *Tsukamoto et al., 2015*) and were applied in other research (*De la Torre et al., 2020a*; *De la Torre et al., 2020b*): (i) presented acute osteomuscular pathology in the lower limbs or lumbar spine, which may alter the outcome of the stabilometric platform, (ii) presented any amputation in the lower limbs, or (iii) presented oncological pathology or was in active treatment with chemotherapy, radiotherapy, or hormonal therapy.

The present study was approved by the Research Ethics Committee of the Community of Aragon (CEICA) (January 16, 2019) and complied the ethical standards of the Declaration of Helsinki (*General Assembly of the World Medical Association, 2014*).

Prior to the start of the tests, the participants signed a consent form sheet that involved accepting the tests and understanding the purpose of them. The participants in this study has given written informed consent to publish these case details.

Table 1 Patients diagnosis according to their deficits.

Peripherical deficit $n = 32$ (76.2%)	Central deficit $n = 10$ (23.8%)
BPPB $n = 15$ (36%)	Ictus $n = 6$ (14%)
Ménière syndrome $n = 8$ (19%)	Neoplasia $n = 2$ (5%)
Vestibular hypofunction $n = 6$ (14%)	Demyelinating disease $n = 2$ (5%)
Otoesclerosis $n = 3$ (7%)	

Notes.
BPPV, benign paroxysmal peripheral vertigo.

Table 2 Participant anthropometric characteristics: mean (SD).

Characteristics	Patients ($n = 42$)
Gender (men/women)	15/27
Age (yr)	57.1 (8.7)
Height (cm)	162.8 (7.9)
Weight (kg)	76.5 (15.8)
Body Mass Index (kg/m^2)	28.7 (5.6)
Foot length (cm)[a]	25.3 (1.1)
Abdominal perimeter (cm)	97.1 (14.9)
Deficit (Peripherical / Central)	32/10

Notes.
[a]Foot length measurements were taken between the proximal and distal points on the foot outline (*Pawar & Dadhich, 2012*).

The choice of 42 patients was based on the following formula for sample size calculation involving qualitative variables (*Charan & Biswas, 2013*):

$$N = \frac{\left(Z_{1-\alpha/2}\right)^2 p(1-p)}{d^2} \quad (1)$$

Here, $Z_1 - \alpha/2$ is the standard normal variate with $p_v < 0.05$ (type I error); p is the expected proportion of the research goal in population; and d is the absolute error or precision, as determined by the researcher. We selected a $Z_1 - \alpha/2$ value of 1.96 (standard, given that py values are considered statistically significant below 0.05), a p-value of 0.7 (based on initial criteria that showed an accuracy of >70%, as shown in the statistical analysis section), and a d-value of 0.15. The resulting N-value was 39.85.

Instrumentation

The device used was the stabilometric platform MoveHuman-Dyna UZ, which was designed and manufactured by the IDERGO (Research and Development in Ergonomics, University of Zaragoza, Spain) research group (see Fig. 1). It is a static posturography device designed for research, which comprises four load cells and a lightweight aluminium structure, whose dimensions and characteristics are detailed in the study of *De la Torre et al. (2017)*. The findings of this device can be replicated in a straightforward manner by other researchers, which enhances the applicability of this study. The acquisition and processing of the

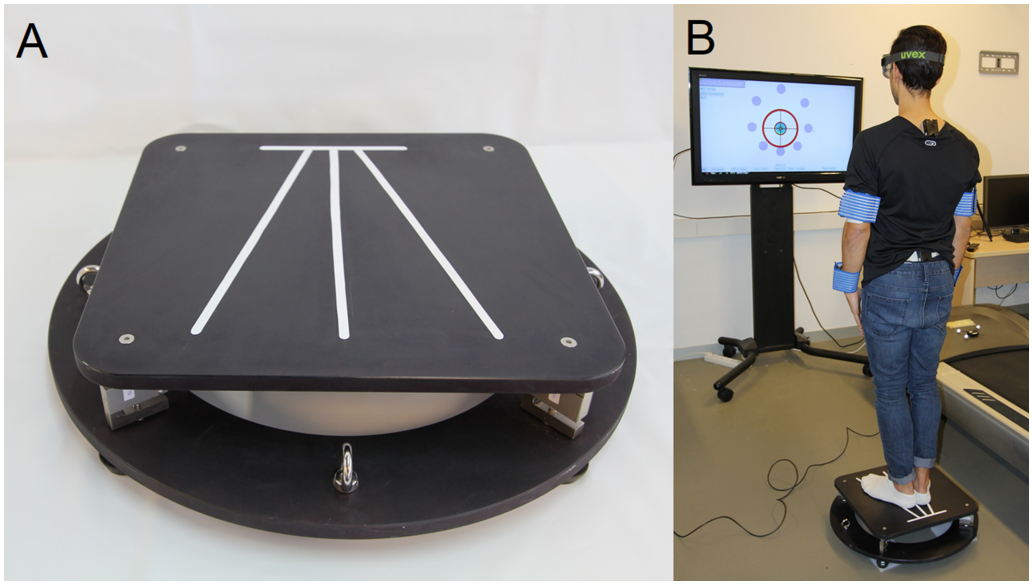


Figure 1 **(A) Stabilometric platform and (B) test work environment.**
Full-size DOI: 10.7717/peerj.10916/fig-1

platform data, as well as the format and method of exporting them, have been carried out according to the procedure used by *De la Torre et al. (2017)*.

Likewise, in accordance with the aforementioned study, the stabilometric platform 'meets the standards established by the International Society for Posture and Gait Research (ISPGR) for its clinical application' (*Scoppa et al., 2013*) in relation to various parameters, such as accuracy, precision, linearity, dimensions, resolution, sampling, and so on. The precision parameters (accuracy, precision, linearity, dimensions and resolution) were obtained through a reliability experiment in which the metrological characteristics of the platform were tested with a gold standard force platform, as well as the error of measurement *De la Torre et al. (2017)*. Processing the force data in function of the cells' position means we can calculate the real-time position of the trajectory that describes the position of the COP by applying the appropriate formula (*López & Calidonio, 2009*; *Ma et al., 2016*). The stabilometric platform has been used in several research projects with patients in different hospitals since 2018, both public (hospital Miguel Servet and university hospital Lozano Blesa (Zaragoza, Spain)) and private (hospital MAZ (Zaragoza, Spain)); all the research projects have been approved by the CEICA Committee. In addition, the characteristics of the platform and its portability make it suitable for clinical use where, for example, the medical office space is limited (*Scoppa et al., 2013*; *De la Torre et al., 2017*).

Protocol

Patients were evaluated by clinician 1 on two different days (sessions) spaced three months apart (first session: pre-session; second session: post-session). After the pre-session, clinician 1 prescribed the rehabilitation treatment according to the specific balance disorder of each patient. Patients with vertigo of peripheral or central origin performed vestibular rehabilitation exercises (*Boomsaad, Telian & Patil, 2017*). For patients with a specific

diagnosis of benign paroxysmal peripheral vertigo (BPPV), the Epley manoeuvre was performed in addition to vestibular rehabilitation exercises (*Orejas et al., 2020*; *Hansson, Persson & Malmström, 2013*).

After the evaluation by clinician 1, in each session (pre and post), the patients conducted a set of balance evaluation tests with a stabilometric platform (three months apart between the pre- and post-session). The tests were performed by the PM&R of the Alcañiz Hospital between February and July in 2019. The fieldwork was performed by a team of a clinician (clinician 2), a nurse, and a technician in the same hospital.

The static and dynamic balance were both assessed with a set of tests previously applied in other studies (*De la Torre et al., 2020b*; *De la Torre et al., 2017*).

Static balance control was assessed with a test based on the Romberg test and the Modified Clinical Test of Sensory Interaction in Balance (CTSIB-M). In the test patients must maintain their COP within the support base throughout the assessment period - 40 s (*De la Torre et al., 2017*). Static balance control was assessed in four different conditions, examined consecutively: (1) rigid surface with eyes open (RSEO), (2) rigid surface with eyes closed (RSEC), (3) soft surface with eyes open (SSEO), and (4) soft surface with eyes closed (SSEC).

On the other hand, the dynamic postural balance, which is vital for motor control, was assessed measuring the limits of stability (LOS) that a patient is able to reach and with it, the management capacity of COP (*Ku, Abu Osman & Wan Abas, 2016*). The inclusion of the LOS, complementary to the assessment of the static balance control, provides additional value to the balance assessment protocol (*Lin et al., 2008*; *Tesio et al., 2013*; *Salehi et al., 2010*).

The specific protocol applied in the tests: the position of the body, arms and feet during the test (*De la Torre et al., 2017*), environmental conditions (e.g., noise, space, etc.) and the additional instrumentation used as a foam rubber for soft surface or instruments for anthropometric data collection, is the same that (*De la Torre et al., 2020a*; *De la Torre et al., 2020b*) used for this stabilometric platform (Fig. 1). This protocol fulfils certain clinical conditions (*Swanenburg et al., 2008*; *Hoving et al., 2005*; *Benvenuti et al., 1999*; *Doyle, Newton & Burnett, 2005*; it must be fast and should not require multiple repetitions to issue a definite, consistent result (*Swanenburg et al., 2008*).

At the start of the tests, clinician 2 provided patients with instructions on how to perform the tests, according to similar studies (*De la Torre et al., 2020a*; *De la Torre et al., 2020b*). The patients were instructed on how to place their feet on the platform according to the mark placed in the stabilometric platform shown in Fig. 1A. The patients had to be in a standing position, with the arms extended and close to the body. The patients were instructed to keep as still as possible in the static tests; in the LOS, they were instructed to, using only the movement of the ankles without lifting the feet, follow a moving target LOS as explained by *De la Torre et al. (2020a)*. Clinician 2 also provided instructions on how to stimulate abdominal toning, since this has an influence on stability and balance (*Ayllón & Fernández, 2006*). Patients completed a practice run of each test so that clinician 2 could verify that they understood the procedure, assumed the correct posture, and executed the tests correctly. This also gave the patients the opportunity to get used to the platform and

environment, which are considered relevant factors in some balance studies (*Taylor et al., 2015*).

Variables

The variables selected for the present study were those determined by *De la Torre et al. (2020a)*. to be more significant in balance assessment studies, which details, and method of obtaining are also explained in the same study. The variables selected for the assessment of the static and dynamic balance were the range of displacement in the anteroposterior and mediolateral directions in mm, area in cm^2 (surface area covered by the trajectory of the COP), average speed of the COP in mm/s, and RMS position in mm. Additionally, in the LOS test, two more variables were assessed: the COP limits in mm (maximum displacement reached along each axis of the octagon radii), and the "success" variable in percentage (quantification of the management and coordination of the COP along each axis of the octagon radii), both defined in previous studies (*De la Torre et al., 2020a*; *De la Torre et al., 2020b*).

MCQ-Balance assessment method

Figure 2 presents the application outline of the MCQ-Balance assessment, which consists of three stages in which the progression of a patient's balance is Measured (M), Classified (C), and Qualified (Q). The method input is the variables provided by the set of balance tests in two temporal points, that is, the values of the variables in the pre-session and post-session. The variables are analysed individually until stage two, where they are grouped at the test level until the end of the assessment. The application outline shows the inputs and outputs of each stage, as well as the processes (P1-P5) applied to them. It also includes the type of information that is handled and the interpretative changes during the process.

Stage 1: Measure

The first stage of the method involves measuring the progression of each variable of the balance tests set by detecting relevant changes between two measures of each variable recorded at different temporal points (e.g., a measure of 26.4 for one session and 27.2 for another session). For this purpose, the process (P1) used in this stage is the statistical method MBD, as described in the Spreadsheet for Monitoring an Individual's Changes (*Hopkins, 2017*) (formerly known as magnitude-based inferences) (*Hopkins, 2019*). According to the MBD method, some inputs are required for each analysed variable:

- Xdif: difference between the measures taken in two temporal points: pre-value (pre-session) and post-value (post-session) (Eq. (1)).

$$Xdif = X_{post} - X_{pre} \tag{2}$$

- MBD threshold: for this method, a threshold (numerical value) must be defined from which a change is considered relevant. In our case, we selected the minimal detectable change (MDC) (Eq. (2)). The implications of this election are explained in the discussion section.

$$MDC = 1.96\sqrt{2}SEM; SEM = SDpool\sqrt{1 - ICC} \tag{3}$$

Where the standard deviation (SDpool) is the pooled average between the standard deviation of the test and retest, ICC is the intraclass correlation coefficient (specifically, the calculated coefficient was ICC3, k (similar to ICC2.1) (*Ruhe, Fejer & Walker, 2010*); the statistical software used for the ICC calculations was the IBM SPSS statistics (*IBM Corp, 2017*) and the ICC results were classified according to *Cicchetti (1994)*, who provided the following intervals to characterize the ICC inter-rater agreement measures; and SEM is the standard error of measurement. Following the exposed calculation procedure, ICC, SEM and MDC values were obtained in a previous test-retest study (*De la Torre et al., 2020a*; *De la Torre et al., 2020b*).

- Short-term typical error (STTE): this represents the error/deviation in the subject's repeated measurements in a short period for a sample of measurements instead of just one measurement per session, without any substantial change between them (as an intervention, for a long time between measurements, etc.) As proposed by *Hopkins (2000)* and *Hopkins (2017)*, this input was obtained with a previous short-term reliability study of the balance test set; similar study to the calculation of variables for the MDC (*De la Torre et al., 2020a*; *De la Torre et al., 2020b*).

To detect whether the change is relevant between two recorded measures, clinical MBD is followed (*Hopkins & Batterham, 2016*). This allows us to determine whether the detected progression is positive (beneficial), negative (harmful) or inconclusive.

First, with the value and sign (positive or negative) of Xdif, we determine the tendency of the change towards a positive or negative progression. In the MCQ-Balance assessment method, we follow the following criteria based on (*De la Torre et al., 2020a*; *De la Torre et al., 2020b*): for the static balance group, a positive progression is considered if Xdif has a negative sign, and for the dynamic balance group, a positive progression is considered if Xdif has a positive sign.

Subsequently, following the calculation method set forth by *Hopkins (2017)*, the probability of change (PoC in %) is obtained, which can be defined as the probability that the difference between the two values is relevant. This probability corresponds to the percentage of the confidence interval of the difference (calculated using the Xdif and STTE) that is outside of the range (+MDC, -MDC).

Once the PoC is calculated in the method, criteria must be established to consider a positive, negative, or null (unclear) progression of each variable. In a case study following the clinical MBD, a positive PoC that is greater than or equal to 25% corresponds to a relevant positive change, whereas a negative PoC that is greater than or equal to 5% corresponds to a relevant negative change in the patient. In contrast, if the positive PoC is less than 25% or the negative PoC is less than 5%, the change is considered 'unclear'. The asymmetry between the two intervals is because, in 'Clinical MBD the effects have an unacceptable risk of harm' (*Hopkins & Batterham, 2016*).

Stage 2: Classify

The second stage of the method consists of classifying the progression of each patient using a scoring. First, a specific score for each variable is calculated individually. Subsequently, from the scores of each variable, a score is obtained for each test. Finally, the test score is

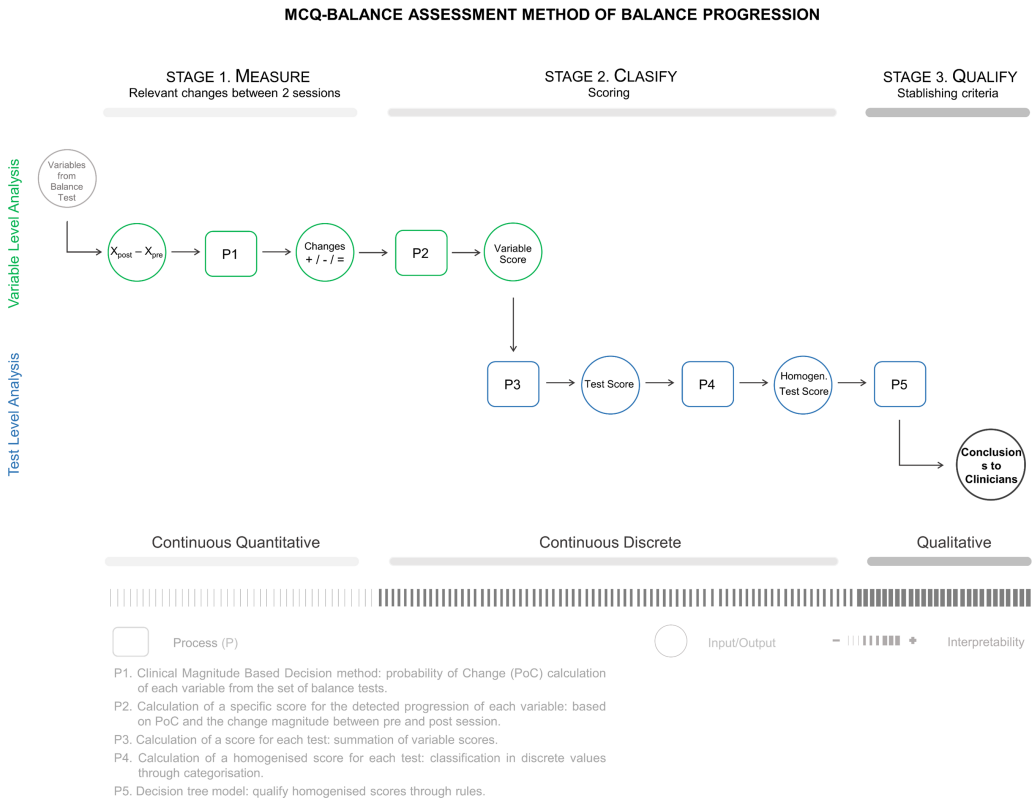


Figure 2 **MCQ-Balance assessment method: processes, inputs, outputs from the different stages.**
Full-size DOI: 10.7717/peerj.10916/fig-2

simplified, and a homogenised score (a discrete variable with the values -2, -1, 0, +1 and +2) is calculated for each of them, making it possible to compare the tests with different numbers of variables.

To determine the specific score for each variable (Score.v.mor the score of the variable m), Eq. (3) (P2) was used:

$$Score_{v_m} = PoC + CQ \tag{4}$$

- PoC: Probability of change for one unit (calculated in 2.4).
- CQ: Quantification of the change that represents the dimensionless difference between the pre- and post-sessions (for one unit) calculated using Eq. (4), in which Xdif is divided by the maximum value of the pre- or post-session. If Xdif is very large (tending to infinity), CQ approaches 1:

$$CQ = \frac{Xdif}{Max(X\ post; X\ pre)} \tag{5}$$

Considering Eqs. (2) and (3), the range of $Score_{v_m}$ is0 to +2 (positive progression) or -2 to 0 (negative progression). The score per variable is a continuous quantitative variable.

As mentioned above, the present study included five tests (four variants of the Romberg test and the LOS test); therefore, through a calculation based on the variable scores (P3),

we obtained five values referred to as $Score_{Test_n}$. In the static balance tests, four situations were considered in which five variables were obtained in each one. In the LOS test, 20 variables were obtained. Equation (5) shows how to calculate the value for $Score_{Test_n}$.

$$Score_{Test_n} = \sum_{m}^{N_{test}} Score_{v_m} \tag{6}$$

where N_{test} is the number of variables per test. Likewise, in Eqs. (6) and (7), the maximum and minimum scores that the $Score_{Test_n}$ can reach are shown.

$$MaxScore_{Test_n} = 2N_{test} \tag{7}$$

$$MinScore_{Test_n} = (-2)N_{test} \tag{8}$$

For the static balance tests, the maximum and minimum scores were +10 and -10, respectively. For the LOS test, the maximum and minimum scores were +40 and -40, respectively.

Due to the different ranges of scores for each test, it is necessary to perform a classification that homogenises and simplifies the scores independently of the number of variables selected in the previous phases. For this, a process (P4) is conducted in which the global scores are transformed into a discrete quantitative variable through categorisation (*González et al., 2020*), establishing a classification of five scores between -2 and +2. The proposed intervals are shown in brackets, which were defined based on statistical criteria, the processing and analysis of the data and the view of the clinician 2 involved in the present study:

- -2: high negative progression from $Test_n$ (30% $MinScore_{Test_n} > Score_{Test_n}$).
- -1: negative progression from $Test_n$ (30% $MinScore_{Test_n} \leq Score_{Test_n} < 10\% MinScore_{Test_n}$).
- 0: no progression from $Test_n$ (10% $MinScore_{Test_n} \leq Score_{Test_n} \leq 10\% MaxScore_{Test_n}$).
- +1: positive progression from $Test_n$ (10% $MaxScore_{Test_n} < Score_{Test_n} \leq 30\%$ $MaxScore_{Test_n}$).
- +2: high positive progression from $Test_n$ (30% $MaxScore_{Test_n} < Score_{Test_n}$).

Stage 3: Qualify

The third and final stage involves using established criteria to qualify the progression based on the resulting scores from stage two. For this purpose, rules based on a decision tree model (see Fig. 3) are proposed to qualify the progression of the balance in a patient and the influence of the involved BSS.

As mentioned above, balance is supported by the visual, proprioceptive and vestibular systems. Consequently, in the set of tests presented in Section 2.2, the patient was deprived successively of one or more BSS:

- RSEO: no BSS altered.
- RSEC: ES altered. The balance depends on the VS and PS.
- SSEO: PS altered. The balance depends on the VS and ES.
- SSEC: ES and PS altered. The balance depends only on the VS.

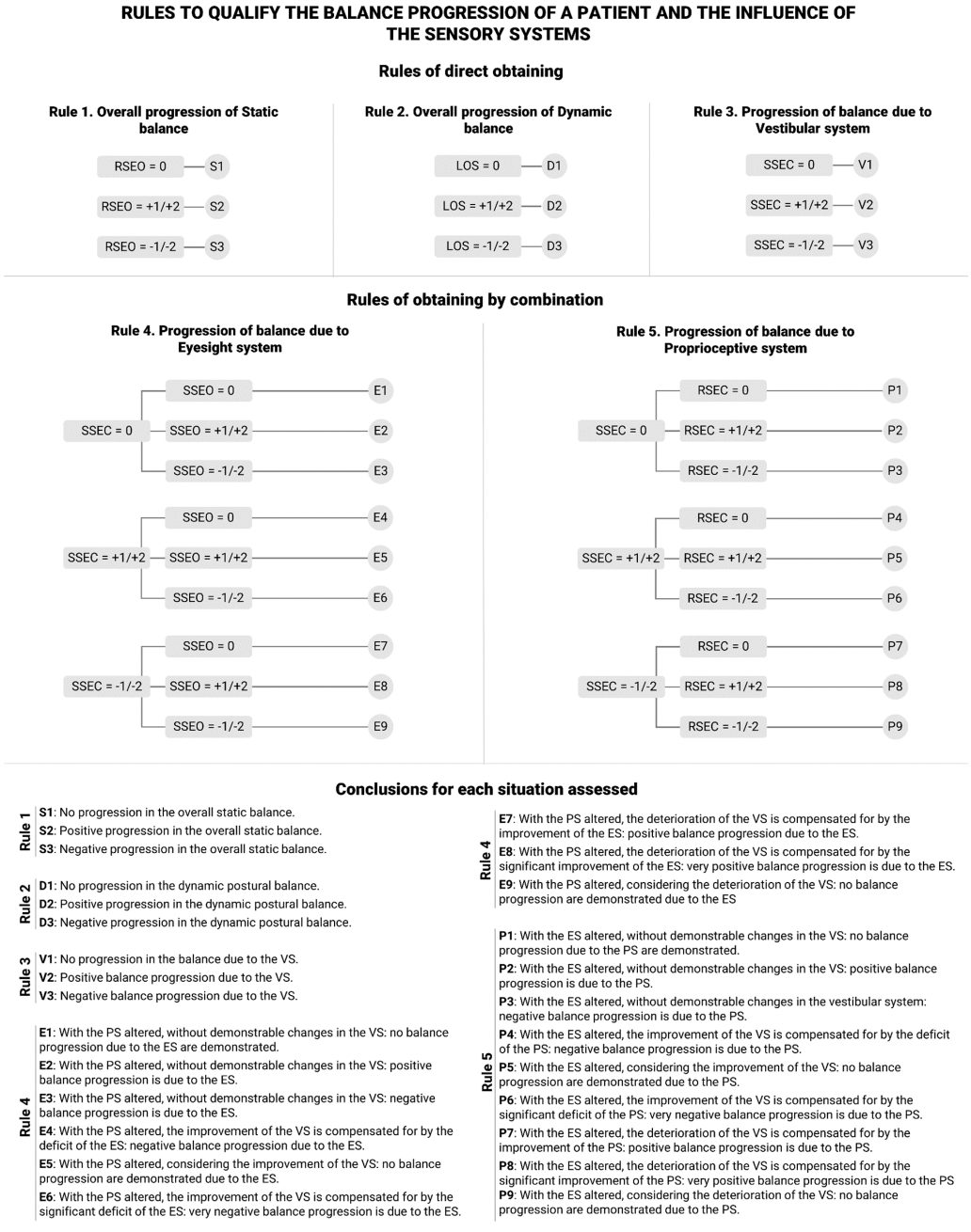


Figure 3 Rules to qualify the balance progression of a patient and the influence of the balance sensory systems. RSEO: Rigid Surface, Eyes Open; RSEC: Rigid Surface, Eyes Closed; SSEO: Soft Surface, Eyes Open; SSEC: Rigid Surface, Eyes Closed. VS: vestibular system; ES: visual system; PS: proprioceptive system. E1–E3: conclusions for the progression of static balance; D1–D3: conclusions for the progression of dynamic postural balance; V1–V3: conclusions for the progression of balance due to VS; S1–S9: conclusions for the progression of balance due to ES; P1–P9: conclusions for the progression of balance due to PS.

Full-size DOI: 10.7717/peerj.10916/fig-3

- LOS: no BSS altered. Unique dynamic postural balance test.

Thus, five rules are proposed that lead to their corresponding conclusions (see 'Conclusions for each situation assessed' in Fig. 3). The clinicians of the present study developed these conclusions. In addition, the rules are divided into two groups: those directly obtained (1, 2, and 3) and those obtained in combination (4 and 5).

Rules 1 and 2 allow to obtain a global assessment of the progression of the static balance control and the dynamic postural balance of a patient from the RSEO and LOS tests, respectively. Rule 3 allows to obtain an assessment of the influence of the VS on the progression of a patient's balance, analysing the SSEC test. Rules 4 and 5 assess the influence of the ES and PS, respectively, on the progression of a patient's balance. These rules result from the combination of SSEC with SSEO (Rule 4) and with RSEC (Rule 5), first analysing the SSEC test and then the corresponding one according to the rule.

Comparison between the MCQ-Balance assessment and clinician judgment

To analyse the application of the MCQ-Balance assessment, the patient results provided by this method have been compared with the assessment of a clinical expert (clinician 3).

The pre- and post-session data collected by clinician 1 (history and physical examination, diagnosis and functional assessment tests) were assessed by clinician 3 at the end of the field work, which allowed an assessment of the balance progression of each of the 42 patients. To avoid the results being influenced or contaminated by the interaction between the clinicians, there was no contact between them during the research.

The assessment of clinician 3 established three possible categories to evaluate patient progression: positive, null or negative progression (represented by "+", "=" and "-", respectively). Regarding the MCQ-Balance assessment, the RSEO variant of the static balance test and LOS test was chosen to make the comparison. This decision was motivated by the fact that, in the RSEO test, the subject has all the BSSs necessary to maintain stability, which corresponds to the standard situation where all BSSs are intact; it is a more favourable test and more consistent with the performance of daily living activities. In addition, in the LOS test (where the capacity or stability limits of patients are measured), the patient is also not deprived of any BSS; therefore, both tests are performed under the same conditions, which we consider in favour of the assessment used in this study (between the results of the pre-treatment and post-treatment session).

Likewise, and since clinician 3 could only establish a classification in three categories, the MCQ-Balance assessment scores have been simplified to a positive (+2 and +1 simplified to '+'), null (0 simplified to '=') and negative (-2 and -1 simplified to '–') progression in order to properly conduct the comparison.

Statistical analysis

We used the statistical software IBM SPSS statistics Version 25 (*IBM Corp, 2017*) for the statistical analysis of the data. To make the comparison between the MCQ-Balance assessment results and the assessment of clinician 3, the Cohen's Kappa statistical coefficient (κ) was chosen (*Cantor, 1996*), which is used to measure inter-rater reliability for qualitative

(categorical) items. Likewise, the confusion matrix was calculated to obtain the accuracy and percentage of false negatives.

Regarding the results of the comparison, it would be reasonable to obtain a Cohen's Kappa coefficient of a moderate or higher category (index above 0.4), as well as an accuracy of more than 70% to minimize the number of false negatives.

RESULTS

The results of the statistical analysis of the comparison between the MCQ-Balance assessment and the evaluation of clinician 3 are presented below.

Stage 1

Regarding phase 1, the average PoC is presented for each patient's tests (see Table 3). The motivation for the choice of PoC is the main output of phase 1 and, therefore, the most representative variable. Due to the volume of information handled, it was not possible to include the information at the variable level as explained in the method; however, the information of each variable from the pre- and post-sessions (pre-value, post-value, difference, MDC, STTE, PoC, CQ and the scores of each variable) of the patients' tests has been calculated and compiled as supplementary material.

Stage 2

The results related to stage 2 correspond to the homogenised scores of the five tests of the 42 patients, as presented in Table 4. This score is a discrete value between -2 and +2; negative values (-2 and -1) indicate negative progression, null values (0) indicate no progression and positive values (1 and 2) indicate positive progression.

Stage 3

Qualification of the scores of each patient, a process conducted in stage 3, is presented in Table 4 with the same identifying code detailed in Fig. 3, where the conclusions are presented based on the scores obtained.

Comparison between the MCQ-Balance assessment and clinician judgment

The results of the comparison between the MCQ-Balance assessment and the assessment of clinician 3 for the RSEO and LOS tests are presented in Table 5 and 6, respectively. They include the confusion matrix, Cohen's Kappa coefficient with its significance (p-value) and the number of false negatives.

As shown in Table 6, for the RSEO test, Cohen's Kappa coefficient is 0.752 (between 0.61–0.80 as substantial (*McHugh, 2012*), the accuracy is 83.4% between the two assessments and there are no false negatives.

As shown in Table 6, for the LOS test, Cohen's Kappa coefficient is 0.581 (between 0.41–0.60 as moderate (*McHugh, 2012*)), the accuracy is 72.9% between the two assessments and there are four false negatives, including three cases where the method did not detect changes and the clinical expert estimated worsening as well as one case where the method detected positive progression and the clinical expert estimated worsening.

Table 3 Stage 1 results: Probability of change of each patient and test.

ID	Def	RSEO	RSEC	SSEO	SSEC	LOS	ID	Def	RSEO	RSEC	SSEO	SSEC	LOS
01	P	0.12	0.79	0.16	0.40	0.11	22	P	0.22	0.29	0.26	0.70	0.72
02	P	−0.06	−0.12	−0.17	−0.38	−0.08	23	P	0.58	0.02	0.06	0.35	0.16
03	P	0.19	−0.03	0.15	−0.05	0.09	24	C	0.09	0.77	0.69	−0.74	0.01
04	P	−0.21	−0.80	−1.00	−0.52	−0.23	25	P	−0.30	−0.40	−0.60	−0.37	0.05
05	P	−0.49	−0.49	−0.30	0.06	−0.10	26	C	0.21	−0.06	−0.06	0.22	0.19
06	P	−0.14	−0.26	0.02	−0.62	−0.14	27	P	−0.15	−0.39	0.48	−0.31	−0.15
07	P	0.39	−0.02	0.46	0.18	0.34	28	P	0.00	−0.19	−0.03	0.35	−0.07
08	P	0.13	−0.31	0.21	−0.22	0.15	29	P	0.18	1.00	0.58	0.23	−0.07
09	P	−0.66	−0.02	0.15	−0.46	0.13	30	C	0.62	0.40	0.76	0.26	−0.21
10	P	−0.57	−0.02	0.15	−0.93	0.15	31	P	0.05	−0.15	−0.34	−0.13	0.06
11	P	−0.05	−0.05	−0.03	0.04	−0.07	32	P	−0.78	0.09	−0.23	−0.94	−0.32
12	P	−1.00	n/a	n/a	n/a	n/a	33	P	−0.04	−0.73	0.19	−0.06	0.00
13	P	0.25	−0.06	−0.39	−0.52	0.16	34	C	−0.80	−0.69	0.28	n/a	n/a
14	C	−1.00	−0.98	n/a	n/a	n/a	35	P	0.23	−0.19	0.26	0.18	0.17
15	P	−0.14	−0.64	−0.90	0.98	−0.05	36	P	0.10	−0.43	0.22	−0.14	−0.16
16	P	0.19	0.94	0.59	0.99	0.16	37	P	0.13	1.00	−0.03	0.29	0.16
17	C	0.17	0.00	−0.07	0.75	0.43	38	C	−0.86	−0.32	0.07	−0.71	0.00
18	P	−0.07	−0.30	0.84	−0.26	0.13	39	P	0.15	0.55	−0.06	−0.08	0.17
19	P	−0.40	−0.25	0.50	0.04	−0.26	40	C	−1.00	n/a	n/a	n/a	n/a
20	C	−0.20	−0.22	−0.64	−0.43	0.58	41	C	−0.68	−0.63	−0.84	n/a	n/a
21	P	0.07	0.05	−0.38	−0.08	0.11	42	P	−0.19	−0.11	−0.11	−0.03	0.01

Notes.

ID, patient identifier; Def, vertigo deficit; P, peripherical deficit; C, central deficit; n/a, test not performed; RSEO, rigid surface eyes open; RSEC, rigid surface eyes closed; SSEO, soft surface eyes open; SSEC, soft surface eyes closed; LOS, limits of stability.

DISCUSSION

In this study, the MCQ-Balance assessment showed an accuracy of 83.4% compared to evaluation by an expert clinician for the detection of relevant changes in balance in patients with balance disorders. The methodology used in this study is easily reproducible, given the wide availability of the resources used.

Few studies have focused on the clinical utility of posturography at the individual patient level (*Visser et al., 2008*). Likewise, although posturography is considered the gold standard, limitations exist regarding its use as a functional assessment (*Climent Barbera, 2003*). Thus, MCQ-Balance assessment method proposed, focuses on the individualised monitoring of patients, try to respond to this problem. Indeed, the transformation of information from continuous quantitative variables to conclusions in medical language facilitates the clinical interpretation of the results, providing greater intelligence to posturography devices (which is a limitation detected in posturography reports) (*Climent Barbera, 2003*).

Stages two and three of the method are adapted to clinical needs because they are the result of multidisciplinary work involving clinicians and technicians. This highlights the relevance of the conclusions that the MCQ-Balance method can generate from the results of the balance tests, which have been defined and written by the clinicians involved in the

Table 4 Stage 2 results—homogenised scores—and stage 3 results—conclusions—of each patient.

			STAGE 2: CLASIFY					STAGE 3: QUALIFY				
ID	VO	CA	RSEO	RSEC	SSEO	SSEC	LOS	R1	R2	R3	R4	R5
01	P	=	0	2	0	1	0	S1	D1	V2	E4	P5
02	P	=	0	0	−1	−1	0	S1	D1	V3	E9	P7
03	P	+	1	0	0	0	1	S2	D2	V1	E1	P1
04	P	–	−1	−2	−2	−2	−1	S3	D3	V3	E9	P9
05	P	–	−2	−2	−1	0	−1	S3	D3	V1	E3	P3
06	P	=	−1	−1	0	−2	0	S3	D1	V3	E7	P9
07	P	+	2	0	2	1	1	S2	D2	V2	E5	P4
08	P	=	0	−2	1	−1	0	S1	D1	V3	E8	P9
09	P	+	−2	0	1	−1	1	S3	D2	V3	E8	P7
10	P	–	−2	0	1	−2	0	S3	D1	V3	E8	P7
11	P	+	0	0	0	0	0	S1	D1	V1	E1	P1
12	P	–	−2	n/a	n/a	n/a	n/a	S3	n/a	n/a	n/a	n/a
13	P	+	1	0	−1	−2	1	S2	D2	V3	E9	P7
14	C	–	−2	−2	n/a	n/a	n/a	S3	n/a	n/a	n/a	n/a
15	P	+	−1	−2	−2	2	0	S3	D1	V2	E6	P6
16	P	+	1	2	2	2	1	S2	D2	V2	E5	P5
17	C	+	1	0	0	2	1	S2	D2	V2	E4	P4
18	P	=	0	−1	2	−1	0	S1	D1	V3	E8	P9
19	P	=	−2	−1	2	0	−1	S3	D3	V1	E2	P3
20	C	–	−1	−1	−2	−1	2	S3	D2	V3	E9	P9
21	P	=	0	0	−1	0	0	S1	D1	V1	E3	P1
22	P	+	1	1	1	2	2	S2	D2	V2	E5	P5
23	P	+	2	0	0	1	1	S2	D2	V2	E4	P4
24	C	=	0	2	2	−2	0	S1	D1	V3	E8	P8
25	P	–	−1	−2	−2	−1	0	S3	D1	V3	E9	P9
26	C	+	1	0	0	1	1	S2	D2	V2	E4	P4
27	P	–	−1	−1	2	−1	−1	S3	D3	V3	E8	P9
28	P	=	0	−1	0	1	0	S1	D1	V2	E4	P6
29	P	+	1	2	2	1	0	S2	D1	V2	E5	P5
30	C	+	2	1	2	1	−1	S2	D3	V2	E5	P5
31	P	=	0	0	−1	0	0	S1	D1	V1	E3	P1
32	P	–	−2	0	−1	−2	−1	S3	D3	V3	E9	P7
33	P	=	0	−2	1	0	0	S1	D1	V1	E2	P3
34	C	–	−2	−2	1	n/a	n/a	S3	n/a	n/a	n/a	n/a
35	P	+	1	−1	1	1	1	S2	D2	V2	E5	P6
36	P	=	0	−2	1	−1	−1	S1	D3	V3	E8	P9
37	P	+	1	2	0	1	1	S2	D2	V2	E4	P5

(*continued on next page*)

Table 4 (*continued*)

			STAGE 2: CLASIFY					STAGE 3: QUALIFY				
ID	**VO**	**CA**	**RSEO**	**RSEC**	**SSEO**	**SSEC**	**LOS**	**R1**	**R2**	**R3**	**R4**	**R5**
38	C	–	−2	−1	0	−2	0	S3	D1	V3	E7	P9
39	P	+	1	2	0	0	0	S2	D1	V1	E1	P2
40	C	–	−2	n/a	n/a	n/a	n/a	S3	n/a	n/a	n/a	n/a
41	C	=	−2	−2	−2	n/a	n/a	S3	n/a	n/a	n/a	n/a
42	P	+	−1	0	0	0	0	S3	D1	V1	E1	P1

Notes.
ID, patient identifier; Def, vertigo deficit; P, peripherical deficit; C, central deficit; n/a, test not performed; RSEO, rigid surface eyes open; RSEC, rigid surface eyes closed; SSEO, soft surface eyes open; SSEC, soft surface eyes closed; LOS, limits of stability; R1...R5, Rules from stage 3, consult figure 3; S, D, V, P, E, consult conclusions from figure 3.

Table 5 MCQ-Balance assessment and clinician judgment comparative: rigid surface eyes open test.

			MCQ-Balance Assessment			**Total**
			–	=	+	
Clinical Expert Assessment	–	**N**	**12**	0	0	12
		%	**28.6%**	0%	0%	28.6%
	=	**N**	3	**10**	0	13
		%	7.1%	**23.8%**	0%	31%
	+	**N**	3	1	**13**	17
		%	7.1%	2.4%	**31%**	40.5%
Total		**N**	18	11	13	**42**
		%	42.9%	26.2%	31%	**100%**
Symmetric measure	**Kappa**	**0.752**	*P*-Value	0.000	**False Negatives**	**00%**

Notes.
N, count of each case; %, percentage of total.

present study. Likewise, the definitions of the intervals of the homogenised scores have been adjusted according to the patients that have been assessed by the clinician 2.

The proposed method has advantages over traditional posturography; however, it is necessary to discuss certain issues and decisions related to the application process, which are explained below.

The first consideration refers to the chosen MBD threshold, a numerical value from which a change is considered relevant. Regarding this, the MDC has been selected as the reference value in the present study because it represents the random balance variability in addition to the measurement errors of the device and the experiment (*Furlan & Sterr, 2018*; *Steffen & Seney, 2008*). We choose the MDC, rather than the minimal important difference (MID), as the MBD threshold (*De Vet & Terwee, 2010*), consistent with previous studies (*De la Torre et al., 2020a*; *De la Torre et al., 2020b*).

The scoring proposed in the present work makes it possible to simplify the interpretation of the results of balance monitoring at the patient level. For this, the scoring allows the results to be standardized to enable a comparison between tests of the same patient and even between studies of different patients.

In the present work, and according to *De la Torre et al. (2017)*, the considered variables have the same importance and are assigned the same weight. However, future studies

Table 6 MCQ-Balance assessment and clinician judgment comparative: limits of stability test.

			MCQ-Balance Assessment			Total
			–	=	+	
Clinical Expert Assessment	–	**N**	**4**	3	1	8
		%	**10.8%**	8.1%	2.7%	21.6%
	=	**N**	2	**10**	0	13
		%	5.4%	**27%**	0%	32.4%
	+	**N**	1	3	**13**	17
		%	2.7%	8.1%	**35.1%**	45.9%
Total		**N**	7	16	14	**37**
		%	18.9%	43.2%	37.8%	**100%**
Symmetric Measure	**Kappa**	**0.581**	*P*-Value	0.000	**False Negatives**	**410.8%**

Notes.
N, count of each case; %, percentage of total.

might advise assigning a different weight to each variable depending on its importance in improving the sensitivity of the MCQ-Balance method for diagnostic purposes. In this case, the maximum and minimum achievable score for each test would be based on the weights assigned to each variable.

The choice of the five intervals to establish the homogenised scores was medically motivated. Clinically, it makes sense to make a five-level classification because the progression of the patient is towards improvement, maintenance, or deterioration of the patient's clinical picture (*Porta, 2014*), assessing the existing graduation in improvement or deterioration. The multidisciplinary agreement reached in the present work combined with the experience of fieldwork and data processing has been concluded at the presented intervals.

Regarding the conclusions in medical language resulting from the method, the ability to portray the influence of the three BSS involved in balance is highlighted in the progression of a patient's balance. In this way, the method facilitates the clinician to adapt medical treatment, focusing on the balance disorder of the patient.

MCQ-Balance assessment exceeded 70% accuracy (relative to the assessment of clinician 3) for both the RSEO test and the LOS test, and its Cohen's Kappa coefficient was >0.4. Therefore, the MCQ-Balance assessment met the accuracy goals we initially established. However, the differences between the two comparisons should be highlighted. While there were no false negatives in the comparison with the RSEO test, with the LOS test, there were four (10.8% of the sample). This is explained by the possible learning factor associated with this test (*Wrisley, 2007*), although 4 of the 37 patients who completed this test is not a representative sample; similar to the comparison with RSEO, there are more cases in which the method determined a negative progression (worsening) where clinician 3 did not. This may be due to the increased sensitivity of the method when detecting worsening that is not visible to the clinician with traditional assessment tools. Finally, we would like to establish that the decision to choose these two tests has been motivated because all BSSs are intact, a situation more in line with the performance of daily living activities. In our opinion is the

best adaptation to the assessment of the clinician 3. Although we consider the reliability obtained in this study adequate (>70%), delving into this type of comparison could result in further improved accuracy.

The simplicity of the MCQ-Balance assessment, as well as its portability and reproducibility, make it possible to systematize its use in the clinic as a complementary evaluation tool. However, future research should focus on verifying the viability of continued clinical use of this assessment, as well as its incorporation into the dynamics of a hospital rehabilitation service.

The influence of participant characteristics has not been analysed because there is no significant difference (gender) and it is not within the scope of the research; however, it was observed that older patients showed less positive progression relative to younger patients. The analysis of the possible influences of the anthropometric variables will be addressed in a future study.

Regarding the progression of the patients, it can be observed that there is no trend in improvement (positive progression) of the sample. The main reason lies in the nature of the prescribed treatments. To achieve effectiveness in rehabilitative treatment, patients need to be constant in performing the prescribed treatment, which is a great handicap of rehabilitation (regardless of subspecialty) (*Tapias, 2014*; *Essery et al., 2017*). Likewise, some cases of fear in the patients were detected in the post-session due to a negative experience in the pre-session. This explains certain cases that present a negative progression provided by the method. This problem is frequent in studies of balance disorders (*Visser et al., 2008*; *Timothy & Hain, 2019*). However, we tried to minimise the problem with additional safety measures, such as the presence of the clinician 2 and a nurse around the patient during the tests.

The lack of portability of current posturography devices is problematic. More portable devices would reduce costs (given the quicker installation process and smaller space requirements) and allow the sharing of devices between different medical centres. However, the high price of more portable devices limits their accessibility and applicability (*Uebbing, 2016*). The reduced cost of the device used in this study, as well as its portability, supports use in lower income countries that may be unable to invest in high-cost posturography equipment.

We acknowledge the major limitation inherent to the applied treatments, although the purpose of the study was not to assess the efficacy of treatments for balance disorders. Likewise, in the assessment of those patients diagnosed with BPPV to whom the Epley manoeuvre was applied, no greater positive progression was detected than the rest of the sample due to the use of a specific treatment. The effectiveness of the treatments will be addressed in a subsequent study with a sample similar to that of the present study. Likewise, future studies should compare the MCQ-Balance assessment with the BESTest (*Padgett, Jacobs & Kasser, 2012*). Besides, new output measures should be added, such as the sway directional index, sway vector (*Błaszczyk, 2016*), or even fractal dimension (*Błaszczyk & Klonowski, 2001*); as well as introduce cognitive tasks (*Raymakers, Samson & Verhaar, 2005*); De la Torre, Bonnet et al., 2020). Finally, future studies should investigate the

possibility of further improving the accuracy MCQ-Balance assessment by incorporating machine learning techniques.

Regarding the implications and possibilities of the assessment method MCQ-Balance, note that it is extrapolated to other cases of balance assessment with different tests, variables, and perspectives (e.g., balance during gait or by combining the test with cognitive tasks). Therefore, the conclusions transcend the present study.

CONCLUSIONS

This study assessed the accuracy and clinical utility of the MCQ-Balance assessment for measuring balance progression in patients with balance disorders. The results obtained with the MCQ-Balance assessment showed remarkable similarity to the assessment of an expert clinician, demonstrating the validity of this new method. We conclude that the proposed method provides objective information that facilitates the monitoring of patients with balance disorders and measurement of alterations in BSS.

Abbreviations

BPPV	Benign Paroxysmal Peripheral Vertigo
BSS	Balance Sensory System
CEICA	Research Ethics Committee of the Community of Aragon
COP	Centre Of Pressure
CQ	Quantification of the Change
CTSIB-M	Modified Clinical Test of Sensory Interaction in Balance
ES	Eye-Sight System
ICC	Intraclass Correlation Coefficient
IDERGO	Research and Development in Ergonomics
ISPGR	International Society for Posture and Gait Research
κ	Cohen's Kappa Statistical Coefficient
LOS	Limits of Stability
MBD	Magnitude-Based Decision
MCQ	Measure, Classify and Qualify
MDC	Minimal Detectable Change
MID	Minimal Important Difference
PM&R	Physical Medicine and Rehabilitation Service
PoC	Probability of Change
PS	Proprioceptive System
RSEC	Rigid Surface with Eyes Closed
RSEO	Rigid Surface with Eyes Open
SDpool	Pooled Average between the Standard Deviation of the Test and Retest
SEM	Standard Error of Measurement
SSEC	Soft Surface with Eyes Closed
SSEO	Soft Surface with Eyes Open
STTE	Short-Term Typical Error
VS	Vestibular System
Xdif	Difference Between the measures taken in two Temporal Points

ACKNOWLEDGEMENTS

We thank the I3A—University Institute of Research of Engineering of Aragon, University of Zaragoza, Zaragoza, Spain, for the materials they provided. We also thank Alcañiz Hospital (Teruel, Spain) for allowing us to use their facilities.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

The project was co-financed by the Government of Aragon, the European Regional Development Fund, and the University of Zaragoza (Spain). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
The Government of Aragon.
The European Regional Development Fund.
The University of Zaragoza (Spain).

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Juan De la Torre conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Javier Marin conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.
- Marco Polo conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Eva M. Gómez-Trullén conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Jose J. Marin conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.

Human Ethics

The following information was supplied relating to ethical approvals (i.e., approving body and any reference numbers):

The present study was approved by the Research Ethics Committee of the Community of Aragon (CEICA) (January 16, 2019).

Data Availability

The following information was supplied regarding data availability:

Raw data are available in as a Supplemental File.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.10916#supplemental-information.

REFERENCES

Aftab Z, Robert T, Wieber PB. 2016. Balance recovery prediction with multiple strategies for standing humans. *PLOS ONE* **11**:e0151166 DOI 10.1371/journal.pone.0151166.

Allum J, Zamani F, Adkin A, Ernst A. 2002. Differences between trunk sway characteristics on a foam support surface and on the Equitest® ankle-sway-referenced support surface. *Gait & Posture* **16**:264–270 DOI 10.1016/S0966-6362(02)00011-5.

Ayllón FN, Fernández DF. 2006. Función y entrenamiento de la musculatura abdominal. Una visión científica. *Journal of Human Sport and Exercise* **1(1)**:15–23.

Bartual J, Pérez N. 1998. El sistema vestibular y sus alteraciones. *Tomo* **1**:21–22.

Baumgartner BTR. 2019. Peripheral vertigo. *Available at https://www.ncbi.nlm.nih.gov/books/NBK430797/*.

Benvenuti F, Mecacci R, Gineprari I, Bandinelli S, Benvenuti E, Ferrucci L, Baroni A, Rabuffetti M, Hallett M, Dambrosia JM. 1999. Kinematic characteristics of standing disequilibrium: reliability and validity of a posturographic protocol. *Archives of Physical Medicine and Rehabilitation* **80**:278–287 DOI 10.1016/S0003-9993(99)90138-7.

Błaszczyk JW. 2016. The use of force-plate posturography in the assessment of postural instability. *Gait & Posture* **44**:1–6 DOI 10.1016/j.gaitpost.2015.10.014.

Błaszczyk JW, Klonowski W. 2001. Postural stability and fractal dynamics. *Acta Neurobiologiae Experimentalis* **61**:105–112.

Boomsaad Z, Telian S, Patil P. 2017. Treatment of intractable vertigo. In: *Youmans and Winn Neurological Surgery*. 7th edition. Philadelphia: Elsevier.

Cantor AB. 1996. Sample-size calculations for Cohen's kappa. *Psychological Methods* **1**:150–153 DOI 10.1037/1082-989X.1.2.150.

Charan J, Biswas T. 2013. How to calculate sample size for different study designs in medical research? *Indian Journal of Psychological Medicine* **35**:121–126 DOI 10.4103/0253-7176.116232.

Choi H, Lee J. 2020. Immediate effect of balance taping using kinesiology tape on dynamic and static balance after ankle muscle fatigue. *Healthcare* **8**:162.

Cicchetti DV. 1994. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment* **6**:284–290 DOI 10.1037/1040-3590.6.4.284.

Climent Barbera JM. 2003. Revista de Biomecánica. Posturografía: algo se mueve? Valencia: Instituto de Biomecánica de Valencia.

Da Costa Barbosa R, Vieira MF. 2017. Postural control of elderly adults on inclined surfaces. *Annals of Biomedical Engineering* **45**:726–738.

De la Torre J, Bonnet V, Mauti R, Chabaud P, Robert T. 2020a. Implementing a test to assess reaction, attention and inhibition capacity in elderly. *Computer Methods in Biomechanics and Biomedical Engineering* **23**:S294–S296 DOI 10.1080/10255842.2020.1815313.

De la Torre J, Marin J, Marin JJ, Auria JM, Sanchez-Valverde MB. 2017. Balance study in asymptomatic subjects: Determination of significant variables and reference patterns to improve clinical application. *Journal of Biomechanics* **65**:161–168 DOI 10.1016/j.jbiomech.2017.10.013.

De la Torre J, Marin J, Polo M, Marín JJ. 2020b. Applying the minimal detectable change of a static and dynamic balance test using a portable stabilometric platform to individually assess patients with balance disorders. *Healthcare* **8**:402 DOI 10.3390/healthcare8040402.

De Vet H, Terwee CB. 2010. The minimal detectable change should not replace the minimal important difference. *Journal of Clinical Epidemiology* **63**:804–805 DOI 10.1016/j.jclinepi.2009.12.015.

Degani AM, Leonard CT, Danna-dos Santos A. 2017. The effects of early stages of aging on postural sway: a multiple domain balance assessment using a force platform. *Journal of Biomechanics* **64**:8–15 DOI 10.1016/j.jbiomech.2017.08.029.

Derebery MJ. 2000. Diagnóstico y tratamiento del vértigo. *Revista Cubana de Medicina* **39**:238–253.

Di Fabio RP. 1996. Meta-analysis of the sensitivity and specificity of platform posturography. *Archives of Otolaryngology–Head & Neck Surgery* **122**:150–156 DOI 10.1001/archotol.1996.01890140036008.

Dieterich M, Eckhardt-Henn A. 2004. Neurologische und somatoforme Schwindelsyndrome. *Der Nervenarzt* **75**:281–302 DOI 10.1007/s00115-003-1678-z.

Dounskaia N, Peterson D, Bruhns RP. 2018. Destabilization of the upright posture through elevation of the center of mass. *Annals of Biomedical Engineering* **46**:318–323 DOI 10.1007/s10439-017-1957-7.

Doyle TL, Newton RU, Burnett AF. 2005. Reliability of traditional and fractal dimension measures of quiet stance center of pressure in young, healthy people. *Archives of Physical Medicine and Rehabilitation* **86**:2034–2040 DOI 10.1016/j.apmr.2005.05.014.

El-Kashlan HK, Shepard NT, Asher AM, Smith-Wheelock M, Telian SA. 1998. Evaluation of clinical measures of equilibrium. *The Laryngoscope* **108**:311–319 DOI 10.1097/00005537-199803000-00002.

Essery R, Geraghty AW, Kirby S, Yardley L. 2017. Predictors of adherence to home-based physical therapies: a systematic review. *Disability and Rehabilitation* **39**:519–534 DOI 10.3109/09638288.2016.1153160.

Furlan L, Sterr A. 2018. The applicability of standard error of measurement and minimal detectable change to motor learning research—a behavioral study. *Frontiers in Human Neuroscience* **12**:95.

General Assembly of the World Medical Association. 2014. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *The Journal of the American College of Dentists* **81**:14–18.

González MÁM, Villegas AS, Atucha ET, Fajardo FJF (eds.) 2020. *Bioestadística amigable*. Amsterdam: Elsevier.

Hamburg MA, Collins FS. 2010. The path to personalized medicine. *New England Journal of Medicine* **363**:301–304 DOI 10.1056/NEJMp1006304.

Hanes DA, McCollum G. 2006. Cognitive-vestibular interactions: a review of patient difficulties and possible mechanisms. *Journal of Vestibular Research* **16**:75–91.

Hansson EE, Persson L, Malmström EM. 2013. Influence of vestibular rehabilitation on neck pain and cervical range of motion among patients with whiplash-associated disorder: a randomized controlled trial. *Journal of Rehabilitation Medicine* **45**:906–910 DOI 10.2340/16501977-1197.

Hickey S, Ford G, Buckley J, O'connor AF. 1990. Unterberger stepping test: a useful indicator of peripheral vestibular dysfunction? *The Journal of Laryngology & Otology* **104**:599–602 DOI 10.1017/S0022215100113337.

Hopkins WG. 2000. Measures of reliability in sports medicine and science. *Sports Medicine* **30**:1–15 DOI 10.2165/00007256-200030010-00001.

Hopkins WG. 2017. A spreadsheet for monitoring an individual's changes and trend. *Sportscience* **21**:5–9.

Hopkins WG. 2019. Compatibility intervals and magnitude-based decisions for standardized differences and changes in means. *Sportscience* **1**:1–5.

Hopkins WG, Batterham AM. 2016. Error rates, decisive outcomes and publication bias with several inferential methods. *Sports Medicine* **46**:1563–1573 DOI 10.1007/s40279-016-0517-x.

Hoving JL, Pool JJ, Van Mameren H, Devillé WJ, Assendelft WJ, Vet HCde, De Winter AF, Koes BW, Bouter LM. 2005. Reproducibility of cervical range of motion in patients with neck pain. *BMC Musculoskeletal Disorders* **6**:59 DOI 10.1186/1471-2474-6-59.

IBM Corp. 2017. *IBM SPSS Statistics for Windows, Version 25.0.* Armonk: IBM Corp.

Ito T, Sakai Y, Ito Y, Yamazaki K, Morita Y. 2020. Association between back muscle strength and proprioception or mechanoreceptor control strategy in postural balance in elderly adults with lumbar spondylosis. *Healthcare* **8**:58 DOI 10.3390/healthcare8010058.

Jacobs JV, Horak FB, Tran VK, Nutt JG. 2006. Multiple balance tests improve the assessment of postural stability in subjects with Parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry* **77**:322–326.

Jull GA. 2004. Grieve's modern manual therapy: the vertebral column. *Available at http://vlib.kmu.ac.ir/kmu/handle/kmu/83023*.

Karlsson A, Frykberg G. 2000. Correlations between force plate measures for assessment of balance. *Clinical Biomechanics* **15**:365–369 DOI 10.1016/S0268-0033(99)00096-0.

Khalaj N, Osman NAA, Mokhtar AH, Mehdikhani M, Abas WABW. 2014. Balance and risk of fall in individuals with bilateral mild and moderate knee osteoarthritis. *PLOS ONE* **9(3)**: e92270 DOI 10.1371/journal.pone.0092270.

Ku PX, Abu Osman NA, Wan Abas WAB. 2016. The limits of stability and muscle activity in middle-aged adults during static and dynamic stance. *Journal of Biomechanics* **49**:3943–3948 DOI 10.1016/j.jbiomech.2016.11.006.

Lin D, Seol H, Nussbaum MA, Madigan ML. 2008. Reliability of COP-based postural sway measures and age-related differences. *Gait & Posture* **28**:337–342 DOI 10.1016/j.gaitpost.2008.01.005.

López LFC, Calidonio JDL. 2009. CgMed Design and construction of a platform to determine the position of the center of gravity in standing. *Revista Ingeniería Biomédica* **3**:26–36.

Ma CZ, Wong DW, Lam WK, Wan AH, Lee WC. 2016. Balance improvement effects of biofeedback systems with state-of-the-art wearable sensors: a systematic review. *Sensors* **16(4)**:434 DOI 10.3390/s16040434.

Martínez Carrasco Á. 2016. Análisis del riesgo de caídas en ancianos institucionalizados mediante escalas de marcha y equilibrio. Proyecto de Investigación, *Available at http://hdl.handle.net/10201/47708*.

McHugh ML. 2012. Interrater reliability: the kappa statistic. *Biochemia Medica* **22(3)**:276–282.

Medical Subject Headings (MeSH). 2020. *Vertigo MeSH Descriptor*. *Available at https://meshb.nlm.nih.gov/record/ui?ui=D014717*.

Muir JW, Kiel DP, Hannan M, Magaziner J, Rubin CT. 2013. Dynamic parameters of balance which correlate to elderly persons with a history of falls. *PLOS ONE* **8**:e70566 DOI 10.1371/journal.pone.0070566.

Neuhauser H. 2016. *The epidemiology of dizziness and vertigo Handbook of clinical neurology*. Elsevier, 67–82.

Orejas JIB, Varea JA, Rodrigo JV, Navas AC. 2020. Resultados y seguimiento de la rehabilitación vestibular. *Revista ORL* **11**:107–114.

Padgett PK, Jacobs JV, Kasser SL. 2012. Is the BESTest at its best? A suggested brief version based on interrater reliability, validity, internal consistency, and theoretical construct. *Physical Therapy* **92**:1197–1207 DOI 10.2522/ptj.20120056.

Patrícia Paludette D, Fabrício Santana da S, Carlos Bolli M. 2015. Comparação do equilíbrio postural entre grupos de mulheres com diferentes faixas etárias / Comparison of postural balance among groups of women with different age ranges / Comparación del balance postural en un grupo de mujeres de franjas etarias distintas. *Fisioterapia e Pesquisa* **22**:392–397.

Pawar PK, Dadhich A. 2012. Study of correlation between human height and hand length in residents of mumbai. *International Journal of Biological and Medical Research* **3(3)**:2072–2075.

Penger M, Strobl R, Grill E. 2017. Country-specific and individual determinants of dizziness in Europe: results from the Survey of Health Ageing and Retirement in Europe (SHARE). *Public Health* **149**:1–10 DOI 10.1016/j.puhe.2017.04.002.

Peterson CL, Ferrara MS, Mrazik M, Piland S, Elliott R. 2003. Evaluation of neuropsychological domain scores and postural stability following cerebral concussion in sports. *Clinical Journal of Sport Medicine* **13**:230–237 DOI 10.1097/00042752-200307000-00006.

Porta M. 2014. *A dictionary of epidemiology*. Oxford: Oxford university press.

Raymakers J, Samson M, Verhaar H. 2005. The assessment of body sway and the choice of the stability parameter (s). *Gait & Posture* **21**:48–58 DOI 10.1016/j.gaitpost.2003.11.006.

Reiley AS, Vickory FM, Funderburg SE, Cesario RA, Clendaniel RA. 2017. How to diagnose cervicogenic dizziness. *Archives of Physiotherapy* **7**:12 DOI 10.1186/s40945-017-0040-x.

Rubenstein LZ, Josephson KR. 2002. The epidemiology of falls and syncope. *Clinics in Geriatric Medicine* **18**:141–158 DOI 10.1016/S0749-0690(02)00002-2.

Ruhe A, Fejer R, Walker B. 2010. The test–retest reliability of centre of pressure measures in bipedal static task conditions–a systematic review of the literature. *Gait & Posture* **32**:436–445 DOI 10.1016/j.gaitpost.2010.09.012.

Salehi R, Ebrahimi-Takamjani I, Esteki A, Maroufi N, Parnianpour M. 2010. Test-retest reliability and minimal detectable change for center of pressure measures of postural stability in elderly subjects. *Medical Journal of The Islamic Republic of Iran* **23**:224–232.

Saxena A, Prabhakar MC. 2013. Performance of DHI score as a predictor of benign paroxysmal positional vertigo in geriatric patients with dizziness/vertigo: a cross-sectional study. *PLOS ONE* **8**:e58106 DOI 10.1371/journal.pone.0058106.

Scoppa F, Capra R, Gallamini M, Shiffer R. 2013. Clinical stabilometry standardization: basic definitions–acquisition interval–sampling frequency. *Gait & Posture* **37**:290–292 DOI 10.1016/j.gaitpost.2012.07.009.

Shumway-Cook A, Brauer S, Woollacott M. 2000. Predicting the probability for falls in community-dwelling older adults using the Timed Up & Go Test. *Physical Therapy* **80**:896–903 DOI 10.1093/ptj/80.9.896.

Shumway-Cook A, Woollacott M, Shumway-Cook A, Woollacott M. 2001. Mobility function. In: *Motor control: theory and practical applications*. Philadelphia: Lippincott Williams & Wilkins, 305–337.

Solomon D. 2000. Distinguishing and treating causes of central vertigo. *Otolaryngologic clinics of North America* **33**:579–601 DOI 10.1016/S0030-6665(05)70228-0.

Stanton MFA. 2020. Vertigo. *Available at https://www.ncbi.nlm.nih.gov/books/NBK482356/*.

Steffen T, Seney M. 2008. 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. *Physical Therapy* **88**:733–746 DOI 10.2522/ptj.20070214.

Stewart MG, Chen AY, Wyatt JR, Favrot S, Beinart S, Coker NJ, Jenkins HA. 1999. Cost-effectiveness of the diagnostic evaluation of vertigo. *The Laryngoscope* **109**:600–605 DOI 10.1097/00005537-199904000-00015.

Strupp M, Dieterich M, Brandt T. 2013. The treatment and natural course of peripheral and central vertigo. *Deutsches Ärzteblatt International* **110**:505–510.

Swanenburg J, De Bruin ED, Favero K, Uebelhart D, Mulder T. 2008. The reliability of postural balance measures in single and dual tasking in elderly fallers and non-fallers. *BMC Musculoskeletal Disorders* **9**:162.

Swanenburg J, De Bruin ED, Stauffacher M, Mulder T, Uebelhart D. 2007. Effects of exercise and nutrition on postural balance and risk of falling in elderly people with decreased bone mineral density: randomized controlled trial pilot study. *Clinical Rehabilitation* **21**:523–534 DOI 10.1177/0269215507075206.

Tapias LM. 2014. Factores individuales que influyen en la adherencia al tratamiento fisioterapéutico. Una revisión del tema.(Determining individual aspects treatment adherence in population with disabilities in Medellin city: a literature review). *CES Movimiento y Salud* **2**:23–30.

Taylor MR, Sutton EE, Diestelkamp WS, Bigelow KE. 2015. Subtle differences during posturography testing can influence postural sway results: The effects of talking, time before data acquisition, and visual fixation. *Journal of Applied Biomechanics* **31**:324–329.

Tesio L, Rota V, Longo S, Grzeda MT. 2013. Measuring standing balance in adults: reliability and minimal real difference of 14 instrumental measures. *International Journal of Rehabilitation Research* **36**:362–374.

Timothy C, Hain M. 2019. Moving platform posturography testing Computerized Dynamic Posturography (CDP). *Available at https://www.dizziness-and-balance.com/testing/posturography.html*.

Tinetti ME. 2003. Preventing falls in elderly persons. *New England Journal of Medicine* **348**:42–49 DOI 10.1056/NEJMcp020719.

Tsukamoto HF, Viviane de Souza Pinho C, Silva Junior , Alexandre da R, Pelosi GG, Marchiori , Luciana Lozza de Moraes , Vaz CRS, Fernandes KBP. 2015. Effectiveness of a vestibular rehabilitation protocol to improve the health-related quality of life and postural balance in patients with vertigo. *International Archives of Otorhinolaryngology* **19**:238–247 DOI 10.1055/s-0035-1547523.

Uebbing TJ. 2016. User experience in smart environments: design and prototyping. Master's thesis, University of Twente.

Vaupel JW, Loichinger E. 2006. Redistributing work in aging Europe. *Science* **312**:1911–1913 DOI 10.1126/science.1127487.

Vellas BJ, Wayne SJ, Romero L, Baumgartner RN, Rubenstein LZ, Garry PJ. 1997. One-leg balance is an important predictor of injurious falls in older persons. *Journal of the American Geriatrics Society* **45**:735–738 DOI 10.1111/j.1532-5415.1997.tb01479.x.

Visser JE, Carpenter MG, Van der Kooij H, Bloem BR. 2008. The clinical utility of posturography. *Clinical Neurophysiology* **119**:2424–2436.

Von Lubitz D, Wickramasinghe N. 2006. Healthcare and technology: the doctrine of networkcentric healthcare. *International Journal of Electronic Healthcare* **2**:322–344 DOI 10.1504/IJEH.2006.010440.

Wipperman J. 2014. Dizziness and vertigo. *Primary Care: Clinics in Office Practice* **41**:115–131 DOI 10.1016/j.pop.2013.10.004.

Wolf SL, Barnhart HX, Kutner NG, McNeely E, Coogler C, Xu T, Atlanta FIC-SIT Group. 1996. Reducing frailty and falls in older persons: an investigation of Tai Chi and computerized balance training. *Journal of the American Geriatrics Society* **44**:489–497 DOI 10.1111/j.1532-5415.1996.tb01432.x.

Wrisley D. 2007. Vision of balance rehabilitation a Paradigm Shift: Technology based intervention improving outcomes for dysequilibrium, dizziness, mobility, balance, and falls. Thesis, University of Florida.

Yin M, Ishikawa K, Wong WH, Shibata Y. 2009. A clinical epidemiological study in 2169 patients with vertigo. *Auris Nasus Larynx* **36**:30–35.

4.4. ARTÍCULO 4

De la Torre, J.; Marin, J.; Ilarri, S.; Marin, J.J. Applying Machine Learning for Healthcare: A Case Study on Cervical Pain Assessment with Motion Capture. *Appl. Sci (S.I. Medical Informatics and Data Analysis)* **2020**, *10*, 5942, doi: 10.3390/app10175942.

APPLIED SCIENCES

ISSN: 2076-3417

Factor de Impacto JCR (2019):

2.474

Áreas Temáticas:

Multidisciplinary Engineering: Q2 (32/91)

Applied Physics: Q2 (63/155)

Multidisciplinary Materials Science: Q3 (161/314)

Multidisciplinary Chemistry: Q2 (88/177)

El artículo 4 se encuentra actualmente en segunda ronda de revisión. La versión del artículo que se presenta a continuación es la elaborado tras una primera ronda de revisión, siendo la más actual a fecha de depósito de la tesis doctoral.

applied sciences

MDPI

Article

Applying Machine Learning for Healthcare: A Case Study on Cervical Pain Assessment with Motion Capture

Juan de la Torre [1,*], Javier Marin [1], Sergio Ilarri [2,3] and Jose J. Marin [1,4]

1 IDERGO-Research and Development in Ergonomics, Biomechanical Laboratory, I3A-University Institute of Research of Engineering of Aragon, University of Zaragoza, 50018 Zaragoza, Spain; 647473@unizar.es (J.M.); jjmarin@unizar.es (J.J.M.)
2 Computer Science for Complex System Modelling (COSMOS), I3A-University Institute of Research of Engineering of Aragon, University of Zaragoza, 50018 Zaragoza, Spain; silarri@unizar.es
3 Department of Computer Science and Systems Engineering, University of Zaragoza, 50018 Zaragoza, Spain
4 Department of Design and Manufacturing Engineering, University of Zaragoza, 50018 Zaragoza, Spain
* Correspondence: 627471@unizar.es

Received: 9 July 2020; Accepted: 24 August 2020; Published: 27 August 2020

Abstract: Given the exponential availability of data in health centers and the massive sensorization that is expected, there is an increasing need to manage and analyze these data in an effective way. For this purpose, data mining (DM) and machine learning (ML) techniques would be helpful. However, due to the specific characteristics of the field of healthcare, a suitable DM and ML methodology adapted to these particularities is required. The applied methodology must structure the different stages needed for data-driven healthcare, from the acquisition of raw data to decision-making by clinicians, considering the specific requirements of this field. In this paper, we focus on a case study of cervical assessment, where the goal is to predict the potential presence of cervical pain in patients affected with whiplash diseases, which is important for example in insurance-related investigations. By analyzing in detail this case study in a real scenario, we show how taking care of those particularities enables the generation of reliable predictive models in the field of healthcare. Using a database of 302 samples, we have generated several predictive models, including logistic regression, support vector machines, k-nearest neighbors, gradient boosting, decision trees, random forest, and neural network algorithms. The results show that it is possible to reliably predict the presence of cervical pain (accuracy, precision, and recall above 90%). We expect that the procedure proposed to apply ML techniques in the field of healthcare will help technologists, researchers, and clinicians to create more objective systems that provide support to objectify the diagnosis, improve test treatment efficacy, and save resources.

Keywords: data mining; data anonymization; health; cervical injury; neck pain; inertial sensors

1. Introduction

In the field of healthcare, the exponential increase in the data that health centers must produce and manage is significant. The need has arisen to develop procedures that make this process easier and that take advantage of all the data generated [1], detecting unknown and valuable information in health data [2]. Thus, the volume of data generated is such that its processing and analysis by traditional methods is too complex and overwhelming [3]. To tackle this challenge, data mining (DM) can play a key role, as it allows the discovery of patterns and trends in large amounts of complex data and the extraction of hidden information to help in making decisions that can improve the quality of the care processes [4–7]. It is closely linked with the scientific discipline in the field of artificial intelligence called machine learning (ML), which "employs a variety of statistical, probabilistic and optimization

techniques that allow computers to learn from past examples and to detect hard-to-discern patterns from large, noisy or complex data sets" [8].

Consequently, ML is generating growing interest in the field of healthcare (e.g., see [9,10]) for relevant special issues related to this topic), mainly derived from its possible applications, such as assessing the effectiveness of treatments, detecting fraud and abuse in health insurance, managing healthcare, making lower-cost medical solutions available to the patients, detecting symptoms and diseases [11], discovering treatment patterns from electronic medical records [12], detecting groups of incidents [13], and identifying medical treatment methods [2,3]. Likewise, ML also presents health benefits: (1) a potential reduction in the time and effort required for diagnosis and treatment, (2) the ability to examine multiple areas simultaneously, (3) a decreased potential for human error, and (4) data that are accessible anytime and anywhere [14]. Besides, DM and ML are key in the path towards personalized medicine, where the goal is to customize treatments to the specifics of each individual [15–18].

However, to take full advantage of the benefits offered by ML in the field of healthcare, several considerations are necessary. Among others, the following aspects can be highlighted:

- The data have to be structured and organized in order to properly process and transform them into suitable variables, which is essential in the development of any pattern recognition software and a highly problem-dependent task [19,20].
- Moreover, the secure treatment and management of data acquires special relevance in the field of healthcare, where the privacy of the patient must be ensured. The management of sensitive data contrasts with other fields of application of ML (fraud detection, stock prediction, etc.), where anonymization treatments may be sometimes not necessary or critical. Therefore, a specific treatment involving the anonymization and categorization of the data must be performed in order to ensure the privacy of the patients [21]. Due to privacy policies, on certain occasions if a suitable anonymization of the data is not performed and/or the required authorizations to access some data are not obtained, the needed health studies cannot be carried out. Therefore, data availability should also be considered as a key factor [21].
- Another difference with other ML applications is the existence of different costs of failures; in the health area, the cost of a false negative (e.g., failing to detect that a patient has a specific disease) is usually much higher than the cost of a false positive (e.g., if a person is initially considered to have a disease that he/she does not really have, additional tests will be performed to rule this out, which may be costly but usually less harmful than failing to diagnose an existing disease).
- It should also be considered that ML hardly ever follows a linear sequence ending at the first attempt; instead, it is rather an iterative feedback process where the stages interact with each other. Furthermore, in the field of healthcare, where the flow of data is continuous and constant, it is reasonable to assume that the model can be designed to be a "learning" model that must be continuously updated to improve its predictions over time. Therefore, the different stages needed to generate a reliable predictive model should be properly structured, from the acquisition of raw data to decision-making, which is essential to achieve the effectiveness of the model [22].

All this motivates the need to define the particularities of the application of ML techniques in the field of healthcare, where different stages in the ML workflow must be correctly defined and structured. The proper application of ML techniques would be beneficial for the clinicians, researchers, developers, and designers involved in the field of health, where the management of information acquires a transcendental role. It would favor the design of new products and services for improving healthcare access [23], creating truly accessible technological solutions [24], and enhance the relationship between health systems and people by providing adequate services at the right time [23].

Based on the above, the aims of this study are the following: (1) to show and develop the particularities of applying ML techniques in the field of healthcare, detailing all the stages that comprise this process, from the acquisition of raw data to the decision-making derived from the predictive model

generated; and (2) to demonstrate and show its practical application in a real use case. Specifically, the ML process is applied in a cervical pain assessment study with patients affected by whiplash pathologies derived from traffic accidents or other causes. This case study shows the proposed methodology in action to solve a specific relevant problem. Moreover, the applied procedure can be used as an ML application guide for other similar studies in the field of healthcare. We believe that the combination of the use case study and the machine learning methodology is a relevant contribution of this paper. We do not remain in the theoretical/methodological part only or limit our work to apply different machine learning algorithms and compare the results, as many other works do; instead, we describe the whole machine learning process, highlighting the aspects that are more relevant for our use case but at the same time providing a general framework that could be used in other health-related projects. In this way, we think that the paper could be relevant both as a specific case study and also as a reference and guideline for other similar projects.

The structure of the rest of this paper is as follows. In Section 2, we present the use case scenario studied and the clinical methods used for data collection. In Section 3, we describe the proposed procedure to develop predictive models for healthcare, illustrating each step with our work on the case study. In Section 4, we present an overall discussion of the proposal and the lessons learnt. Finally, in Section 5, we summarize our conclusions, the limitations of the study, and some ideas for future work.

2. Use Case Scenario and Clinical Methods

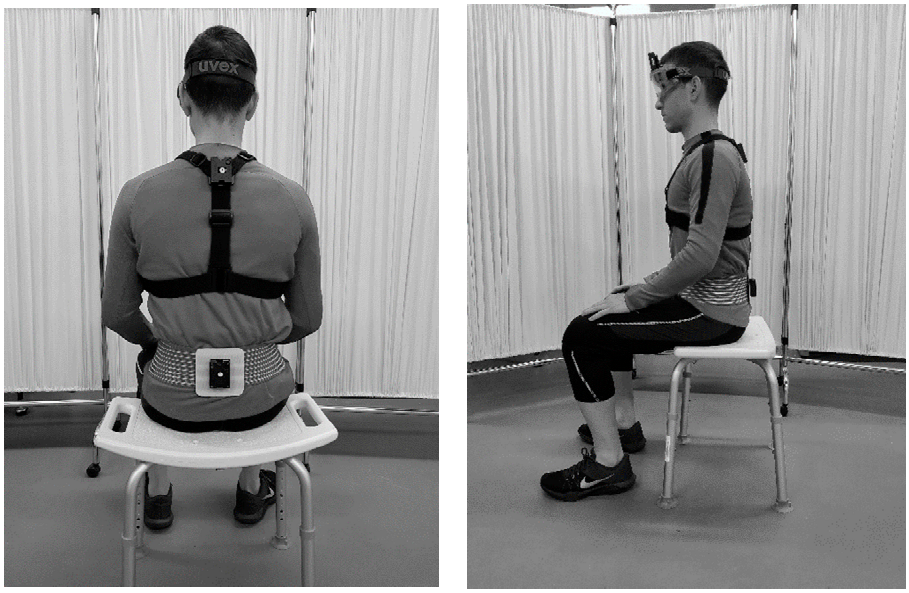
To illustrate the particularities of using ML techniques in the health area, a case study related to the detection of cervical pain is considered. The goal is to try to estimate automatically the presence of cervical pain, which can help to objectify a diagnosis and to clarify issues in case of insurance litigation. The selection of cervical pathology as a case study in this work is motivated by the fact that musculoskeletal disorders of the cervical spine have a high incidence and prevalence and are considered a public health problem, especially in developed countries [25,26]. Likewise, cervical injuries (usually due to whiplash after a traffic accident) are difficult to diagnose [26] because traumatic cervical spine injuries and their associated symptoms are diverse [25].

A real dataset was collected by evaluating the movement of the cervical spine in 151 patients (60 asymptomatic subjects, 42 with cervical pain resulting from a traffic accident, and 49 with neck discomfort due to other causes). Cervical movement assessment tests were performed by using an MH-sensor motion capture system [27,28] (see Figure 1). The participants performed a sequence of functional cervical Range of Motion (ROM) tests of the following movements: flexion-extension, rotation, and lateralization (Figure 2). The patients were collaborating subjects in order to avoid disturbances produced by non-collaborating subjects immersed in a judicial process with an insurance company [29]. The medical test was performed twice with each patient, giving a total of 302 samples.

Moreover, all the participants, who were either asymptomatic or had cervical pain, were also assessed with a clinical examination to verify that they met the inclusion criteria:

- age between 18 and 65 years;
- not immersed in a judicial process;
- no presence of surgery and/or cervical fracture.

The medical inspection, assessment by scales/clinical tests, and development of the clinical profile of the patients were conducted by clinicians. All the participants received information about the experiment and signed a consent agreement prior to the testing. The study received a favorable verdict from the Bioethics Committee of Aragón in Spain (CEICA) on 25 July 2017.



(**a**) (**b**)

Figure 1. Move Human (MH)-sensor motion capture system, cervical assessment. (**a**) Back view. (**b**) Lateral view.

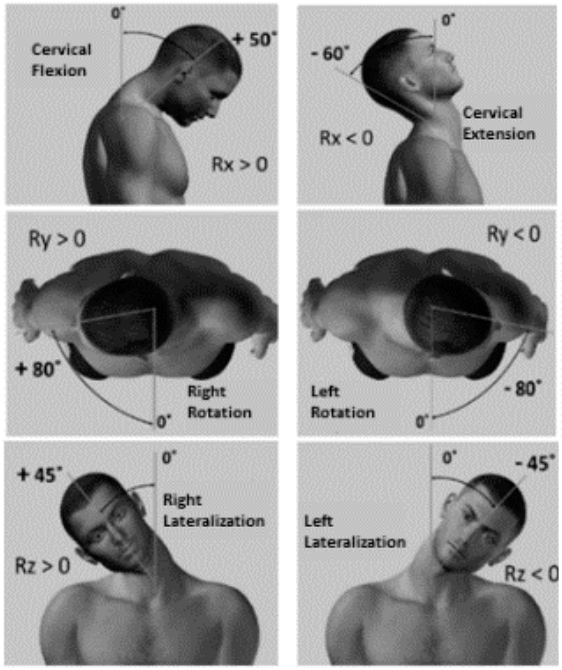


Figure 2. Cervical movements evaluated.

3. Proposed Procedure to Develop Predictive Models in Healthcare

A predictive model to support decision-making in the field of healthcare should be able to make predictions relative to relevant target clinical variables. The final goal is to deploy a system that can help in clinical decision-making (e.g., objectifying diagnoses, testing the efficacy of treatments, saving resources, providing suitable and customized treatments, etc.). The proposed procedure for continuous use in the field of healthcare is summarized and outlined in Figure 3.

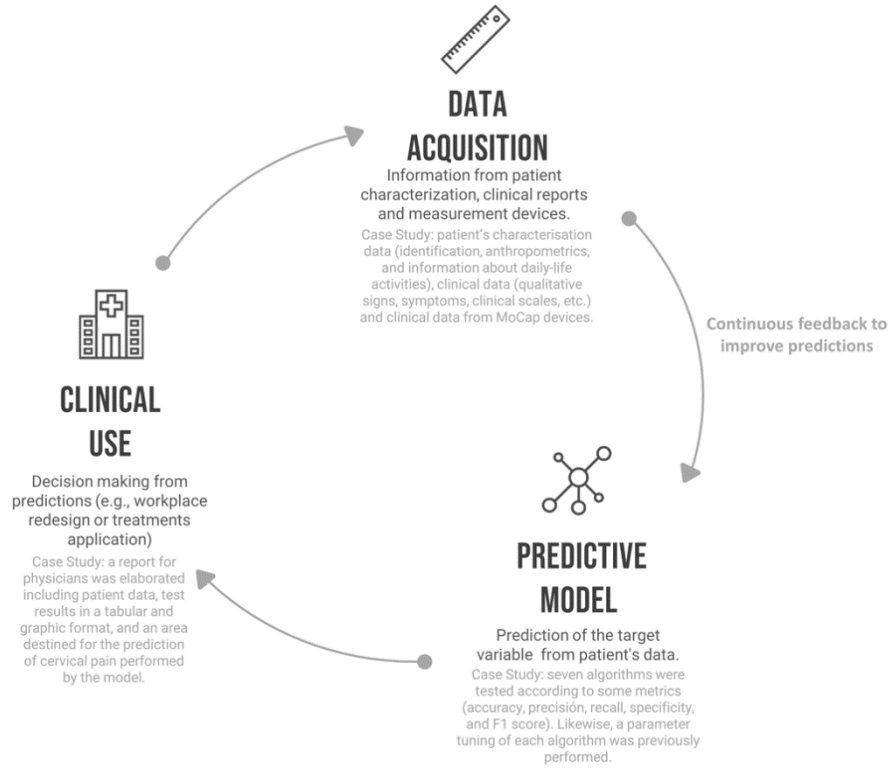


Figure 3. Application of predictive models for clinical decision-making. Icons made by monkik, smashicons and mynamepong.

The complete ML process has been considered, with the particularities of its application in the healthcare area. It is based on seven stages that range from the definition of the target to the clinical use of the system, as shown in Figure 4. Each stage is explained in the following subsections. Besides, this paper is accompanied by electronic Supplementary Material to facilitate the understanding of the different stages of application in the case study considered; specifically, we provide sample data files obtained at the end of different stages of the process (anonymized datasets) and an example report that a health professional could obtain as the output of the process.

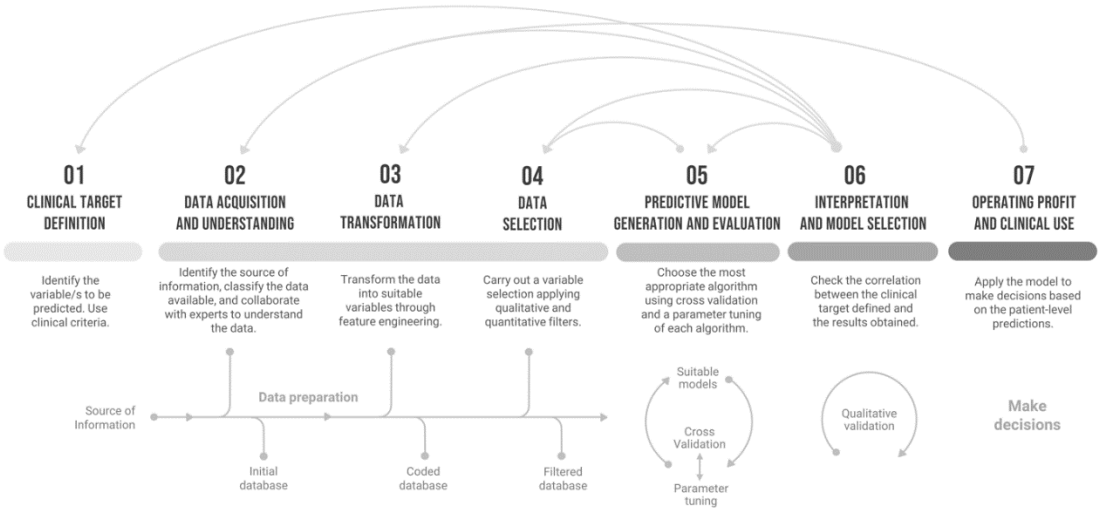


Figure 4. Project management procedure proposed for the application of machine learning in healthcare.

In order to adapt and particularize the usual process of applying ML techniques in the healthcare field and develop the project management procedure, some outstanding studies such as [20,30], as well

as the most widespread DM processes, such as knowledge discovery in databases (KDD) [31,32]; sample, explore, modify, model, and assess (SEMMA) [14,32]; and the cross-industry standard process for DM (CRISP-DM) [32–35], have been considered as a reference. Likewise, the project management procedure scheme proposed (shown in Figure 4) has been inspired by different outlines of clinical applications proposed by different authors [4,19,36–38] and adapted and extended according to our own experience and the work performed with clinicians in our case study and other related collaborations, such as the project "Mobile units for functional assessment of the musculoskeletal system" (CEICA reference of the project: OTRI-2019/0108) in collaboration with the hospital MAZ (Mutua de Accidentes de Zaragoza, Zaragoza, Spain), whose goal was to predict the degree of collaboration of patients in insurance litigation.

From the related proposals mentioned above, the CRISP-DM process has been our main inspiration to develop the project management procedure proposed in this paper. This is to be expected because CRISP-DM sets a general common framework that can be adapted to different scenarios. Thus, there are similarities between the six stages in CRISP-DM and our seven-stage proposal. For example, the CRISP-DM stage 6 "deployment" is closely related to our last stage, which is "operating profit and clinical use". As another example, stage 6 of CRISP-DM establishes that the creation of the model is not the end of the project and, similarly, in healthcare the knowledge provided to the clinician through the application of the predictive model is not the end of the process, since the system is continuously acquiring new data to improve the clinical performance. As the main difference between both procedures, we put more emphasis on data management aspects, since this is a key point in healthcare, and consider the whole process from the perspective of its application in a healthcare scenario. While only one stage for data management is considered in the CRISP-DM process (stage 3, "data preparation"), data management is the focus of three stages in our proposal (stage 2 "data acquisition and understanding", stage 3 "data transformation", and stage 4 "data selection").

The works mentioned in this section have inspired our proposal, which extends existing models by including a thorough analysis of all the data management challenges, as well as an illustration of each step through a real practical case study. Although the particularities of applying ML techniques in healthcare are exemplified in a specific case study, the procedure presented is flexible enough to adapt to any healthcare case.

3.1. Stage 1: Clinical Target Definition

In the first stage, the aim of the system is established—that is, the variables with clinical significance that the system should be able to predict are identified. Likewise, the final performance of the model and the statistical measures that will define its performance must also be defined. Measures such as the accuracy, precision, or recall are usual metrics used to assess the performance of a classification model, and metrics such as the Mean Absolute Error (MAE) or Root Mean Squared Error (RMSE), to cite two examples can be used to evaluate the accuracy of a numeric prediction.

In predictive classification in a healthcare domain, it is usual that some metric must be highlighted in such a way that the model must always be generated with the aim to minimize it. This metric is usually the number of false negatives (affecting the recall metric). The reason for this is that, in healthcare, false negatives and false positives have no similar costs, which has always been an issue that clinicians have had to deal with [39]. Moreover, the clinical need must be identified (e.g., to classify a certain type of pathology, to predict a pattern or behavior, etc.). In addition, the sample size and viability of the project must be assessed prior to its realization [34].

Application to our Case Study

In our case study, the aim of the prediction model is to predict the presence of cervical pain (the target variable) in patients who have suffered whiplash or suffer from chronic cervical pathology. In our collected dataset, the cervical pain is a binary variable (the presence or absence of pain) which

has been reported by the collaborating subjects, who were real patients undergoing an assessment process in a hospital.

Predicting the presence of cervical pain is of interest especially in the forensic field, as the incidence and prognosis of whiplash injury from motor vehicles is relevant to insurance litigations for pain and suffering [29]. The aim is to determine the presence or absence of pain with enough confidence to be able to aid clinicians to detect possible magnifications of the injury by the affected individuals and thus establish an unbiased compensation for the cervical pain [40]. It can help to identify and objectify pain in patients with a high degree of anxiety and in hypochondriac patients.

Without a loss of generality, the following target metrics have been determined for the purpose of this study, whose required threshold values have been stablished according to the criteria of clinical experts (for this particular case study, they considered that achieving this quality criteria would be enough for the system to be used as a decision-support system in production):

- Accuracy: greater than 85%.
- Precision: greater than 85%.
- Recall: greater than 90%.

The sample size is 302; although this is not a very large dataset, it contains a lot of variables with relevant information to characterize the presence of cervical pain during insurance litigation, which allows predicting the target variable, thus considering the project viable.

3.2. Stage 2: Data Acquisition and Understanding

The second stage implies identifying different sources that will allow access to the data necessary to feed the model, both in the initial phase of model design (initial training), as well as regarding a future continuous feedback when it reaches the operational stage of application in the field of healthcare. Likewise, the typology of these data will also be identified and selected. The level of accuracy that these sources of information can provide and the frequency of data collection (which may be determined by aspects such as the cost of the equipment needed, the availability of collaborating patients, etc.) will be considered in the choice [34].

In the field of healthcare, due to the diverse origins of data, their typology, their consistency, and even their veracity, the process of categorization of the data acquires special relevance for their correct structuring, understanding, and subsequent treatment.

When there are patient's personal data that make the identification of the patient possible, proper anonymization techniques must be applied. Apart from direct identification data (such as the name and last name of the patient, his/her history number, or his/her card ID), other data such as the age, nationality, height, weight, diagnosis, etc., can be used for indirect patient identification. In stage 3 "Data transformation" (see Section 3.3), certain techniques are presented to safeguard the privacy of patients and avoid the loss of useful information for the generation of the model, but these sensitive variables must be identified in this second phase.

Although there are classifications of information and data in the healthcare environment [30], alternative complementary classifications are proposed in this paper for a better understanding and structuring of the data. This has been motivated by the needs of medical staff, as well as by specialists of the medical legal/forensic field collaborating with us in our case study: a greater variety of classifications is of interest in order to include the perspectives of all the parties involved.

Clinical Data and Patient Data

Firstly, two types of data related to patients can be identified when we distinguish between clinical data and other data related to the patient:

- Patient characterization data. They are generally static data (although there may be small fluctuations in certain data values over large time intervals, such as in the case of the weight of the patient). They can be grouped in the following categories:

    - Identification data: name, age, gender, educational level, nationality, etc.
    - Temporary data: dates of control or highlighted clinical evolution, visits to the hospital, start and end of treatments, etc.
    - Anthropometric data: measurements of the size and proportions of the human body, such as the height, weight, percentage of fat and muscle, foot length, abdominal perimeter, etc., that usually require instrumentation to be obtained (scale, tape measure, etc.).
    - Daily life activities (DLA) data: data usually reported by the patient related to the habits and activities that he/she usually performs on a daily basis. In some cases, some of these data can be measured using wearables sensors or other devices deployed in smart homes.
- Clinical data: data of a medical nature that may require instrumentation and medical tests for their acquisition.

Data According to the Degree of Objectivity

Another possible classification is to categorize the data according to the degree of objectivity:

- Measures: objective data that do not require assessment by a clinician. These data are not affected by the reproducibility factor. Examples are test or clinical scales, test results, or data collected by medical instrumentation. Data recorded by sensor devices provide objective data on some measurable dimensions of the patient and can be of different types: motion capture (MoCap) sensors, surface electromyography (EMG), stabilometric platforms, dynamometers, etc.
- Assessed data: information that depends on the assessment of the clinician, such as diagnoses and treatments.
- Reported data: subjective information provided by the patient regarding his/her condition (perceived symptoms).

Data According to Clinical Considerations

We also present a classification that groups the collected data according to clinical considerations:

- Clinical profile: data about symptoms and clinical signs of the patient that can lead to a diagnosis by the clinician. We refer to symptoms when they are of subjective nature, reported by the patient, and to signs if they are objective and obtained by the clinician about the pathology. In addition, the signs can be qualitative (binary) or (discrete or continuous) quantitative (e.g., the temperature of a thermometer, image tests, other measurements, etc.).
- Treatment data: data about the treatment that the clinician has prescribed, such as the type of treatment, number of rehabilitation sessions, drugs received, surgery, etc.
- Clinical scales, tests, or surveys: data resulting from scales or validated and protocolized tests whose objective is to obtain objective information about the patient (e.g., the timed up and go test, the Unterberger test, psychological tests, etc.).
- Medical history: data concerning the patient's clinical history, ordered chronologically (e.g., first hospital visit, imaging test for diagnosis, treatment administration after diagnosis, etc.).

Data According to their Data Types

Finally, the different data variables can be grouped according to their types, independently of the specifics of the health area. For example, we could consider:

- Qualitative variables: also called categorical variables, they are variables that are not numerical. They describe data that fit into categories (e.g., educational level, the level of development of a disease, the level of invasiveness of a treatment, etc.).
- Quantitative variables: also called measurement or numerical variables, they represent quantities of different nature. They can be divided into discrete variables that can only take a finite number of values (e.g., the number of rehabilitation sessions, score on a clinical scale, age, etc.),

and continuous variables, which can take values in an infinite/continuous range of possible values (e.g., the temperature of a thermometer, weight, Body Mass Index (BMI), etc.).

- Textual data: data that are directly collected in text format, such as handwritten annotations in medical histories.

In this stage, data profiling [41] and cleaning [42] must be applied to detect potential problems and, if possible, fix them. By categorizing the data, using one of the proposed classifications (or another one that could be useful for a specific use case) independently, or several of them at the same time, the process of the understanding and assimilation of the data available and their scope is facilitated. This step must be carried out to obtain a preliminary database, containing the data collected, that will be called the initial database. For this task, having the support of both a clinical expert and a technologist is recommended. In some cases, when the amount of data to handle is large or coming from different data sources, a data warehouse can be created to integrate all the information and allow the easy and efficient analysis of the data stored [43,44].

During data collection for classification tasks, it is also important to collect enough instances/samples to represent in an appropriate way the different classes that must be predicted. In case there is imbalance regarding the number of samples in the different target classes, this should be identified as part of the data profiling, and some strategies could be applied to deal with this issue [45,46] (e.g., to try to prevent the majority class from dominating the predictions in a harmful way).

Application to Our Case Study

In this case study, the initial database was prepared according to the final goal, which was predicting the presence/absence of cervical pain. The categorization of the data was jointly agreed by physicians and technologists, considering the application of this research in the legal field and the degree of objectivity of the data. According to the four classifications presented, the data from our case study could be classified as shown in Table 1, following the exposed criteria:

Table 1. Possible classifications of the case study data.

1. Clinical Data and Patient Data	2. Data According to the Degree of Objectivity	3. Data According to the Clinical Considerations	4. Data According to Their Data Types
Patient characterization data: - Identification: name, age, gender, educational level, etc. - Temporary: accident date, visits to the hospital, date of the range of motion (ROM) test, etc. - Anthropometric: weight, height, body mas index, foot length, etc. - Daily life activities: physical activity intensity, workplace, etc.	Measures: all the data from the MoCap sensors or Whiplash scale (WDQ).	Clinical profile: - Symptoms: pain periodicity, feeling of instability, limitation of mobility, etc. - Signs: contracture, limitation of mobility, spinal column alterations, etc.	Qualitative variables: educational level, limitation of mobility, contracture, etc.
Clinical data: feeling of instability, surgery, all the data from the MoCap sensors, etc.	Assessed data: contracture, limitation of mobility, Jackson contraction, etc.	Treatment: n/a.	Quantitative variables: - Discrete: WDQ, age, etc. - Continuous: all the data from the MoCap sensors, weight, height, etc.
	Reported data: pain periodicity, feeling of instability, etc.	Clinical scales or tests: WDQ.	Textual data: n/a.
		Medical history: accident date, visits to the hospital, etc.	

The classifications shown in Table 1 illustrate different useful perspectives of the data collected in the case study. Specifically, and considering the final goal of objectively predicting the presence of cervical pain, the classification that best suited the point of view of the health professionals participating in our case study was the following (a combination of the first two classification approaches described):

- Patient characterization data: the data measured and reported, such as the identification information (e.g., ID, age, gender, etc.), anthropometrics (e.g., height, weight, BMI, etc.), and data relative to the activities of daily life (e.g., weekly physical activity and its intensity, workplace, etc.).
- Assessed and reported data: such as the characterization of a cervical accident (e.g., the time since the accident, type of impact of the traffic accident, position of the head, type of effect, etc.), qualitative signs (e.g., column alterations and contractures), clinical scales (e.g., whiplash scale—WDQ), and symptoms (e.g., instability, limitation of mobility, etc.).
- Measured data: such as data from a cervical assessment test with MoCap or from each movement studied, for example the angles reached and the speeds of the movements (e.g., maximum values, minimum values, average values, etc.).

For the purposes of medical legal/forensic assessments, a classification according to the degree of objectivity of the clinical data (assessed and reported) is of interest. Thanks to the understanding of an expert technologist in insurance litigation, it has been possible to identify objective information in the case study, such as the presence of a contracture, column alterations, the WDQ, the characterization of the traffic accident (in case such an event had occurred), etc.

The initial database is presented as Supplementary Material File S1, which collects the dataset that has been considered for the development of the model.

3.3. Stage 3: Data Transformation

Once the data to be considered in the initial database have been selected, certain transformations of the data must be performed in order to handle empty values, perform data transformations to define the required variables and adapt them to the required format, and ensure the anonymization of the data.

The information obtained from the different sources can correspond to variables already defined and structured or to raw data, and it can be presented as numerical, text, curves, images, etc. [47]. Transforming raw data into the format required for the application of specific ML algorithms is a common pre-processing step to be performed, and it could also be useful because the volume of data to be handled could be reduced, and its predictive power could be significantly increased, making it possible to have a significantly lower volume of data to achieve a reliable and stable predictive model [5,19]. To be exploited by traditional DM and ML algorithms, textual data can be transformed into structured variables and different text mining techniques can be applied [48,49]. Besides, depending on the purpose, unsupervised learning approaches can be applied on the texts—for example, for dimensionality reduction (e.g., using the Self-Organizing Map (SOM) method [50]), for clustering documents according to their similarity, or for discovering topics in documents (e.g., probabilistic topic modeling by using Latent Dirichlet allocation) [51].

This stage of transformation of raw data is known as feature engineering and is performed prior to modelling [52]. In some cases, several data mining techniques can be applied to extract features from raw data. For example, Principal Component Analysis (PCA), like the SOM method mentioned above, is a dimensionality reduction method that is often used to reduce the dimensionality of large data sets; it can be used to tackle the problem of high dimensionality that appears in some projects when the number of variables is excessively high compared to the total number of samples (see Section 3.5) [39]. In the field of healthcare, the following types of transformations can be highlighted:

- Texts that become “positive or negative opinions”, concepts, dates, values, etc., through the application of text mining techniques [53,54].

- Images of medical tests that are converted into binary variables related to a pathology or other binary representations.
- Curves or other graphical representations from which information can be extracted as statistical variables, such as the mean, standard deviation, skewness, quartiles, etc.
- Different imputation techniques [18,55–57] that can be applied in order to fill empty values, either by means of interpolation or by using other procedures. Alternatively, some records may need to be discarded (e.g., if several key data values are missing).
- The potential addition or creation of variables based on the knowledge of clinical experts and technologists, either as a combination of existing variables or based on experience acquired in similar studies [30].
- Data anonymization, applied in order to preserve the privacy of patients [58]. The trade-off between privacy and information loss should be considered. A detailed analysis of data anonymization techniques for health data is out of the scope of this work but, for illustration purposes, some examples of transformations that can be applied to guarantee the privacy of sensitive information are:

    - Transformation of continuous variables (e.g., age, size, weight, income, etc.) into ordinal variables by defining suitable ranges. Normalization (e.g., min-max normalization) or standardization (z-score scaling) techniques could also be applied to transform the quantitative variables into variables in the range from 0 to 1.
    - Transformation of qualitative variables (e.g., diagnosis, treatment, education level, etc.), that could be classified, according to a scale, into ordinal variables (e.g., severity of diagnosis, treatment risks, range of studies (“high school”, ”BS”, ”MS”, “PhD”), etc.). In this way, these variables cannot be associated with a specific patient, thus preserving the patient’s privacy.
    - Transformation of qualitative variables (e.g., nationality, physical description, address, etc.), that could not be classified according to a scale into groups (e.g., by continents or country grouping, by groups according to a general description, by postal code, etc.) so that these variables cannot be associated with a specific patient.
    - The previous anonymization techniques are examples of the generalization of attribute values. Other possible techniques and privacy-preservation methodologies include adding noise [59], k-anonymity [60], differential privacy [61], etc.

Another important aspect to mention here is the need for the normalization of quantitative attributes; if we have quantitative variables with very different scales, the variables that can take larger values could end up dominating others when learning a predictive model. This is unsuitable because it mistakenly attributes more importance to those variables just because their usual values are higher than the values of other variables.

After applying the specified transformations, the initial database is transformed to an encoded (modified) database, which includes all the available information in the format of multiple variables.

Application to Our Case Study

In the case of the cervical pain assessment study, this process consisted of transforming the data from the MoCap sensors into variables (i.e., mean, deviation, maximum, minimum, etc., of the range of the movement) and transforming the clinical and characterization data of the patients into variables considering sensitive variables and their anonymization. Some ranges of variables were associated with a numeric coding to improve the anonymization process and the generation of the predictive model. The process was carried out jointly by a clinical expert and a technologist. As in our case study we have transformed all the quantitative variables into discrete variables, no additional normalization

was needed. The following are examples of the transformations performed (the numeric coding used is indicated in brackets):

- Age transformation in a three-tier classification: under 30 (0), between 30 and 55 (1), and over 55 (2).
- Transformation of the level of weekly exercise into a classification of three levels as a function of the time required and the intensity of the exercise: slight (0), moderate (1), and intense (2).
- Transformation of the level of studies in a classification of three levels with their assigned numerical coding: basic/high school (0), medium/bachelor (1), and superior/university studies (2).
- Weight transformation in a three-tier classification: less than 60 Kg (0), between 60 and 85 Kg (1), and greater than 85 Kg (2).
- Height transformation in a three-level classification: less than 158 cm (0), between 158 and 185 cm (1), and greater than 185 cm (2).
- Transformation of the body mass index (BMI) into a three-tier classification: under 24 (0), between 24 and 30 (1), and over 30 (2).
- Grouping of the cervical pain periodicity to create a variable of three levels: sporadic (0), discontinuous (1), and frequent (2).

The completeness and quality of the data recorded is also a key aspect in any ML pipeline, and particularly in the health area. In our case study, some variables were initially incomplete, due to the lack of collected data from certain patients (five patients). These incomplete data were related to specific features (e.g., the workplace of the patient, his/her age, the dominant side of his/her body, the date of the traffic accident, etc.), and were later collected by a telephone call.

The encoded database is presented as Supplementary Material File S2, where the variables are obtained after the transformations that are carried out from the initial database are collected and the anonymized variables are highlighted (see Variable_View).

3.4. Stage 4: Data Selection

The next stage is to filter the encoded database obtained in the previous phase and select the most useful variables, applying different filters in a way that will lead to obtaining a filtered database, which will be the basis of the predictive model. Possible successive filters to be used include the following:

1. Filters due to ethical and legal issues: Discard personal or private variables and those that are unimportant for the purpose of the predictive model, such as names, clinical history numbers, telephone numbers, and addresses. The filtered database must be anonymous with respect to existing regulations on the protection of personal data, so a previous anonymization process becomes essential in order to keep as much important data as possible. Notice that during the previous step (data transformation, see Section 3.3), some data are transformed to increase privacy; in this step, privacy might need to be further increased by not selecting some sensitive data in case those data have not been properly transformed previously or in the case of other sensitive data that are irrelevant for predictions.
2. Manual selection: Screening based on the needs set by the target of the prediction, removing outliers or variables with a lot of missing data [5]. It is highly recommended that this filtering be conducted by an expert in the healthcare field.
3. Automated attribute selection: specific software and algorithms can be used for filtering, calculating the gain ratio for each of the variables and rejecting those with low predictive power—for example, using regression techniques.

Application to Our Case Study

In our case study, the encoded database initially included 230 variables in the initial dataset (the "Coded database"), which were reduced to 28 after the following consecutive filtering steps (see Table 2):

1. The removal of variables related to personal and ethical data not anonymized previously and with no predictive power. The variables name, surname, telephone number, address, and email were removed in this filtering.
2. Manual filtering performed by physicians and technologists, corresponding to non-objective or inappropriate variables in the medical legal/forensic field (e.g., the sensation of instability, mobility limitation, pain periodicity, etc.), as well as variables with missing data that could not be completed (e.g., the position of the head and type of effect in case of a traffic accident, intensity of work, etc.). In this filtering, 74 variables were removed according to the criteria indicated.
3. Finally, a filtering was applied based on the gain ratio of the variables. We used the IBM SPSS modeler software [62] (v. 18), discarding 123 variables with low predictive power (we selected those with a gain ratio higher than 0.95 out of 1). Variables such as the average angle, standard deviation, complementary angle, weight, height, etc., were selected. The selection of these 28 variables is consistent with the target variable (the presence or absence of pain) and its associated requirements, since it is a desirable situation for clinicians that these variables are objective, represent the main cervical movements, and correspond to clinical data objectively acquired by the clinicians.

Table 2. Final variables considered in the case study after feature selection.

Patient Characterization Data	**Gender**	**Age**	**Educational Level**	
Clinical data: assessed and reported data	Contracture	Traffic accident	Spinal column alterations	WDQ
Clinical data: data measured with sensors	Mean Speed [°/s] in:	Flex.-Ext.	Rotation	Lateralization
	Max Speed [°/s] in:	Flexion	Right Rotation	Right Lateral
		Extension	Left Rotation	Left Lateral
	Max Angle [°] in:	Flexion	Right Rotation	Right Lateral
		Extension	Left Rotation	Left Lateral
	Total Range [°] in:	Flex.-Ext.	Rotation	Lateralization
	Total Length [°] in:	Flex.-Ext.	Rotation	Lateralization

Table 2 shows the 28 final variables of the filtered database; the detailed information of each variable, with the different associated values for the different data instances, is included as Supplementary Material File S3.

3.5. Stage 5: Predictive Model Generation and Evaluation

In the next stage, a predictive model according to the established objective is designed based on the filtered database obtained in the previous stage. To do this, we must select those algorithms that are considered viable for the specific clinical project, such as decision trees, neural networks, or support vector machines (SVM), among others [63]. If the volume of data is very high, the use of specific support software can facilitate selecting a suitable algorithm by performing iterations before generating the full predictive model. For example, the IBM SPSS modeler classifier node (v. 18) can create several models and then compare them to select the best approach for a particular analysis.

In this stage, the performance of the selected algorithms should be tested to choose the most convenient one to implement in the predictive model. To evaluate their stability and effectiveness, different cross-validation approaches can be considered [19,30,64–66]. The simplest method consists of separating the sample into two sub-samples: one to train the model and another one to test it (holdout method). Other more advanced methods include dividing the sample into k sub-samples (k-fold cross validation), stratifying the sample with the same percentage of each class (stratified k-fold cross validation), or even making as many combinations as the number of data instances (leave-one-out cross

validation). Furthermore, a validation set could be used (besides the "test set" and the "training set"), which is a set of examples used to tune the parameters of a classifier [67]. This is useful because the performance of the selected algorithms can be improved through parameter tuning, which consists of varying values of parameters of the algorithms in order to find the most suitable configuration for each of them. Once the suitable parameter configuration for each algorithm is selected, the performance of the different algorithms can be compared.

It must be stressed that the most appropriate prediction method depends on the data. Besides, overfitting (a phenomenon that occurs when the adjustment of the model to the training data is too strong and, as a consequence, finds difficulties in obtaining suitable conclusions about unobserved data) should be avoided, as this would lead to a model that will only be able to make predictions for the data with which it has been trained [52]. There are several techniques to deal with this, such as regularization (smoothing the models), data augmentation (increasing the training data), early stopping, etc. As an example, early stopping implies that the training process of the model must be stopped before overfitting [30,68,69], that is, before the model adjusts too much to the training data (i.e., before the performance of model gets worse on the validation data).

Over-parametrization is a recurrent problem in the application of ML techniques in healthcare, where there are cases where the ratio between variables and data is very high and overfitting effects can occur. The opposite can also happen when the volume of data is very high, but the number of variables is excessively reduced and/or has little predictive power. Therefore, if needed, depending on the results obtained we could come back to Stage 4 "Data Selection" to reduce the number of variables. As a guideline, to avoid over-parametrization the 1 to 10 ratio between variables (attributes or features) and data (number of instances) should not be exceeded [20,52]. If there was a smaller ratio between the variables and data, a more exhaustive screening of variables would have to be carried or a larger sample would have to be obtained. In terms of classification, having an excessive number of variables will lead to not solving the problem or not achieving the proposed objective because the model will only be able to classify the training data correctly [70].

To select the most suitable algorithm for the project, it is necessary to follow an organization strategy, storing each previous version and modification of the project [52]. In this way, the results are contrasted (training data, validation, errors, etc.). After the selection of the most effective algorithm, we will be able to generate the definitive predictive model that incorporates the already-existing filtered database. If the predictive model were evaluated positively, a continuous learning process could begin by receiving periodic information from the assessment tests performed on future patients, as shown in Figure 3.

Application to Our Case Study

In view of the target of our clinical study to classify the presence of cervical pain, only supervised learning algorithms must be selected, dismissing unsupervised learning algorithms. The algorithms selected in our study were logistic regression [71], decision trees [72], random forests [73], SVM [73], neural networks (MLP neural networks) [74], k-Nearest Neighbors (KNN) [73], and Gradient Boosting Algorithm (GBA) [75]. All these are popular supervised machine learning approaches. The main parameters selected to perform the parameter tuning of those algorithms are shown in Table 3. In our experimental evaluation, we combined the parameters of each algorithm presented in Table 3 using the software tool Weka [76] (v. 3.8); specifically, for the parameter tuning we used the Weka Experimenter user interface.

Table 3. ML approaches considered and their main parameters.

Approach	Main Parameters
Logistic regression	Ridge value in the log-likelihood: from 10^{-4} to 10^{-12} (parameter change every 10^{-2}).
Decision tree (C4 pruned)	Number of instances per leaf: 3/5/10/15/20. Confidence factor for pruning (Conf.): 0.15/0.25/0.35.
Random forest	Maximum depth of the tree: 3/4/5. Number of trees: 25/50/100/200.
Support vector machine (SVM)	Tolerance: 10^{-3}. Kernel function: radial basis function (RBF). Epsilon for round-off error: 10^{-12}. Complexity (C): 0.25/0.5/1/2/4. Gamma (kernel width): 0.01/0.25/0.5/1/2.
Neural Network (MLP neural network)	Type: multilayer perceptron (MLP). Learning Rate (LR, the amount the weights are updated): 0.2/0.3/0.4/0.5. Momentum (Mom., applied to the weights during updating): 0.1/0.2/0.3. Number of epochs for training: 500. Number of hidden layers: 15. Auto-built option in Weka set to true.
K-Nearest Neighbors (KNN)	Number of Neighbors (K): 1/3/5/7/9/11/13/15/20. Distance function: Euclidean distance, Manhattan distance.
Gradient Boosting Algorithm (GBA)	Iterations (Iter.): 10/20/50/100. Weight threshold (W.T.): 50/100/200. AdaBoost implementation provided by Weka.

In our case study, the predictive models generated by applying the previously selected algorithms are shown in Table 3. As mentioned previously, the DM software used in this study was Weka (v. 3.8). Considering the current availability of data for our case study, the performance of each parameter configuration of each algorithm was tested using a validation set which is the same as the test set; using the same set for validation and testing is not the ideal situation, but we decided not to reserve a part of the available dataset for validation because of the moderate size of our sample. We considered the accuracy metric (i.e., the percentage of instances correctly classified) for determining the most suitable algorithm configuration. Figure 5 shows the performance of the algorithms obtained during parameter tuning using a k-fold cross validation (k = 10).

Consequently, the parameter configuration selected for each algorithm is shown in the first row of Table 4. The effectiveness of the models generated by k-fold cross validation (k = 10) was evaluated. The following metrics were computed in order to determine the most suitable algorithms (see Table 4): the accuracy (the percentage of instances correctly classified), the precision (the percentage of patients with pain correctly classified over all the patients labelled by the algorithms as patients with pain), the recall/sensitivity (the percentage of patients with pain correctly classified over all the patients with real pain), the specificity (the percentage of healthy patients correctly classified over all the patients who are really healthy), and the F1-score (the harmonic average of the precision and recall).

We have noticed using the software Weka, which provides the attribute weights of each model in the results display section, that the variables with greater predictive power in all the predictive models evaluated are the following: maximum speed in all the movements, the existence of a traffic accident, and the presence of a contracture. In our case study, there is no indication of over-parametrization; as described in Section 3.4, we have a sample of 302 instances and 28 selected variables, which complies with the 1 to 10 ratio between the variables and data.

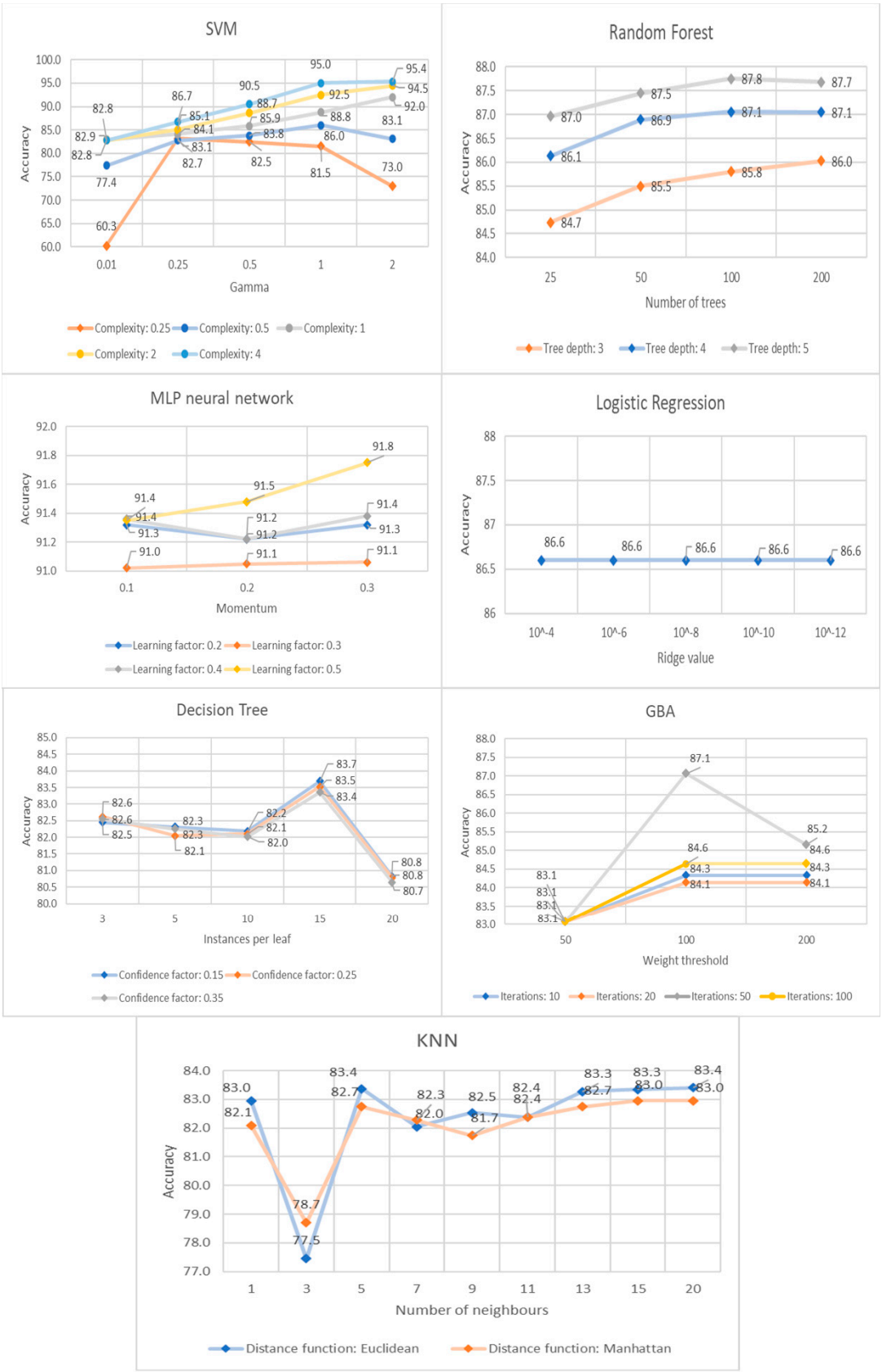


Figure 5. Parameter tuning of the seven algorithms selected.

Table 4. Metrics considered for algorithm selection and the results obtained for each algorithm.

	Logistic Regression	SVM	Decision Tree	Random Forest	MLP Neural Network	KNN	GBA
Parameter Selection	Ridge: 10^{-8}	C: 4; Gamma: 2	Instances per leaf: 15; Conf.: 0.15	Trees: 200 Depth: 5	LR: 0.5; Mom.: 0.3	K: 20; Euclidean	Iterat.: 50 W.T.:100
Accuracy	86.6%	95.4%	83.7%	87.7%	91.8%	83.4%	87.1%
Precision	88.6%	95.4%	86.9%	86.1%	93%	83.5%	87.1%
Recall/Sensitivity	89.6%	97.8%	84.1%	92.3%	92.5%	91.2%	92.9%
Specificity	82.5%	91.7%	80.8%	76.6%	90.5%	71.7%	78.3%
F1 Score	89.1%	95.3%	85.5%	88.9%	92.8%	83.2%	86.9%

3.6. Stage 6: Interpretation of Results and Model Selection

Once a statistically acceptable predictive model has been created, it can be deployed to be used in production, since it can provide predictions with good accuracy to support decision-making (Figure 4, stage 6). However, for the exploitation of the model, it is necessary first to evaluate the degree of correlation between the results obtained and the previously defined clinical target (in our case study, the prediction of cervical pain). Besides, some models (e.g., the tree-based classifiers) could be explored to analyze which particular features are the most decisive factors in the classification model.

If the evaluation of the correlation does not yield satisfactory results—that is, if the consonance of the results with the target is not achieved or the model does not have enough predictive power according to the minimum goals established initially—the previous stages must be repeated with the objective of improving and optimizing the process according to the target initially set, leading to an iterative process (see Figure 4) [20,77]. There can be several causes of low correlation or poor predictive power:

- The target is difficult to achieve or too complex;
- Patients are inadequately characterized;
- The sample is insufficient;
- The data include variables that are not necessary (e.g., irrelevant variables, confounding factors, or redundant variables) or do not include those that are;
- Problems exist in the predictive model generation stage regarding the selected algorithm [78], overfitting, or over-parameterization.

For effective deployment, not only the accuracy of the models but also the resources must be considered. For example, the scalability requirements are key when you have high volumes of data to avoid very long training times, lack of memory, etc. This is important in terms of productivity, portability, cost reduction, the minimization of staff involvement, etc. In cases where large volumes of data have been collected for many years (cancer studies, studies of discharge from the hospital and sick leaves, etc.), scalability acquires a transcendental role [30].

Application to Our Case Study

Based on the results collected in Table 4, a quite good performance of all the algorithms can be observed for most metrics. Considering the large number of variables available initially, the rigor with which all the data were obtained by the multidisciplinary team, and the adequate selection of variables carried out in previous phases (choosing the objective variables and with a greater predictive power), a suitable prediction performance was initially expected.

Considering the minimal target values of the statistical measures established in Stage 1 "Clinical target definition", several algorithms fulfil the requirements in terms of the target metrics:

- Accuracy (>85%): logistic regression, SVM, random forest, MLP neural network, and GBA.
- Precision (>85%): logistic regression, SVM, decision tree random forest, MLP neural network, and GBA.

- Recall (>90%): SVM, random forest, MLP neural network, kNN, and GBA.

According to the results, the four algorithms that fulfil the requirements are SVM, random forest, the MLP neural network, and GBA. From all these algorithms, SVM achieves the best results in all the metrics considered, and therefore it could be the model selected for production.

3.7. Stage 7: Operating Profit and Clinical Use

The last stage (Figure 4, stage 7) concerns the adaptation and organization of the acquired knowledge and the predictive capacity of the system to make it accessible to the physician [47]. At this point, it is important to emphasize the key role played by the expert medical professionals (final decision makers) in interpreting the results, in determining whether certain patterns observed make medical sense and are relevant, or in clearly distinguishing between correlation and causality (as studied for different heath topics [79,80]). The "intelligence" provided is not intended to replace the physician but to advise and guide his/her decisions, which will always prevail [20].

Application to Our Case Study

The results obtained directly by the model may involve difficulties when interpreted by the physicians. That is why the information presented to them must be intuitive, simple, and easily interpretable. In this regard, a concise graphic and clear report, where the results of the test and the prediction of the pathology/target variable to be predicted are presented, can help the clinical to more easily interpret the test performed and the results of the predictive model. This paper is accompanied by a report example as Supplementary Material File S4, showing a possible example of a report for our case study that includes patient data, test results in a tabular and graphic format, and an area showing the prediction of cervical pain obtained by the model.

Once the model enters production, it will be possible to add data regarding new patients both with or without cervical pain and verified diagnosis, which would increase the sample and thus the predictive power.

4. Discussion and Lessons Learnt

In this paper, the particularities of applying ML techniques in the field of healthcare are shown, developing all the stages that comprise it, to generate reliable and stable models. It has been exemplified through a case study of cervical pain evaluation, where we have been able to predict the presence of cervical pain with accuracy, precision, and recall above 85% with the approaches based on SVM, random forest, MLP neural networks, and GBA.

In order to clarify and structure the knowledge acquired during the development of the current study, a summary of some key aspects and lessons learnt regarding DM and ML in the field of healthcare is shown in Table 5. Every key aspect has been categorized in a general classification, followed by a description, the real situation exemplified in our case study, and some important related references.

Table 5. Summary of the key aspects and lessons learnt.

Category	Key Aspect	Description	Case Study	Sample References
Clinical target	Proper selection	Clinical target definition according to the aims and clinical needs. This facilitates the subsequent selection of data.	Presence of cervical pain. Only collaborating subjects and objective variables were selected.	[34,40]
	Definition of statistical measures	Minimum metrics to be fulfilled by the model according to the clinical target. Metrics are checked in stage 6 (interpretation of results) after the predictive model generation.	Performance required: accuracy: greater than 85%; precision: greater than 85%; recall: greater than 90%.	[39]

Table 5. *Cont.*

Category	Key Aspect	Description	Case Study	Sample References
Data	Identification and understanding	Diversity of the origins of data in healthcare (regarding typology, consistency, and veracity) and need to correctly understand the data. A health expert is required in this stage.	Prior to the field work, relevant clinical information was identified and the tests to perform were determined.	[34]
	Clear and concise structure	Categorization of data using appropriate variables/features, applying classifications motivated by medical needs. This is essential to carry out an adequate analysis of the information.	Data classification motivated by clinical staff and forensic experts: patient characterization data, assessed and reported clinical data, measured clinical data.	[30]
	Data transformations in healthcare	Feature engineering. Reducing the raw data to be handled and adapting them to the required format in order to increase the predictive power.	Variables such as the age, level of studies, weight, height, etc., were transformed into discrete variables. Ranges of variables were associated with a numeric code.	[5,19,47,52]
	Anonymization	Preservation of sensitive patient data by transforming values of data variables into scales or groups, thus avoiding patient identification. This is a key aspect in healthcare data management.	No quantitative variables remained after the anonymization process (through transformation into discrete variables and the removal of identifying attributes) that could be associated with patients.	[58]
	Selection	After data transformation, the selection of variables applying a filter according to the target:ethical and legal issues, manual selection, automated attribute selection.	The volume of data in the case study was reduced from 230 variables to 28 after applying the three successive aforementioned filters.	[19,21]
	Normalization	Normalization of quantitative attributes avoiding situations where variables that can take larger values could end up dominating others.	No quantitative variables remained after anonymization.	[55–57]
	Completeness	Completeness and quality of the data recorded as a key aspect in the health area.	Incomplete data from 5 patients related to specific features were collected through a telephone call.	[81]
	Over-parametrisation	Need not to exceed the 1 to 10 ratio between variables and data to avoid overfitting. The dimensionality is an issue in studies with a high number of variables compared to the total number of samples.	This was a real issue in our case study because of the volume of data provided by sensors. A sample of 302 and 28 variables was finally selected, which complies with the 1 to 10 ratio.	[20]

Table 5. *Cont.*

Category	Key Aspect	Description	Case Study	Sample References
	Scalability	Support for handling large amounts of data (efficient and effective collection, storage, management, and exploitation). Depending on the project duration, and especially if it is intended to have an adaptive character (projects with data collected for many years), scalability is a key issue to consider.	The current project is still in an initial stage, with no large-scale deployment. No scalability problems have been detected.	[30]
Predictive model	High recall and relatively high precision	Minimization of the number of false negatives (increasing recall). This is a key goal in healthcare, since false negatives and false positives have no similar costs in this area. The precision should also be suitable, as a high number of false positives would lead to false alarms, the performance of needless procedures, and increasing costs and discomfort for the patients.	The selected algorithms (SVM, random forest, MLP neural network, and GBA) have recall >90%.	[39]
Project work procedure	Multidisciplinary work as a key point	Composing teams involving technical people and diverse health professionals. This is required, but not always possible. Insufficient collaboration could be diminished by applying the stages assigned to each of the professionals in a concise and structured way.	There was interaction between professionals in almost all the stages. Nevertheless, more interventions could be encouraged because clinical experts were not present in stage 5.	[34,52]
	Continuous data collection	Improvement of the model performance thanks to a continuous learning process.	There is an intention to improve the current system by incorporating data of new collaborating patients.	[30]
MoCap	Sensors/devices in healthcare	Complementary objective tests to help physicians.	We expect the applicability of the proposal in the forensic field as an objective system of application to aid in judicial processes.	[19,82]

Data structuring. A clear and concise structuring of the data is essential to carry out an adequate analysis of the information as well as to make this information really useful for the purposes of the predictive model. In our case study, and prior to the field work, the relevant clinical information was identified and the tests to perform were determined (cervical ROM in three different planes with a MoCap system of inertial sensors), so that its processing and subsequent structuring were easier. It is essential to accurately structure the available information in a suitable way (using appropriate variables/features) when working with large volumes of data, and to classify the different variables in different groups (using appropriate categories) to facilitate access (for all the parties involved) in a more effective and useful way. Data management is so important that the preparation of the data covers three of the seven stages of the methodology (stages 2, 3, and 4).

Selection of variables. Likewise, the selection of the most adequate information to predict a certain characteristic, as well as its transformation in terms of variables, is essential. In relation to this adaptation, sensitive patient data must be anonymized for their use in the generation of a predictive

model [19,21]. Although converting continuous predictors to discrete variables (specifically binary variables) is not always recommended [20], the necessary transformation of data for privacy reasons in the field of health conditions the stage of data transformation. The reduction in the volume of data in the case study was from 230 variables to 28 due to the large number of variables provided by the inertial sensors and the clinical data. The variable reduction applied (based on the gain ratio of the variables) after the corresponding selection of the data for ethical and legal issues and the screening made by an expert was necessary to fulfil the 1 to 10 ratio between the variables and data. This key point allowed us to ensure that the predictive models work properly, maximizing the predictive power and avoiding overfitting.

Selection of the predictive model. Regarding the selection of the predictive model, after performing a parameter tuning of the seven selected algorithms and comparing the most suitable configuration of each of them (see Figure 5), it was concluded that the models that meet the established requirements regarding accuracy, precision, and recall for the case study were SVM, random forest, MLP neural networks, and GBA. Our results highlight the low number of false negatives achieved (high recall), a fundamental aspect in healthcare studies [39]. These results are in agreement with other investigations of a similar nature, using the same software (Weka), in terms of the accuracy, precision, recall, and F1 score with the SVM and random forest algorithms [36].

Variability of the measures acquired. It has been detected that it is possible to assess the measurement capacity of our medical equipment in terms of the variability of the measures that it obtains. A *series* variable identifies whether the data are relative to the first measure of each subject or to the second (which was performed consecutively). The results obtained by the predictive models showed that there were no differences between the two series of cervical ROM (this variable had the lowest predictive power among all the variables introduced in the model). This result indicates that the measure has behaved stably in collaborating subjects. This result is interesting in the forensic field due to the following reason. If repeating the cervical ROM test in a patient results in significant differences, they would not be derived from the variability of the test, but by the type of pathology that prevents the patient from repeating the test normally. Alternatively, the patient may try to simulate or magnify the lesion by not showing consistent results between the first and second series. This aspect would be of relevance to judicial experts [29,83].

Data collection in production. Once the system is applied in its context and has been developed, continuous data collection must be planned in production in order to improve the prediction accuracy, resulting in a continuous learning system (Figure 3). In the case study, to increase the sample in the exploitation stage of the model it is possible to include those patients who perform cervical assessment tests in a care or rehabilitation setting; thus, their sincerity can be assumed regarding the degree of cervical pain as well as full collaboration in the performance of the tests. However, for the collection of training data it may be necessary to exclude those patients who are immersed in a judicial and indemnifying process and who report cervical pain because their degree of collaboration or sincerity is unknown. In the future, the system can be used to predict the cervical pain of non-collaborating patients (e.g., patients in a judicial process or patients with a high degree of anxiety or hypochondriacs) from the predictive model previously generated with collaborating patients, serving as objective evidence in judicial proceedings with insurance companies [40].

Multidisciplinarity. For a correct interpretation of the results, multidisciplinary work is a key point, since the contribution of each of the branches of knowledge is necessary in this type of project to optimize the possibilities offered by the model. In this way, it will be possible to assess the statistical quality of the results and their medical utility. For example, in our case study questions were raised regarding the way the data and the results should be represented (solved with the databases defined and the design of a report for physicians, presented as Supplementary Material), the possible interpretation and use of the system by the clinician (problems could have arisen if we had not been worked in close collaboration with the clinicians; however, the target variable and how to present the results were clarified from the beginning of the project), and the overlap with other possible

decision-support systems (if other systems could also provide in the future a prediction indicating that a patient suffers cervical pain, both results would be presented to the clinician and he/she would take the final decision). Through collaboration with the medical experts, these issues have been solved. However, multidisciplinary work is not always possible, since professionals participate in different stages of the entire process according to their degree of knowledge, experience, and training, so in some cases there may be no direct or sufficient interaction between them [34]. This lack of interaction among professionals could be diminished by following the stages assigned to each of the professionals in a concise and structured way, thus avoiding problems that may lead to project failure [52]. In the case study, the rigor followed by the different professionals involved in the different stages have resulted in adequate results. Nevertheless, although there has been interaction between them, it could have been done in a more collaborative way, since clinical experts were not present in the generation of the predictive model and the interpretation of the results, which could have improved the quality of the study thanks to its specific medical knowledge.

Exploitation of sensor data. The use of sensors and devices with the use of ML could be implemented as a complementary objective test to help physicians. This type of test could constitute an aid to the decision-making in the diagnosis or treatment, or if there is doubt about the veracity of the information reported by the patient [19,82]. Although we wanted to show the clinical utility of this type of technology, the lack of studies on the application of ML techniques with motion capture sensors in healthcare, and specifically their applicability in the forensic field as an objective system of application in judicial processes, have further motivated the choice of the case study. Therefore, while our case study focuses on the medical legal/forensic field, the procedure proposed to use ML techniques could be applied in any study of the health field (cancer detection, studies of discharge from the hospital and sick leave, etc.).

Use of resources. Regarding the frequency and regularity of data acquisition, it is necessary to previously estimate it to limit the duration of the project [34], as well as to quantify the necessary storage size, which is a factor with high variability between projects. If the project is intended to have an adaptive character that can be applied or expanded for subsequent research, the scalability and magnitude should be considered. If the scalability of the project is not considered, and the intention is to continue acquiring data and adapting the model to the continuous growth of information, there may come a time when the project is no longer viable because it is unable to assimilate the corresponding increase in data size. This situation is common in epidemiological projects of data collection for large periods of time, where the scalability is transcendental for the future of the project [30]. On the other hand, it is important to consider that, in certain projects as in the case study presented in this paper, the ratio between available variables and data can be high. So, not exceeding the 1 to 10 ratio between variables and data is transcendental to avoid overfitting effects [20].

5. Conclusions and Future Work

Through a practical guide, the stages and particularities to consider for the application of ML techniques in the field of healthcare have been described, considering all the stages involved in the process. This procedure is shown through objective cervical functional assessment tests that use MoCap technology with inertial sensors and a predictive model whose goal is to estimate the presence of cervical pain from the data collected with the test. Four models (SVM, random forest, MLP neural network, and GBA) from the seven models initially generated obtained an accuracy and precision of more than 85% and a recall of more than 90% (i.e., the percentage of false negatives is smaller than 10%). The approach and the results obtained could help objectify diagnoses, improve test treatment efficacy, and save resources in healthcare systems. The procedure, which has been applied to data derived from a cervical assessment study for verification and evaluation, is also appropriate for any healthcare field regardless of the origin of the data. It can be useful, for example, in gait studies [82], balance studies [84], cardiac failure studies [85], the prediction of events [86], fertility tests [87], etc.

Despite the great usefulness of ML in the field of healthcare, some limitations have been detected in this field. First, a major limitation is how to achieve a suitable flow of data from health centers and hospitals, as well as the accessibility (in relation to privacy policies and authorizations) [21], gathering, and integration of the data [3]. If a global information collaboration policy were established between hospitals [6,69,88], the problem of access to information could be solved, and it would be possible to share more data and feed the predictive models applied to the field of healthcare more efficiently. The explainability of predictions [89,90] is an important issue in a health care domain, as it could increase the trust in the ML systems (for both clinicians and patients) and even lead to the acquisition of new knowledge; however, more research on how to achieve explainability while considering the potential trade-off with accuracy must be performed, especially in the health domain.

Regarding the limitations of the conclusions obtained with this case study, once the model is in the exploitation stage it would be advisable to carry out an external validation to verify the viability of the model in terms of geography, temporality, etc. [30,77]. Regarding the data sample used in our case study, its size has been large enough to obtain good results, but it would be relevant to see the impact of increasing it, as new data about patients becomes available, to enable a continuous learning process that could lead to better results over time. Besides, a study is currently being conducted to check the accuracy of the proposed models with a sample of non-collaborating patients.

Concerning the target variable (presence of cervical pain), which is a binary variable, it could be defined in a more granular way considering not only the presence of pain but also its intensity (as a continuous variable or as a discrete variable with several pain degrees). The problem with pain intensity is that pain scales are highly dependent on the subjectivity of the patient, and this issue could be further exacerbated with non-collaborating subjects. However, as a future goal of our research, it could be useful to tackle this issue and introduce some statistical techniques, such as the numerical measurement z-score, to normalize the subjective values from pain intensity scales (e.g., the Visual Analogue Scale) provided by the patients. The z-score could help to reduce the bias of patients, allowing us to include pain intensity as a target variable in our proposal.

Finally, as future work, it could also be interesting to extend the range of experiments performed and analyze the potential interest of other ML methods; for example, we could consider applying different classifiers applied over different categories of data proposed in the paper and combine them into an ensemble.

Supplementary Materials: The following materials are available online at http://www.mdpi.com/2076-3417/10/17/5942/s1, File S1: Initial database. File S2: Coded database. File S3: Filtered database. File S4: Report for physician.

Author Contributions: Conceptualization, J.d.l.T., J.M., S.I., and J.J.M.; data curation, J.d.l.T., J.M., and J.J.M.; formal analysis, J.d.l.T., J.M., S.I., and J.J.M.; funding acquisition, S.I. and J.J.M.; investigation, J.d.l.T.; methodology, J.d.l.T., J.M., and J.J.M.; project administration, J.J.M.; resources, S.I. and J.J.M.; software, J.d.l.T. and J.J.M.; supervision, S.I. and J.J.M.; validation, J.d.l.T.; visualization, J.d.l.T.; writing—original draft preparation, J.d.l.T., J.M., and J.J.M.; writing—review and editing, J.d.l.T., J.M., S.I., and J.J.M. All authors have read and agreed to the published version of the manuscript.

Funding: This project was co-funded by the Government of Aragon, the European Regional Development Fund and the University of Zaragoza. We thank the support of the project TIN2016-78011-C4-3-R (AEI/FEDER, UE) and the Government of Aragon (Group Reference T64_20R, COSMOS research group).

Acknowledgments: We thank Mapfre insurance company (Zaragoza, Spain) for allowing us to use their facilities. We also thank the I3A—University Institute of Research of Engineering of Aragon, University of Zaragoza, Zaragoza, Spain—for the materials and support they provided.

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

Abbreviations

The following abbreviations are used in this manuscript:

BMI	Body Mass Index
CEICA	Bioethics Committee of Aragón
CRISP-DM	CRoss-Industry Standard Process for Data Mining
DLA	Daily life activities
DM	Data Mining
EMG	Surface Electromyography
GBA	Gradient Boosting Algorithm
KDD	Knowledge Discovery in Databases
KNN	K-Nearest Neighbors
MAE	Mean Absolute Error
ML	Machine Learning
MLP	MultiLayer Perceptron
MoCap	Motion Capture
PCA	Principal component analysis
RMSE	Root Mean Squared Error
ROM	Range of Movement
SEMMA	Sample, Explore, Modify, Model, and Assess
SVM	Support Vector Machine
WDQ	Whiplash Scale

References

1. Kayyali, B.; Knott, D.; Van Kuiken, S. *The Big-Data Revolution in US Health Care: Accelerating Value and Innovation*; Mc Kinsey Co.: New York, NY, USA, 2013; Volume 2, pp. 1–13.
2. Tomar, D.; Agarwal, S. A survey on Data Mining approaches for Healthcare. *Int. J. Bio-Sci. Bio-Technol.* **2013**, *5*, 241–266. [CrossRef]
3. Koh, H.C.; Tan, G. Data mining applications in healthcare. *J. Healthc. Inf. Manag.* **2011**, *19*, 65.
4. Maity, N.G.; Das, S. Machine learning for improved diagnosis and prognosis in healthcare. In Proceedings of the 2017 IEEE Aerospace Conference, Big Sky, MT, USA, 4–11 March 2017; pp. 1–9.
5. Yoo, I.; Alafaireet, P.; Marinov, M.; Pena-Hernandez, K.; Gopidi, R.; Chang, J.-F.; Hua, L. Data Mining in Healthcare and Biomedicine: A Survey of the Literature. *J. Med. Syst.* **2012**, *36*, 2431–2448. [CrossRef]
6. Sen, I.; Khandelwal, K. Data Mining in Healthcare. 2018. Available online: https://www.researchgate.net/publication/322754945_DATA_MINING_IN_HEALTHCARE (accessed on 26 August 2020).
7. Clavel, D.; Mahulea, C.; Albareda, J.; Silva, M. A Decision Support System for Elective Surgery Scheduling under Uncertain Durations. *Appl. Sci.* **2020**, *10*, 1937. [CrossRef]
8. Cruz, J.A.; Wishart, D.S. Applications of Machine Learning in Cancer Prediction and Prognosis. *Cancer Inform.* **2006**, *2*, 59–77. [CrossRef]
9. Wang, F.; Stiglic, G.; Obradovic, Z.; Davidson, I. Guest editorial: Special issue on data mining for medicine and healthcare. *Data Min. Knowl. Discov.* **2015**, *29*, 867–870. [CrossRef]
10. Rosales, R.E.; Rao, R.B. Guest Editorial: Special Issue on impacting patient care by mining medical data. *Data Min. Knowl. Discov.* **2010**, *20*, 325–327. [CrossRef]
11. Alotaibi, S.; Mehmood, R.; Katib, I.; Rana, O.; Albeshri, A. Sehaa: A Big Data Analytics Tool for Healthcare Symptoms and Diseases Detection Using Twitter, Apache Spark, and Machine Learning. *Appl. Sci.* **2020**, *10*, 1398. [CrossRef]
12. Huang, Z.; Dong, W.; Bath, P.; Ji, L.; Duan, H. On mining latent treatment patterns from electronic medical records. *Data Min. Knowl. Discov.* **2015**, *29*, 914–949. [CrossRef]
13. Bentham, J.; Hand, D.J. Data mining from a patient safety database: The lessons learned. *Data Min. Knowl. Discov.* **2012**, *24*, 195–217. [CrossRef]
14. Obenshain, M.K. Application of Data Mining Techniques to Healthcare Data. *Infect. Control. Hosp. Epidemiol.* **2004**, *25*, 690–695. [CrossRef] [PubMed]
15. Zhang, P.; Wang, F.; Hu, J.; Sorrentino, R. Towards Personalized Medicine: Leveraging Patient Similarity and Drug Similarity Analytics. *AMIA Summits Transl. Sci. Proc.* **2014**, *2014*, 132. [PubMed]
16. Hamet, P.; Tremblay, J. Artificial intelligence in medicine. *Metab. Clin. Exp.* **2017**, *69*, S36–S40. [CrossRef] [PubMed]

17. Joyner, M.J.; Paneth, N. Seven Questions for Personalized Medicine. *JAMA* **2015**, *314*, 999–1000. [CrossRef]
18. Weiss, J.C.; Natarajan, S.; Peissig, P.L.; McCarty, C.A.; Page, D. Machine Learning for Personalized Medicine: Predicting Primary Myocardial Infarction from Electronic Health Records. *AI Mag.* **2012**, *33*, 33–45. [CrossRef]
19. Mannini, A.; Sabatini, A.M. Machine Learning Methods for Classifying Human Physical Activity from On-Body Accelerometers. *Sensors* **2010**, *10*, 1154–1175. [CrossRef]
20. Moons, K.; Kengne, A.P.; Woodward, M.; Royston, P.; Vergouwe, Y.; Altman, U.G.; Grobbee, D.E. Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. *Heart* **2012**, *98*, 683–690. [CrossRef]
21. Wilkowska, W.; Ziefle, M. Privacy and data security in E-health: Requirements from the user's perspective. *Heal. Inform. J.* **2012**, *18*, 191–201. [CrossRef]
22. Dolley, S. Big Data Solution to Harnessing Unstructured Data in Healthcare. IBM Report. 2015. Available online: https://assets.sourcemedia.com/31/a6/cb1b019c4d6cb338fab539eea360/ims14428usen.pdf. (accessed on 26 August 2020).
23. Andersen, R.M.; Davidson, P.L.; Baumeister, S.E. Improving access to care. In *Changing the US Health Care System: Key Issues in Health Services Policy and Management*; John Wiley & Sons: Hoboken, NJ, USA, 2013; pp. 33–70.
24. Marin, J.; Blanco, T.; Marin, J.J. Research Lines to Improve Access to Health Instrumentation Design. *Procedia Comput. Sci.* **2017**, *113*, 641–646. [CrossRef]
25. Cassidy, J.D.; Carroll, L.; Côté, P.; Lemstra, M.; Berglund, A.; Nygren, Å. Effect of Eliminating Compensation for Pain and Suffering on the Outcome of Insurance Claims for Whiplash Injury. *N. Engl. J. Med.* **2000**, *342*, 1179–1186. [CrossRef]
26. Moreno, A.J.; Utrilla, G.; Marin, J.; Marin, J.J.; Sanchez-Valverde, M.B.; Royo, A.C. Cervical Spine Assessment with Motion Capture and Passive Mobilization. *J. Chiropr. Med.* **2018**, *17*, 167–181. [CrossRef] [PubMed]
27. Utrilla, G.; Marín, J.J.; Sanchez-Valverde, B.; Gomez, V.; JAuria, J.M.; Marin, J.; Royo, C. *Cervical Mobility Testing in Flexion-Extension and Protraction-Retraction to Evaluate Whiplash Syndrome Through Motion Capture*; Universidad de Zaragoza: Zaragoza, Spain, 2017.
28. Marín, J.J.; Pina, M.J.B.; Gil, C.B. Evaluación de Riesgos de Manipulación Repetitiva a Alta Frecuencia Basada en Análisis de Esfuerzos Dinámicos en las Articulaciones sobre Modelos Humanos Digitales. *Cienc. Trab.* **2013**, *15*, 86–93. [CrossRef]
29. Marín, J.; Boné, M.; Ros, R.; Martínez, M. Move-Human Sensors: Sistema portátil de captura de movimiento humano basado en sensores inerciales, para el análisis de lesiones musculoesqueléticas y utilizable en entornos reales. In Proceedings of the Sixth International Conference on Occupational Risk Prevention, Galicia, Spain, 14–16 May 2008.
30. Steyerberg, E.W. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*; Springer: Berlin, Germany, 2009.
31. Fayyad, U.M.; Piatetsky-Shapiro, G.; Smyth, P.; Uthurusamy, R. *Advances in Knowledge Discovery and Data Mining*; AAAI Press: Menlo Park, CA, USA, 1996.
32. Azevedo, A.I.R.L.; Santos, M.F. KDD, SEMMA and CRISP-DM: A Parallel Overview. ISCAP—Informática—Comunicações em Eventos Científicos. 2008. Available online: https://recipp.ipp.pt/handle/10400.22/136 (accessed on 26 August 2020).
33. Chapman, P.; Clinton, J.; Kerber, R.; Khabaza, T.; Reinartz, T.; Shearer, C.; Wirth, R. *CRISP-DM 1.0 Step-by-Step Data Mining Guide*; SPSS Inc.: Chicago, IL, USA, 2000.
34. McGregor, C.; Catley, C.; James, A. A process mining driven framework for clinical guideline improvement in critical care. In Proceedings of the Learning from Medical Data Streams Workshop, Bled, Slovenia, 6 July 2011.
35. Catley, C.; Smith, K.; McGregor, C.; Tracy, M. Extending CRISP-DM to incorporate temporal data mining of multidimensional medical data streams: A neonatal intensive care unit case study. In Proceedings of the 2009 22nd IEEE International Symposium on Computer-Based Medical Systems, Albuquerque, NM, USA, 3–4 August 2009; pp. 1–5.
36. Araujo, F.H.; Santana, A.M.; Neto, P.S. Using machine learning to support healthcare professionals in making preauthorisation decisions. *Int. J. Med. Inform.* **2016**, *94*, 1–7. [CrossRef]
37. Bose, I.; Mahapatra, R.K. Business data mining—A machine learning perspective. *Inf. Manag.* **2001**, *39*, 211–225. [CrossRef]

38. Bhatla, N.; Jyoti, K. An analysis of heart disease prediction using different data mining techniques. *Int. J. Eng.* **2012**, *1*, 1–4.
39. Raschka, S.; Mirjalili, V. *Python Machine Learning*; Packt Publishing Ltd.: Birmingham, UK, 2017.
40. Schuller, E.; Eisenmenger, W.; Beier, G. Whiplash Injury in Low Speed Car Accidents: Assessment of Biomechanical Cervical Spine Loading and Injury Prevention in a Forensic Sample. *J. Musculoskelet. Pain* **2000**, *8*, 55–67. [CrossRef]
41. Naumann, F. Data profiling revisited. *ACM SIGMOD Rec.* **2014**, *42*, 40–49. [CrossRef]
42. Rahm, E.D.H. Data Cleaning: Problems and Current Approaches. *Bull. Tech. Comm. Data Eng.* **2000**, *23*, 3–13.
43. Jannot, A.-S.; Zapletal, E.; Avillach, P.; Mamzer, M.-F.; Burgun, A.; Degoulet, P. The Georges Pompidou University Hospital Clinical Data Warehouse: A 8-years follow-up experience. *Int. J. Med. Inform.* **2017**, *102*, 21–28. [CrossRef]
44. Evans, R.S.; Lloyd, J.F.; Pierce, L.A. Clinical Use of an Enterprise Data Warehouse. *AMIA Annu. Symp. Proc.* **2012**, *2012*, 189–198.
45. Batista, G.E.A.P.A.; Prati, R.C.; Monard, M.C. A study of the behavior of several methods for balancing machine learning training data. *ACM SIGKDD Explor. Newsl.* **2004**, *6*, 20–29. [CrossRef]
46. Batista, G.E.A.P.A.; Prati, R.C.; Monard, M.C. Balancing Strategies and Class Overlapping. *Lect. Notes Comput. Sci.* **2005**, *3646*, 24–35. [CrossRef]
47. Bhardwaj, R.; Nambiar, A.R.; Dutta, D. A Study of Machine Learning in Healthcare. *2017 IEEE 41st Annu. Comput. Softw. Appl. Conf.* **2017**, *2*, 236–241. [CrossRef]
48. Dörre, J.; Gerstl, P.; Seiffert, R. Text mining: Finding nuggets in mountains of textual data. In Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), San Diego, CA, USA, 23–27 August 1999; pp. 398–401.
49. Aggarwal, C.C.; Zhai, C. *Mining Text Data*; Springer Science & Business Media: Berlin, Germany, 2012.
50. Janasik, N.; Honkela, T.; Bruun, H. Text mining in qualitative research: Application of an unsupervised learning method. *Organ. Res. Methods* **2009**, *12*, 436–460. [CrossRef]
51. Blei, D.M. Probabilistic topic models. *Commun. ACM* **2012**, *55*, 77–84. [CrossRef]
52. Henri, L. *Data Scientist y Lenguaje R Guía de Autoformación Para el uso de Big Data*; Francisco, J., Piqueres, J., Eds.; Colecciones Epsilon: Cornellá de Llobregat, Spain, 2017.
53. Stavrianou, A.; Andritsos, P.; Nicoloyannis, N. Overview and semantic issues of text mining. *ACM SIGMOD Rec.* **2007**, *36*, 23–34. [CrossRef]
54. Carlos, T.; Sergio, I.; Carlos, S. Text Mining of Medical Documents in Spanish: Semantic Annotation and Detection of Recommendations. In Proceedings of the 16th International Conference on Web Information Systems and Technologies (WEBIST 2020), Budapest, Hungary, 3–5 November 2020.
55. Donders, A.R.T.; Van Der Heijden, G.J.; Stijnen, T.; Moons, K.G. Review: A gentle introduction to imputation of missing values. *J. Clin. Epidemiol.* **2006**, *59*, 1087–1091. [CrossRef]
56. Farhangfar, A.; Kurgan, L.; Dy, J. Impact of imputation of missing values on classification error for discrete data. *Pattern Recognit.* **2008**, *41*, 3692–3705. [CrossRef]
57. Bennett, D.A. How can I deal with missing data in my study? *Aust. N. Z. J. Public Health* **2001**, *25*, 464–469. [CrossRef]
58. Arbuckle, L.; El Emam, K. *Anonymizing Health Data*; O'Reilly Media, Inc.: Newton, MA, USA, 2013.
59. Kargupta, H.; Datta, S.; Wang, Q.; Sivakumar, K. On the privacy preserving properties of random data perturbation techniques. In Proceedings of the Third IEEE International Conference on Data Mining, Melbourne, FL, USA, 19–22 November 2003; pp. 99–106.
60. El Emam, K.; Dankar, F.; Issa, R.; Jonker, E.; Amyot, D.; Cogo, E.; Corriveau, J.-P.; Walker, M.; Chowdhury, S.; Vaillancourt, R.; et al. A globally optimal k-anonymity method for the de-identification of health data. *J. Am. Med. Inform. Assoc.* **2009**, *16*, 670–682. [CrossRef]
61. Dankar, F.K.; El Emam, K. The application of differential privacy to health data. *Jt. EDBT/ICDT Workshops EDBT-ICDT* **2012**, *2012*, 158–166. [CrossRef]
62. IBM. SPSS Software. Available online: https://www.routledge.com/IBM-SPSS-Statistics-26-Step-by-Step-A-Simple-Guide-and-Reference/George-Mallery/p/book/9780367174354 (accessed on 16 February 2020).
63. Wu, X.; Kumar, V.; Quinlan, J.R.; Ghosh, J.; Yang, Q.; Motoda, H.; McLachlan, G.J.; Ng, A.; Liu, B.; Yu, P.; et al. Top 10 algorithms in data mining. *Knowl. Inf. Syst.* **2007**, *14*, 1–37. [CrossRef]

64. Gupta, P. Cross Validation in Machine Learning. 2020. Available online: https://towardsdatascience.com/cross-validation-in-machine-learning-72924a69872f (accessed on 26 August 2020).
65. Arlot, S.; Celisse, A. A survey of cross-validation procedures for model selection. *Stat. Surv.* **2010**, *4*, 40–79. [CrossRef]
66. Refaeilzadeh, P.; Tang, L.; Liu, H. Cross-Validation. *Encycl. Database Syst.* **2009**, *5*, 532–538.
67. Xu, Y.; Goodacre, R. On Splitting Training and Validation Set: A Comparative Study of Cross-Validation, Bootstrap and Systematic Sampling for Estimating the Generalization Performance of Supervised Learning. *J. Anal. Test.* **2018**, *2*, 249–262. [CrossRef] [PubMed]
68. McCaffrey, J. Neural Network Train-Validate-Test Stopping. 2020. Available online: https://visualstudiomagazine.com/articles/2015/05/01/train-validate-test-stopping.aspx (accessed on 26 August 2020).
69. Ferber, R.; Osis, S.T.; Hicks, J.L.; Delp, S.L. Gait biomechanics in the era of data science. *J. Biomech.* **2016**, *49*, 3759–3761. [CrossRef] [PubMed]
70. Reed, R.; MarksII, R.J. *Neural Smithing: Supervised Learning in Feedforward Artificial Neural Networks*; Mit Press: Cambridge, MA, USA, 1999.
71. Christodoulou, E.; Ma, J.; Collins, G.S.; Steyerberg, E.W.; Verbakel, J.Y.; Van Calster, B.; Evangelia, C.; Jie, M. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J. Clin. Epidemiol.* **2019**, *110*, 12–22. [CrossRef]
72. Bae, J.M. The clinical decision analysis using decision tree. *Epidemiol. Health* **2014**, *36*. [CrossRef] [PubMed]
73. Noi, P.T.; Kappas, M. Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery. *Sensors* **2018**, *18*, 18. [CrossRef]
74. Penny, W.; Frost, D.; Penny, W. Neural Networks in Clinical Medicine. *Med. Decis. Mak.* **1996**, *16*, 386–398. [CrossRef]
75. Zhang, Z.; Zhao, Y.; Canes, A.; Steinberg, D.; Lyashevska, O.; AME Big-Data Clinical Trial Collaborative Group. Predictive analytics with gradient boosting in clinical medicine. *Ann. Transl. Med.* **2019**, *7*. [CrossRef]
76. Witten, I.; Frank, E.; Hall, M.; Pal, C. Appendix B: The WEKA workbench. In *Data Mining: Practical Machine Learning Tools and Techniques*, 4th ed.; Morgan Kaufmann: Burlington, MA, USA, 2016.
77. Moons, K.; Kengne, A.P.; Grobbee, D.E.; Royston, P.; Vergouwe, Y.; Altman, U.G.; Woodward, M. Risk prediction models: II. External validation, model updating, and impact assessment. *Heart* **2012**, *98*, 691–698. [CrossRef] [PubMed]
78. Murphy, C.K. Identifying Diagnostic Errors with Induced Decision Trees. *Med. Decis. Mak.* **2001**, *21*, 368–375. [CrossRef] [PubMed]
79. Zhao, L. The gut microbiota and obesity: From correlation to causality. *Nat. Rev. Genet.* **2013**, *11*, 639–647. [CrossRef] [PubMed]
80. Dab, W.; Ségala, C.; Dor, F.; Festy, B.; Lameloise, P.; Le Moullec, Y.; Le Tertre, A.; Médina, S.; Quenel, P.; Wallaert, B.; et al. Air pollution and health: Correlation or causality? The case of the relationship between exposure to particles and cardiopulmonary mortality. *J. Air Waste Manag. Assoc.* **2001**, *51*, 220–235. [CrossRef] [PubMed]
81. Liu, C.; Talaei-Khoei, A.; Zowghi, D.; Daniel, J. Data completeness in healthcare: A literature survey. *Pac. Asia J. Assoc. Inf. Syst.* **2017**, *9*, 5. [CrossRef]
82. Mannini, A.; Trojaniello, D.; Cereatti, A.; Sabatini, A.M. A Machine Learning Framework for Gait Classification Using Inertial Sensors: Application to Elderly, Post-Stroke and Huntington's Disease Patients. *Sensors* **2016**, *16*, 134. [CrossRef]
83. Ramírez, P.C.; Ordi, H.G.; Fernández, P.S.; Morales, M.I.C. Detección de exageración de síntomas en esguince cervical: Pacientes clínicos versus sujetos análogos. *Trauma* **2014**, *25*, 4–10.
84. De La Torre, J.; Marin, J.; Marin, J.J.; Auria, J.M.; Sanchez-Valverde, M.B. Balance study in asymptomatic subjects: Determination of significant variables and reference patterns to improve clinical application. *J. Biomech.* **2017**, *65*, 161–168. [CrossRef]
85. Austin, P.C.; Tu, J.V.; Ho, J.E.; Levy, D.; Lee, D.S. Using methods from the data-mining and machine-learning literature for disease classification and prediction: A case study examining classification of heart failure subtypes. *J. Clin. Epidemiol.* **2013**, *66*, 398–407. [CrossRef]
86. Choi, E.; Bahadori, M.T.; Schuetz, A.; Stewart, W.F.; Sun, J. Doctor AI: Predicting Clinical Events via Recurrent Neural Networks. *JMLR Work. Conf. Proc.* **2016**, *56*, 301–318.

87. Uyar, A.; Bener, A.; Ciray, H.N. Predictive modeling of implantation outcome in an in vitro fertilization setting: An application of machine learning methods. *Med. Decis. Mak.* **2015**, *35*, 714–725. [CrossRef] [PubMed]
88. Abdelaziz, A.; Elhoseny, M.; Salama, A.S.; Riad, A. A machine learning model for improving healthcare services on cloud computing environment. *Measurement* **2018**, *119*, 117–128. [CrossRef]
89. Hall, P.; Gill, N. *An Introduction to Machine Learning Interpretability*; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2018.
90. Guidotti, R.; Monreale, A.; Ruggieri, S.; Turini, F.; Giannotti, F.; Pedreschi, D. A Survey of Methods for Explaining Black Box Models. *ACM Comput. Surv.* **2018**, *51*, 1–42. [CrossRef]

4.5. ARTÍCULO 5

De la Torre, J.; Bonnet, V.; Mauti, R.; Chabaud, P.; Robert, T. Implementing a test to assess reaction, attention and inhibition capacity in elderly. *Computer Methods in Biomechanics and Biomedical Engineering*, 2020, vol. 23 (sup1), S294-S296, doi: 10.1080/10255842.2020.1815313.

COMPUTER METHODS IN BIOMECHANICS AND BIOMEDICAL ENGINEERING

ISSN: 1476-8259; 1025-5842

Factor de Impacto SJR (2019):

0.45

Áreas Temáticas:

Bioengineering: Q3

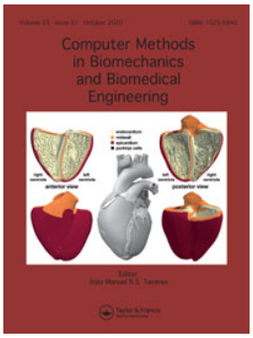
Biomedical Engineering: Q3

Computer Sciences Applications: Q2

Human-Computer Interaction: Q2

Medicine (miscellaneous): Q2

El artículo 5 es fruto de una estancia de investigación de cuatro meses en el *Laboratoire de biomécanique et mécanique des chocs* (Université Gustave Eiffel) en Lyon (Francia).



Taylor & Francis
Taylor & Francis Group

Computer Methods in Biomechanics and Biomedical Engineering

ISSN: 1025-5842 (Print) 1476-8259 (Online) Journal homepage: https://www.tandfonline.com/loi/gcmb20

Implementing a test to assess reaction, attention and inhibition capacity in elderly

Juan De la Torre, Vincent Bonnet, Rafael Mauti, Pascal Chabaud & Thomas Robert

To cite this article: Juan De la Torre, Vincent Bonnet, Rafael Mauti, Pascal Chabaud & Thomas Robert (2020) Implementing a test to assess reaction, attention and inhibition capacity in elderly, Computer Methods in Biomechanics and Biomedical Engineering, 23:sup1, S294-S296

To link to this article: https://doi.org/10.1080/10255842.2020.1815313

Published online: 02 Nov 2020.

COMPUTER METHODS IN BIOMECHANICS AND BIOMEDICAL ENGINEERING
2020, VOL. 23, NO. S1, S294–S296
https://doi.org/10.1080/10255842.2020.1815313

Implementing a test to assess reaction, attention and inhibition capacity in elderly

Juan De la Torre[a,b], Vincent Bonnet[c], Rafael Mauti[d], Pascal Chabaud[d] and Thomas Robert[b]

[a]IDERGO - Research and Development in Ergonomics, University of Zaragoza, Zaragoza, Spain; [b]Univ Lyon, Univ Gustave Eiffel, LBMC UMR_T9406, Lyon, France; [c]LISSI, Univ. of Paris-Est Créteil, France; [d]Univ Lyon, Univ Claude-Bernard Lyon 1, Villeurbanne, France

1. Introduction

Falls and fall-related injuries correspond to major, public health problems, especially among the elderly (Tinetti 2003). These injuries are a considerable burden on public health-care budgets in many western countries increasing every year due to ageing population (Heinrich et al. 2010).

In order to avoid a fall, it requires perception of a postural threat, selection of an appropriate corrective response, and proper response execution. The individual physiological components required to avoid falls—sensory acuity, reaction time, and reactive stepping—have been previously investigated.

The choice stepping reaction time test (CSRT) is an integrated test that is a good indicator of the risk of fall and has correlation with impaired cognitive functions (Lord and Fitzpatrick 2001). However, its integrated aspect does not allow to get insight into the specific deficits to precisely define the subjects profile with risk of fall (Pijnappels et al. 2010).

Therefore, the aim of this work was to complement the CSRT with additional tasks to better assess in elderly people perception and response execution, attention and inhibition capacity. For this purpose, a test with different tasks is proposed. It was tested in two different samples, young and elderly, in order to determinate differences between them and anticipate relevant conclusions.

2. Methods

2.1. Participants

A test, which consists of four different tasks to measure stepping time (ST: time from the appearance of a stimulus until step on a target), has been conducted on two samples of healthy subjects (they could walk without external help and exercised at least 30 min/day). Prior the test, subjects signed a form, consenting to undergo and understand the tests. The research was approved by Univ. Eiffel's ethical committee and complied the ethical standards of the Declaration of Helsinki. Young group (YG), consisted on eleven young subjects 7 males and 4 females (age 24.91 ± 2.77 y.; height 173.55 ± 9.55 cm; dominant hand side: 10 R, 1 L; dominant leg side: 8 R, 3 L). Elderly group (EG), consisted on nine elderly subjects 3 males and 6 females (age 79.89 ± 3.37 y.; height 163.67 ± 8.65 cm; dominant hand side: 8 R, 1 L; dominant leg side: 6 R, 3 L).

2.2. Instrumentation

The device used was a multimodal force platform of postural analysis, which was designed and manufactured in the LISSI (Figure 1(a)). It comprises six low cost pressure sensors, made with a plastic sheet whose electrical resistance varies with pressure. There are four stepped-on areas as targets in the following directions: forward (F), backward (B), rightward (R) and leftward (L); the feet are placed in the remaining two areas where it is detected whether a foot has been lifted. Pressure changes are read by an Arduino (programmed in C language), which measures the ST and the pressure. The graphical interface, which manages the stimulus display, the data recorded, and task implementation are programmed in Java. A repeatability study was carried out (Klotz 2019) between our device and a gold-standard force platform, with ST bias of 26 ms.

2.3. Experimental procedure

The four tasks have been designed based on the literature (Lord and Fitzpatrick 2001; Yechiam et al. 2006; Davranche et al. 2009). In all tasks the subject was required to step as quickly as possible on a designated target, without anticipating stimulus

Figure 1. (a) Multimodal platform. (b) SRT, CSRT and Go – No Go interface. (c) Flanker interface.

appearance. The step foot is a decision of the subject (except in R and L, which had to be with the corresponding leg). Subjects performed the test barefoot. A training trial before each task was carried out, which consisted of one repetition per direction and condition. The task consisted on four repetitions per direction and condition, which were the following:

Simple Reaction Time (SRT): the subject was aware of the target direction to be stepped on. Three trials in each direction were performed one after the other, with the following order: F, B, lateral (depends on the dominant leg). Each direction is represented on screen by an empty square in that direction, and the stimuli were displayed in them as a green signal (Figure 1(b)). This test primarily assesses the perception and response execution capacities.

Choice Stepping Reaction Time (CSRT): similar task to SRT, but the targets were randomized (subjects were not aware of the target to be stepped on). CRST requires additional resources than the SRT, including attention and decisional processes essentially.

Go/No Go: new stimulus introduced in this task, which is a red signal that indicates not to move. Directions and stimulus type are randomized. This task is more complex than the CSRT (two information – direction and colour of the target – have to be perceived and processed) and emphasized inhibitory control.

Flanker (FL): this task introduces congruent and non-congruent stimuli. It has a different display: colour signals are changed by five arrows displayed in the central part of the screen (Figure 1(c)). The central arrow indicates the target, while the side arrows are used to present a congruent or non-congruent information (if they are in the same or in a different direction as the central arrow). Directions and stimulus type (congruent/non-congruent) are randomized. It mainly assesses selective attention and perceptual inhibition.

2.4. Data analysis

A descriptive analysis (mean and SD) of ST for each task and group (YG and EG) has been carried out. Outliers have been removed and not considered in the analysis. A one way repeated-measure analysis of variance (RMAnova) on ST with factor task was carried out, as well as Tukey paired comparison between tasks of each group. Possible differences between groups, for the same task, have been analysed individually using a t-test.

3. Results and discussion

Descriptive and statistical analysis results for YG and EG are shown in Figure 2.

As expected, EG shows significantly higher ST values than YG in all tasks, due to decline in cognitive and functional capacities with age. Likewise, both groups show higher ST values as the difficulty of the tasks increases (Lord and Fitzpatrick 2001). RMAnova, for both groups, shows significant differences in a ST collective comparison.

As expected, YG and EG were very well discriminated by the CSRT. Interestingly, about 2/3 of the ST difference in CSRT could be explained by differences in the SRT, i.e., by the capacity to step slowly on a target due to a motor slowdown due to the decrease in muscle mass. Nevertheless, the remaining can be attributed to decreased cognitive capacities in EG (Lord and Fitzpatrick 2001).

Go/No Go task includes only minor differences in ST for both YG and EG. In EG, this difference was even non-significant in comparison with CSRT (p-value = .884). In addition, no mistakes have been recorded in both groups. It indicates that the task difficulty should be increased to make it more relevant (e.g., increasing the proportion of No Go trials).

Comparison between congruent and no-congruent conditions has been carried out resulting in significant differences in both groups. However, mean differences between types of stimuli are similar in both groups, which

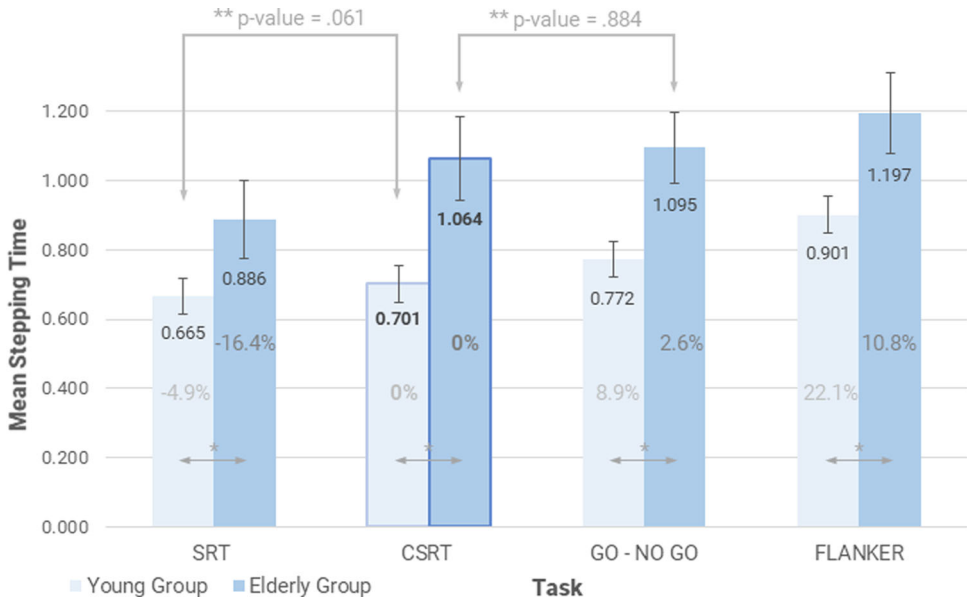


Figure 2. Descriptive and statistical analysis. *Significant differences in t-test between groups. **Tasks pooled by Tukey paired comparison. % Variation in percent. per task with respect to CSRT.

indicates that FL is not especially sensitive comparing EG with YG. In other words, it seems that selective attention is not specifically affected in this particular group of EG.

4. Conclusions

Results confirmed that CSRT is very discriminant for EG and YG and seems particularly relevant to assess the risk of fall. However, it also showed that SRT test add relevant information about the capacity to perform a quick step. It seems a relevant addition to the CSRT to better assess the risk of fall and highlights specific deficiencies. The additional tasks proposed, in the way they were designed, did not bring much to the CSRT or the SRT. However, tracks to improve these tasks have been proposed and should be studied in future studies. The perspective would consist in comparing results from this test to the clinical state of the subjects, with a prospective follow-up of potential falls. Thus, the final goal would be to obtain a test that could complement clinical data resulting in a reliable indicator of risk of fall in elderly to carry out preventive actions improving medical decision-making.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Agence Nationale de la Recherche (grant number ANR-16-CE19-0006-01). We thank Iberus + program of PhD student mobility and Ibercaja-Cai stays of research program for the financial support to the corresponding author. We also thank the Ligue Auvergne Rhône Alpes de rugby à XIII and its programme "Silver XIII Équilibre" for facilitating elderly participants for this study.

References

Davranche K, Hall B, McMorris T. 2009. Effect of exercise on cognitive control during Eriksen flanker task. J Sport Exerc Psychol. 31(5):628–639.

Heinrich S, Rapp K, Rissmann U, Becker C, König H. 2010. Cost of falls in old age: a systematic review. Osteoporos Int. 21(6):891–902.

Klotz S. 2019. Tapis Equilibre. Master mention mecanique - Parcours biomecanique.

Lord SR, Fitzpatrick RC. 2001. Choice stepping reaction time: a composite measure of falls risk in older people. J Gerontol Series. 56(10):627–632.

Pijnappels M, Delbaere K, Sturnieks D, Lord SR. 2010. The association between choice stepping reaction time and falls in older adults. Age Ageing. 39(1):99–104.

Tinetti ME. 2003. Clinical practice. Preventing falls in elderly persons. N Engl J Med. 348(1):42–49.

Yechiam E, Goodnight J, Bates JE, Busemeyer JR, Dodge KA, Pettit G, Newman JP. 2006. A formal cognitive model of the go/no-go discrimination task: Evaluation and implications. Psychol Assess. 18(3):239–249.

KEYWORDS Balance; cognition; reaction time; risk factor; step initiation

627471@unizar.es

5. DISCUSIÓN GENERAL

5.1. CONTRIBUCIONES

A continuación, se exponen y discuten las aportaciones del conjunto de publicaciones presentadas, de cara a mostrar la conexión entre ellas y en relación con los objetivos generales de esta tesis.

5.1.1. OBJETIVO 1: GENERAR Y PROPORCIONAR INFORMACIÓN OBJETIVA SOBRE EL ESTADO DE PACIENTES QUE PRESENTAN TRASTORNOS DEL EQUILIBRIO.

El objetivo de proporcionar información de pacientes con trastornos del equilibrio ha sido abordado principalmente en los artículos 1 (A1), 2 (A2), 3 (A3) y 5 (A5). Las acciones de investigación ejecutadas en los diferentes artículos mencionados fueron realizadas de manera secuencial, todas ellas como respuesta al objetivo 1.

Verificación de estándares para uso en clínica de plataforma estabilométrica.

El punto de partida fue la validación, para su aplicación en el ámbito clínico, de la plataforma estabilométrica, dispositivo utilizado en gran parte de las investigaciones. Para ello, se requería de un proceso de verificación de estándares establecidos por la ISPGR (Scoppa et al., 2013), siendo abordado este aspecto en el A1. Dichos estándares comprenden aspectos diversos, habiendo sido cumplidos tal y como se expone a continuación:

- Características metrológicas como precisión, exactitud, resolución y linealidad. Dichos parámetros fueron calculados para la plataforma mediante un experimento contemplado en el apartado *2.3. Acciones de Investigación,* a través de la ley de propagación de errores.
- Peso permisible: hasta 300 kg en nuestro equipo (recomendación de 200 kg).
- Altura de los sujetos: sin límite en nuestro equipo (recomendación de 250 cm).
- Longitud del pie: 40 cm en nuestro equipo (recomendación de 35 cm).
- Tasa de muestreo: 60 Hz en nuestro equipo (recomendación de 250 cm) (Scoppa et al., 2013; Yamamoto et al., 2017).
- Origen de coordenadas (X, Y): si bien se recomienda su colocación en la esquina posterior izquierda de la plataforma, se propuso que la referencia del sistema de las coordenadas del COP corresponda a su posición al comienzo de la prueba. Tomando como referencia las recomendaciones del ISPGR, si esa posición fuera elegida, las coordenadas resultantes dependerían de las dimensiones de la plataforma y la posición inicial del sujeto al comienzo de la prueba, lo que dificulta la comparación entre diferentes pruebas. En la posición propuesta, las coordenadas (X, Y) del COP se inicializan a cero al comienzo de la captura e inmediatamente después de la "fase de adaptación" del sujeto, que no debe ser menos de 5 s (Taylor et al., 2015), y una vez que el operador ha verificado que el sujeto está en una posición estable. Además, algunas ventajas adicionales que presenta el origen de coordenadas establecido son: las coordenadas (X, Y) del COP se normalizan con el propio sujeto, mostrando los gráficos resultantes la evolución del paciente durante la prueba desde la posición inicial, facilitando así la interpretación clínica; el estabilograma presentado muestra valores positivos y negativos, siendo especialmente útil para visualizar posibles asimetrías; el estatocinesiograma mostrado tiene un punto de origen con valores positivos y negativos, lo que facilita distinguir los movimientos del COP en los cuatro cuadrantes; no afecta los parámetros clásicos (Scoppa et al., 2013), ya que no

dependen del origen establecido. Finalmente, se recomienda *grabar* el desplazamiento del COP con respecto al punto central de la plataforma en el instante inicial cuando se ajusten a cero las coordenadas, pudiendo usar de esta forma el criterio establecido por el ISPGR.

Diseño de pruebas de evaluación funcional del equilibrio y su aplicación en entornos clínicos.

Las pruebas de evaluación del equilibrio propuestas (en el A1) tienen validez a nivel de diagnóstico, ya que pueden proporcionar información sobre el grado de colaboración del paciente, detectando posibles exageraciones de la patología (patrones fisiológicos inconsistentes) (Ramírez et al., 2014), ya sea debido al miedo, ansiedad u otras razones; siendo un punto de interés en el ámbito de medicina forense.

De acuerdo con otros estudios (Peydro de Moya et al., 2005; Goebel et al., 1997) que han identificado patrones no fisiológicos que potencialmente indican una falta de colaboración observando el desplazamiento del COP en diferentes pruebas, se establecieron un conjunto de pruebas de evaluación funcional del equilibrio, tanto para evaluar el equilibrio estático como dinámico. Los tests estáticos, basados en el test de Romberg (Khasnis and Gokula, 2003), muestran una dificultad creciente según se van realizando las diferentes pruebas. De esta forma se pueden estudiar los patrones y la consistencia de las pruebas, ya que mejores resultados en pruebas más complejas denotarían un patrón inconsistente (Baydal-Bertomeu et al., 2004; Peydro de Moya et al., 2005).

Por tanto, estos patrones y el análisis de coherencia sirvieron para medir la consistencia de las pruebas establecidas y, en consecuencia, comprobar la idoneidad de los protocolos establecidos para su aplicación clínica.

Asimismo, en la misma investigación, se propusieron las variables del equilibrio más relevantes a nivel de diagnóstico para ayudar a los profesionales sanitarios a interpretar y evaluar los resultados de las pruebas de equilibrio. Estas variables fueron identificadas en función de una revisión bibliográfica (ver referencias consultadas en A1) de la literatura actual, sirviendo esta selección para simplificar y facilitar el proceso de evaluación.

Por otra parte, y como resultado de la estancia de investigación en la *Université Gustave Eiffel* (entidad colaboradora en el A5), se ha desarrollado una prueba de evaluación del equilibrio para evaluar la capacidad de percepción, atención e inhibición en personas mayores. Se propusieron varias tareas a realizar en esta prueba, a fin de discernir cuales eran más discriminantes para evaluar el riesgo de caída (Lord and Fitzpatrick 2001). Algunas de estas tareas involucraban la capacidad cognitiva de los sujetos, probando su eficacia, pero a la vez mostrando la necesidad de profundizar en este tipo de pruebas (Davranche et al. 2009).

La introducción de tareas que involucren la percepción, atención e inhibición, así como la capacidad cognitiva, permite estudiar otra vía para predecir el riesgo de caída en pacientes con trastornos del equilibrio, sentando las bases para introducir tareas cognitivas en las pruebas propuestas en el A1.

Adaptación de la información resultante de pruebas de evaluación del equilibrio para facilitar su interpretación clínica.

Pocos estudios se han centrado en la utilidad clínica de la posturografía a nivel de paciente (Visser et al., 2008). Además, aunque la posturografía se considera el *Gold Standard*, existen limitaciones en cuanto a su uso como evaluación funcional (Climent Barbera JM, 2003). Así, el método de evaluación MCQ-Balance propuesto en el A3, se centra en el seguimiento individualizado de los pacientes, intentando dar respuesta a esta problemática. En efecto, la transformación de información de variables cuantitativas continuas

(como por ejemplo el desplazamiento del COP en una prueba determinada) a conclusiones que facilitan la interpretación clínica de los resultados, aporta mayor inteligencia a los dispositivos, siendo una limitación detectada en los informes de posturografía (Climent Barbera JM, 2003).

El trabajo multidisciplinar llevado a cabo en esta tesis, fruto de la colaboración con diferentes facultativos sanitarios, así como del trabajo colaborativo con hospitales y centros de salud, ha posibilitado la detección de la falta de adaptación de la información emitida por estos equipos al perfil de los profesionales que van a hacer uso de ella. En numerosas ocasiones, los médicos no utilizan estos sistemas o tecnologías por no sentirse familiarizados con la operativa de uso y con los resultados obtenidos, por ello se desarrolló el método de evaluación MCQ-Balance.

5.1.2. OBJETIVO 2: MONITORIZAR Y ANALIZAR LA PROGRESIÓN DE TRASTORNOS DEL EQUILIBRIO EN PACIENTES INMERSOS EN TRATAMIENTOS DE REHABILITACIÓN.

La monitorización del equilibrio de pacientes se aborda en los A2 y A3. Se ha seguido un conjunto de fases hasta desarrollar un método capaz de monitorizar el equilibrio (emitiendo conclusiones interpretables por los clínicos como se ha expuesto con anterioridad), contrastando los resultados obtenidos con la evaluación de un médico experto.

Test-retest de las pruebas de evaluación funcional del equilibrio.

En primer lugar, se llevó a cabo un estudio de fiabilidad con una muestra de sujetos sanos en el A2, para obtener datos de referencia, calculando el MDC de las diferentes variables analizadas, así como para comprobar que se podían detectar cambios relevantes en el equilibrio entre dos instantes temporales. Los resultados que se obtuvieron en el test-retest fueron satisfactorios en comparación con la estudios analizados en la *review* realizada por por Ruhe, Fejer, & Walker (2010).

La utilidad de una prueba para detectar cambios relevantes entre dos instantes temporales de un paciente depende del valor del MDC, habiendo destacado en la investigación aquellas variables con menor MDC para los test de evaluación funcional del equilibrio (dichas variables fueron seleccionadas a partir de la selección previa llevada a cabo en el A1). Así, cuanto menor sea el MDC de las variables, más útil será la prueba y mayor capacidad tendrá para detectar cambios relevantes. De esta forma, se muestra la variabilidad de la prueba, demostrando que es posible detectar cambios relevantes en el equilibrio entre diferentes instantes temporales.

Se considera necesario profundizar en la importancia y en las limitaciones de uso del MDC, así como en las conclusiones obtenidas de su uso en las diferentes investigaciones.

Decidir sobre la configuración de la experimentación del A2 fue un asunto que despertó interés y discusión entre el equipo multidisciplinar que participó en este estudio:

- El MDC evita que se confundan los cambios detectados debido al "ruido" inherente de la repetición de una prueba a lo largo del tiempo. Porque son afectados por ciertos factores que introducen variabilidad en los resultados, independiente de la posible patología del paciente, tal como se expone en el siguiente punto.
- En el diseño de la experimentación del cálculo del MDC se priorizó que este debería dirigirse a estudiar y limitar cuantitativamente la variación (variabilidad) entre los resultados de las pruebas pre y post; específicamente, solo de aquellos factores intrínsecos al protocolo aplicado y de la instrumentación utilizada. La intención era que estos factores se aislaran del ruido asociado a otros factores como la patología, el tratamiento aplicado entre pre y post test, la progresión de la

enfermedad, etc. Así, en este estudio una muestra de sujetos sanos y jóvenes fue seleccionada, permitiendo determinar la variabilidad provocada por factores como la variabilidad intrínseca del ser humano, las condiciones de las pruebas (posición del cuerpo y del pie, duración de la prueba, condición visual y propiocepción, etc.), el factor de aprendizaje, incidentes normales durante el transcurso de las pruebas, o la propia comprensión del participante de cómo realizar la prueba. Además, se consideró que elegir una muestra de sujetos sanos sería beneficioso para otras pruebas / investigaciones médicas similares ya que, en general, el reclutamiento de personas sanas es más fiable que acceder a una muestra homogénea de pacientes con una patología específica.

- Si hubiéramos obtenido un MDC de un grupo de pacientes con la misma patología y edad similar a los pacientes presentados, solo se hubieran podido evaluar las influencias del tratamiento restringido al marco de aplicación del MDC a estos pacientes, no permitiendo la evaluación de cambios más allá de estos pacientes debido al tratamiento estudiado.
- De acuerdo con lo anterior, se sabe con certeza que cualquier cambio detectado entre las pruebas pre y post en un paciente, que supere el umbral del valor de MDC, no ha sido causado por los factores intrínsecos ya mencionados. Entre los factores extrínsecos de la prueba que puedan afectar antes y después de la prueba, el efecto del tratamiento y / o las fluctuaciones de la patología deben ser evaluados por el clínico para analizar los resultados del método propuesto.
- Sin embargo, hay que considerar que otros factores también pueden haber afectado al paciente durante el período de evaluación, alterando los efectos del tratamiento; específicamente, otras enfermedades que pueden afectar el equilibrio (por ejemplo, una enfermedad degenerativa) o incluso eventos personales (por ejemplo, la muerte de miembros de la familia, accidentes físicos, etc.) que podrían perjudicar al afectar al estado de ánimo del paciente. Respecto a este punto, propusimos en el A2 que estas consideraciones se incluyan en los criterios de exclusión establecidos en este estudio, que podrían ser utilizados en estudios de temática similar, con el fin de limitar los factores extrínsecos a valorar de los relativos al tratamiento administrado por el médico y / o a las fluctuaciones de la patología.

Por tanto, el clínico recibe información útil ya que puede valorar el cambio detectado en un paciente, excluyendo los factores intrínsecos de la prueba, lo que le permite discernir claramente la progresión que ha seguido el paciente. Esto puede ser relevante para el médico, que comprende la situación de cada paciente y puede valorar las razones de estos cambios.

Monitorización del equilibrio y de los BSS en pacientes con trastornos del equilibrio.

Tras calcular el MDC para las distintas variables analizadas a partir del test-retest, se comprobó la posibilidad de aplicar el MDC con el método MBD para detectar cambios relevantes en pacientes. Para ello, se seleccionó una muestra de ocho pacientes, comprobando que el método era capaz de detectar mejora, empeoramiento o ausencia de cambio.

Asimismo, en el A2 se extrajeron ciertas pautas "lógicas" de interés para valorar el grado de cambio de un paciente a nivel individual en el equilibrio, estableciendo dicho cambio como positivo o negativo respecto a la evaluación inicial. Una reducción en el valor de las variables relacionadas con las pruebas del equilibrio estático se asoció con una mejora en el equilibrio. Un aumento en el valor de las variables del test LOS refleja una mejora en el equilibrio dinámico y un mayor control del COP.

Como se ha expuesto con anterioridad, se observó que la información resultante podía ser aprovechada de manera más eficaz, ya que no resultaba de fácil interpretación para los facultativos y se requería de un

especialista en este tipo de pruebas para la interpretación de los resultados. Para ello, se desarrolló un método para cuantificar, clasificar y calificar la progresión del equilibrio de pacientes, centrando el diseño del método en la monitorización del equilibrio y de los BSS, proporcionando información objetiva y fácilmente interpretable.

La gran ventaja del método reside en que es capaz de emitir conclusiones objetivas de la progresión de los trastornos del equilibrio, así como de los BSS afectados, a partir de la realización únicamente dos sesiones de tests, una pre- y otra post-tratamiento. De esta forma, el facultativo puede obtener información acerca de la progresión del equilibrio de un paciente, en un formato que facilita su interpretación clínica. La sencillez del método, así como el nivel de detalle proporcionado para su reproducibilidad por otros investigadores, permiten sistematizar su uso en la clínica como herramienta de evaluación complementaria. Sin embargo, las investigaciones futuras deberían centrarse en verificar su aplicación de forma integrada en un servicio de rehabilitación hospitalaria.

En relación a las conclusiones resultantes del método, es posible destacar la influencia de los BSS en la progresión del equilibrio de un paciente. De esta forma, el método facilita al clínico la adaptación del tratamiento médico, centrándose en la patología concreta que afecta al equilibrio del paciente.

Si bien se ha demostrado la importancia del método MCQ-Balance en la evaluación de pacientes que están inmersos en un tratamiento, se debe destacar que es el médico el que debe tomar la decisión última de cómo proceder, ya que se trata de una herramienta complementaria para el facultativo que no trata de sustituirle.

Comparativa del Método MCQ-Balance con la evaluación de un médico experto.

El método fue aplicado en una muestra de 42 pacientes con trastornos del equilibrio, con diferentes patologías, para comprobar su aplicabilidad en clínica. Asimismo, se realizó una comparativa entre los resultados emitidos por el método y la evaluación de un médico especialista, obteniendo notable similitud entre los mismos. La evaluación de MCQ-Balance mostró una precisión del 83,4% y un coeficiente Kappa de Cohen de 0,752 en comparación a la realizada por un médico especialista.

Finalmente, se considera que se han logrado aunar los objetivos 1 y 2, ya que se ha conseguido monitorizar trastornos del equilibrio, generando información que facilita la interpretación clínica de los resultados de la prueba y a nivel individual de cada paciente.

5.1.3. OBJETIVO 3: GENERAR MODELOS PREDICTIVOS PARA MEJORA DEL DIAGNÓSTICO EN EL ÁMBITO DE LA SALUD

En relación con el creciente volumen de datos que se generan en los centros hospitalarios, y a la posibilidad de mejorar la “inteligencia” de la información resultante de pruebas médicas complementarias como la utilizada en esta tesis se abordó este objetivo, principalmente, en el artículo 4 (A4). Se desarrolló una metodología para generar modelos predictivos en el ámbito de la salud, siendo aplicada a un caso de estudio de pacientes con patología cervical. La elección de esta patología está directamente relacionada con las investigaciones precedentes, puesto que los trastornos del equilibrio están muy relacionados con las patologías cervicales (Moreno et al., 2017). De esta forma, si se demostraba la eficacia de la metodología en este caso de estudio, se podría aplicar en investigaciones con pacientes con trastornos del equilibrio.

Desarrollo de una metodología de ML aplicable en el campo de la salud.

En el A4 se han mostrado las particularidades de la aplicación de técnicas de ML en el campo de la salud, desarrollando las distintas etapas requeridas para generar modelos fiables y estables.

Sin embargo, durante el planteamiento y desarrollo de la investigación para la aplicación de técnicas de ML en salud, se encontraron dificultades específicas en este campo, tales como la adaptación de estas técnicas al tratamiento de datos sensibles de pacientes, la aplicación de técnicas de anonimización, la importancia de los falsos negativos en este campo en comparación con otros ámbitos, etc.

Con el fin de aclarar y estructurar los conocimientos adquiridos durante el desarrollo de la investigación, se ha elaborado un resumen de algunos aspectos clave y lecciones aprendidas sobre DM y ML en el campo de la salud, el cual se puede consultar en la discusión del A4. Asimismo, se han discutido los siguientes aspectos en relación con la aplicación de ML el campo de la salud: estructuración de los datos, selección de variables, selección del modelo predictivo, variabilidad de las medidas obtenidas, colección de datos para producción, multidisciplinariedad, explotación de los datos y uso de los recursos.

Aplicación de la metodología en pacientes con patología cervical.

La motivación de incluir un caso de estudio se fundamenta en ejemplificar las particularidades de uso de las técnicas de ML en el campo de la salud, ya que se trata de un proceso complejo que comprende numerosas fases. Los resultados obtenidos fueron satisfactorios, pudiendo iniciarse la aplicación de la metodología en pacientes con trastornos del equilibrio. Los resultados muestran que se ha podido predecir la presencia de dolor cervical con exactitud, precisión y *recall* por encima del 85% con los enfoques basados en *SVM, random forest, redes neuronales MLP y GBA.*

Como se ha comentado con anterioridad, la elección de un caso de evaluación cervical tuvo una motivación doble. Por un lado, la relación directa entre patologías cervicales y trastornos del equilibrio, (como se ha expuesto en la sección de *1.2. Antecedentes*), ya que el vértigo de origen cervical tiene una alta prevalencia y numerosos pacientes evaluados en el A3 tienen vértigo de origen cervical. Asimismo, la evaluación del rango cervical puede ayudar en la valoración de determinados trastornos del equilibrio. Por otro lado, una de las líneas futuras de investigación está enmarcada en los tratamientos de trastornos del equilibrio de origen cervical mediante el uso de sensores inerciales de captura de movimiento (los mismos que fueron usados en el A4), por lo que este caso de estudio sirve como base para desarrollar esta línea de investigación.

A la vista de lo expuesto, se desprende que se han abordado los objetivos de esta tesis con la realización de las distintas investigaciones publicadas.

5.2. LIMITACIONES

Más allá de sus contribuciones, el trabajo presentado en esta tesis tiene algunas limitaciones que se deben considerar de manera adicional a las limitaciones ya expuestas en los distintos artículos.

Si bien las muestras utilizadas en los diferentes estudios se consideran adecuadas, podría existir cierto sesgo por tratarse de población únicamente española y aragonesa. Por otra parte, sería interesante aplicar el método MCQ-Balance en una muestra con más número de pacientes, abarcando más tipologías de trastornos del equilibrio, como la fístula perilinfática o la laberintitis, con el objetivo de contrastar los resultados del estudio. En relación con la metodología presentada en el A4, y tal y como se expone en las siguientes líneas, resultaría interesante aplicarla en un caso de estudio con pacientes con trastornos del equilibrio.

Debido al marco temporal que concierne a una tesis doctoral, no se han podido incluir todas las investigaciones deseadas, las cuales se detallan en el siguiente apartado como *líneas futuras de investigación*.

5.3. LINEAS FUTURAS DE INVESTIGACIÓN

Derivados de los resultados de esta tesis, se van a llevar a cabo diferentes trabajos relacionados con las temáticas abordadas. Se realizará un estudio sobre la eficacia de un tratamiento específico en pacientes con vértigo de origen cervical, siendo la muestra escogida una parte de los 42 pacientes evaluados con el método MCQ-Balance (A3). Asimismo, estos pacientes realizaron tests adicionales a las pruebas de evaluación del equilibrio, siendo los mismos movimientos de rango cervical realizados por los sujetos del caso de estudio de la investigación para generar modelos predictivos en el campo de la salud (A4), utilizando el mismo equipo de sensores inerciales empleado en dicho estudio.

Por otra parte, y utilizando como base el trabajo realizado en el A5 donde se involucran tareas cognitivas para poder ser capaz de predecir el riesgo de caídas, se prevé realizar un estudio con modelos predictivos para predicción del riesgo de caída en pacientes con trastornos del equilibrio, similares a los abordados en el A3, siendo necesario aumentar la muestra para afianzar los resultados. En dicho estudio se tiene previsto aplicar la metodología para generación de modelos predictivos en el campo de la salud (expuesta en el A4), en combinación con el método MCQ-Balance.

Por último, y en relación con la comparativa entre los resultados emitidos por el método propuesto y la evaluación de un médico especialista, las líneas futuras de investigación se podrían dirigir a la correlación entre ambos resultados en otro estudio. Se tratará de establecer un procedimiento, incluyendo otros métodos y tests (como el *BESTest*), para afianzar la precisión y los resultados obtenidos.

6. CONCLUSIONES

6.1. CONCLUSIONES

La presente tesis tiene como objetivo mejorar la aplicación clínica de métodos e instrumentación para la evaluación de trastornos del equilibrio, así como sentar las bases para aplicar técnicas de *Machine Learning* en el campo de la salud para optimizar el aprovechamiento de los datos generados, todo ello con el objetivo último de mejorar la toma de decisiones médicas (en línea con la estrategia europea de *"Improving health information and better use of health data"*) y enfocado a la medicina personalizada y al envejecimiento saludable (completamente relacionados con las estrategias europeas de "*Personalised Medicine*" y *"Healthy Ageing")*. En relación a los objetivos inicialmente planteados, se han (1) desarrollado y mejorado métodos y equipos que proporcionan información objetiva a los clínicos sobre el equilibrio de los pacientes; (2) se ha propuesto un método para monitorizar la progresión de trastornos del equilibrio en pacientes inmersos en tratamientos de rehabilitación, proporcionando información más comprensible y fácil de interpretar; así como (3) se ha establecido una metodología para aplicar técnicas de *Machine Learning* en el ámbito de la salud, abordando las particularidades inherentes a este campo.

Fruto de esta investigación y de la publicación de los diversos estudios derivados de la misma, se desprenden las siguientes conclusiones:

1) Se considera que es posible desarrollar e implementar una plataforma estabilométrica de coste accesible y portable que cumpla con los estándares establecido por la *International Society for Posture and Gait Research* para su aplicación clínica, y pueda utilizarse de forma habitual por el personal sanitario para valorar objetivamente el nivel de control del equilibrio estático y dinámico. De esta forma, se pretende contribuir a extender el uso de este tipo de instrumentación, la cual aporta valor como prueba complementaria para apoyar los diagnósticos y los tratamientos de rehabilitación.

2) Se han propuesto y formalizado un conjunto de pruebas, basadas en el test de Romberg y en el test de evaluación de los Límites de Estabilidad, para valorar el equilibrio estático y dinámico en posición de bipedestación. Dichas pruebas miden el nivel de control del equilibrio, así como los límites de estabilidad de los pacientes. De manera complementaria, permiten estimar el grado de consistencia y colaboración durante su ejecución. Asimismo, el conjunto de variables definidas para este conjunto de pruebas favorece su aplicación en entornos clínicos y forenses. Dichas variables han sido: rango de desplazamiento del centro de presiones en los ejes anteroposterior y medio-lateral, área recorrida, velocidad media del centro de presiones, raíz media cuadrática y, de manera adicional para los límites de estabilidad, desplazamiento máximo del centro de presiones a lo largo de cada eje de los radios del octágono de equilibrio y la variable *éxito* -que cuantifica la gestión del centro de presiones.

3) Se estima que es posible detectar cambios relevantes entre dos instantes temporales en pacientes con trastornos del equilibrio, adaptando el método *Magnitude-Based Decision* a las singularidades de las variables que permiten caracterizar el equilibrio, mostrando dichos cambios de una forma gráfica para ayudar en la toma de decisiones médicas y al seguimiento individualizado de los pacientes.

4) Se propone un método para medir, clasificar y calificar la progresión del equilibrio en pacientes con trastornos del equilibrio, adaptando la información resultante para hacerla más comprensible y fácil de interpretar, denominado *método de evaluación MCQ-Balance*. El método, el cual ha sido verificado con la evaluación de un médico experto, aporta información objetiva para monitorizar la progresión

del equilibrio de pacientes, así como permite medir las alteraciones de los sistemas sensoriales del equilibrio, mejorando así la toma de decisiones médicas respecto al tipo de tratamiento.

5) Se considera que integrar tareas que involucren esfuerzos cognitivos en test clásicos de evaluación de la capacidad de reacción, como el *choice stepping reaction time test*, permitiría prevenir de manera más eficiente el riesgo de caídas en personas mayores, con el objetivo de llevar a cabo acciones preventivas que mejoren la toma de decisiones médicas. Para ello se han añadido nuevas tareas y estímulos al *choice stepping reaction time test*, para valorar así la capacidad de percepción y respuesta de ejecución, atención e inhibición de personas mayores.

6) Se ha establecido una metodología para aplicar técnicas de *Machine Learning* en el ámbito de la salud, considerando e integrando las particularidades inherentes a este campo. Todas las etapas involucradas en este tipo de procesos se han incluido en la metodología, de manera que es posible obtener, clasificar, comprender y transformar la información de carácter sanitario para generar modelos predictivos complementarios al diagnóstico médico. Asimismo, se ha demostrado su aplicabilidad a través de un caso práctico con pacientes, verificando así que el enfoque y los resultados obtenidos pueden servir para ayudar a objetivar diagnósticos, mejorar la eficacia de los tratamientos y ahorrar recursos en los sistemas de salud. El procedimiento propuesto también se cree apropiado para cualquier ámbito sanitario independientemente del origen de los datos.

6.2. CONCLUSIONS

The present thesis aims to improve the clinical application of methods and instrumentation for the assessment of balance disorders, as well as to lay the foundations for applying Machine Learning techniques in the health field to optimize the use of the data generated, all of this with the ultimate aim of improving medical decision-making (in line with the European strategy of "Improving health information and better use of health data") and focused on personalized medicine and healthy aging (completely related to the European strategies of "Personalized Medicine ”and “Healthy Aging ”). In relation to the objectives initially established, (1) methods and equipment have been developed and improved that provide objective information to clinicians on the balance of patients; (2) a method has been proposed to monitor the progression of balance disorders in patients undergoing rehabilitation treatments, providing information that is easily interpretable by physicians; as well as (3) a methodology has been established to apply Machine Learning techniques in the field of health, addressing the particularities inherent in this field.

As a result of this research and the publication of the various studies derived from it, the following conclusions have been obtained:

1) It is considered that it is possible to develop and implement an accessible and portable stabilometric platform that meets the standards established by the International Society for Posture and Gait Research for clinical application. It can be used regularly by healthcare professionals to assess objectively the level of control of static and dynamic balance. In this way, it is intended to contribute to extending the use of this type of instrumentation, which adds value as a complementary test to support diagnoses and rehabilitation treatments.
2) A set of tests, based on the Romberg test and the Stability Limits evaluation test, have been proposed and formalized to assess static and dynamic balance in the standing position. These tests measure the level of balance control, as well as the stability limits of patients. In a complementary way, they allow estimating the degree of consistency and collaboration during its execution. Likewise, the set of variables defined for this set of tests favours its application in clinical and forensic environments. These variables have been range of displacement of the centre of pressures in the anteroposterior and mediolateral axes, area, average velocity of the centre of pressures, root mean square and, additionally for the limits of stability, maximum displacement of the centre of pressures along each axis of the radii of the equilibrium octagon and the success variable - which quantifies the management of the centre of pressures.
3) It is estimated that it is possible to detect relevant changes between two time instants in patients with balance disorders, adapting the Magnitude-Based Decision method to the singularities of the variables that allow characterizing the balance, and showing these changes graphically to help medical decision-making and individualized follow-up of patients.
4) A method is proposed to measure, classify and qualify the progression of balance in patients with balance disorders, adapting the resulting information to make it more understandable and easy to interpret, called the MCQ-Balance assessment method. The method, which has been verified with the evaluation of an expert doctor, provides objective information to monitor the progression of balance in patients, as well as to measure alterations in the sensory systems of balance, thus improving medical decision-making regarding the type of treatment.

5) It is considered that integrating tasks that involve cognitive efforts in classic tests for evaluating the ability to react, such as the choice stepping reaction time test, would allow more efficiently prevent the risk of falls in older people, with the aim of leading to carry out preventive actions that improve medical decision-making. For this, new tasks and stimuli have been added to the choice stepping reaction time test, to assess the perception and response capacity of execution, attention, and inhibition of older people.
6) A methodology has been established to apply Machine Learning techniques in the field of health, considering and integrating the particularities inherent in this field. All the stages involved in this type of process have been included in the methodology, so that it is possible to obtain, classify, understand, and transform health information to generate predictive models complementary to medical diagnosis. Likewise, its applicability has been demonstrated through a practical case with patients, thus verifying that the approach and the results obtained can be used to help objectify diagnoses, improve the efficacy of treatments, and save resources in health systems. The proposed procedure is also believed to be appropriate for any healthcare setting regardless of the origin of the data.

7. REFERENCIAS

1. Abdelaziz, A., Elhoseny, M., Salama, A. S., & Riad, A. (2018). A machine learning model for improving healthcare services on cloud computing environment. *Measurement, 119*, 117-128.
2. Aftab, Z., Robert, T., & Wieber, P. B. (2016). Balance recovery prediction with multiple strategies for standing humans. *PloS One, 11*(3), e0151166. doi:10.1371/journal.pone.0151166 [doi]
3. Aggarwal, C. C., & Zhai, C. (2012). *Mining text data* Springer Science & Business Media.
4. Alejandro J. Moreno, Gonzalo Utrilla, Javier Marin, Jose J. Marin, Maria B. Sanchez-Valverde, Ana C. Royo. (2017). Cervical spine assessment with motion capture and passive mobilization. *Journal of Chiropractic Medicine,*
5. Allum, J., Zamani, F., Adkin, A., & Ernst, A. (2002). Differences between trunk sway characteristics on a foam support surface and on the equitest® ankle-sway-referenced support surface. *Gait & Posture, 16*(3), 264-270.
6. Alonso, A. C., Luna, N. M. S., Mochizuki, L., Barbieri, F., Santos, S., & Greve, J. M. D. (2012). The influence of anthropometric factors on postural balance: The relationship between body composition and posturographic measurements in young adults. *Clinics, 67*(12), 1433-1441.
7. Alotaibi, S., Mehmood, R., Katib, I., Rana, O., & Albeshri, A. (2020). Sehaa: A big data analytics tool for healthcare symptoms and diseases detection using twitter, apache spark, and machine learning. *Applied Sciences, 10*(4), 1398.
8. Andersen, R. M., Davidson, P. L., & Baumeister, S. E. (2013). Improving access to care. *Changing the US Health Care System: Key Issues in Health Services Policy and Management, 33*
9. Araújo, F. H., Santana, A. M., & Neto, Pedro de A Santos. (2016). Using machine learning to support healthcare professionals in making preauthorisation decisions. *International Journal of Medical Informatics, 94*, 1-7.
10. Arbuckle, L., & El Emam, K. (December 2013). Anonymizing health data. *O'Reilly Media, Inc.,* doi:ISBN: 978-1449363062
11. Arlot, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics Surveys, 4*, 40–79. doi:10.1214/09-SS054
12. Austin, P. C., Tu, J. V., Ho, J. E., Levy, D., & Lee, D. S. (2013). Using methods from the data-mining and machine-learning literature for disease classification and prediction: A case study examining classification of heart failure subtypes. *Journal of Clinical Epidemiology, 66*(4), 398-407. doi:10.1016/j.jclinepi.2012.11.008 [doi]
13. Ayllón, F. N., & Fernández, D. F. (2006). Función y entrenamiento de la musculatura abdominal. una visión científica. *Journal of Human Sport and Exercise,*
14. Azevedo, Ana Isabel Rojão Lourenço, & Santos, M. F. (2008). KDD, SEMMA and CRISP-DM: A parallel overview. *ISCAP - Informática - Comunicações Em Eventos Científicos,*
15. B³aszczyk, J. W., & Klonowski, W. (2001). Postural stability and fractal dynamics. *Acta Neurobiol.Exp, 61*, 105-112.
16. Bae, J. M. (2014). The clinical decision analysis using decision tree. *Epidemiology and Health, 36*, e2014025. doi:10.4178/epih/e2014025 [doi]
17. Baker, R. (2006). Gait analysis methods in rehabilitation. *Journal of Neuroengineering and Rehabilitation, 3*(1), 4.
18. Baloh, R. W., Jacobson, K. M., Enrietto, J. A., Corona, S., & Honrubia, V. (1998). Balance disorders in older persons: Quantification with posturography. *Otolaryngology—Head and Neck Surgery, 119*(1), 89-92.

19. Bartual, J., & Pérez, N. (1998). El sistema vestibular y sus alteraciones. *Tomo, 1*, 21-22.

20. Batterham, A. M., & Hopkins, W. G. (2006). Making meaningful inferences about magnitudes. *International Journal of Sports Physiology and Performance, 1*(1), 50-57.

21. Bauer, C., Gröger, I., Rupprecht, R., & Gaßmann, K. G. (2008). *Intrasession reliability of force platform parameters in community-dwelling older adults* doi:https://doi.org/10.1016/j.apmr.2008.02.033

22. Baumgartner B, T. R. (2019 Jun 3). Peripheral vertigo. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK430797/

23. Baydal-Bertomeu, J. M., Barberà I Guillem, R., Soler-Gracia, C., Peydro De Moya, M. F., Prat, J. M., & Barona De Guzmán, R. (2004a). *Determinación de los patrones de comportamiento postural en población sana española* doi:http://dx.doi.org/10.1016/S0001-6519(04)78520-9

24. Baydal-Bertomeu, J. M., Barberà I Guillem, R., Soler-Gracia, C., Peydro De Moya, M. F., Prat, J. M., & Barona De Guzmán, R. (2004b). *Determinación de los patrones de comportamiento postural en población sana española* doi:http://dx.doi.org/10.1016/S0001-6519(04)78520-9

25. Bennett, D. A. (2001). How can I deal with missing data in my study? *Australian and New Zealand Journal of Public Health, 25*(5), 464-469.

26. Bentham, J., & Hand, D. J. (2012). Data mining from a patient safety database: The lessons learned. *Data Mining and Knowledge Discovery, 24*(1), 195-217.

27. Benvenuti, F., Mecacci, R., Gineprari, I., Bandinelli, S., Benvenuti, E., Ferrucci, L., . . . Dambrosia, J. M. (1999). Kinematic characteristics of standing disequilibrium: Reliability and validity of a posturographic protocol. *Archives of Physical Medicine and Rehabilitation, 80*(3), 278-287.

28. Bhardwaj, R., Nambiar, A. R., & Dutta, D. (2017). A study of machine learning in healthcare.*2*, 236-241.

29. Bhatla, N., & Jyoti, K. (2012). An analysis of heart disease prediction using different data mining techniques. *International Journal of Engineering, 1*(8), 1-4.

30. Błaszczyk, J. W. (2016). The use of force-plate posturography in the assessment of postural instability. *Gait & Posture, 44*, 1-6.

31. Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM, 55*(4), 77-84.

32. Bonnechere, B., Jansen, B., & Jan, S. V. S. (2016). Cost-effective (gaming) motion and balance devices for functional assessment: Need or hype? *Journal of Biomechanics, 49*(13), 2561-2565.

33. Boomsaad, Z., Telian, S., & Patil, P. (2017). Treatment of intractable vertigo. *Youmans and Winn Neurological Surgery.7th Ed.Philadelphia, PA: Elsevier,*

34. Borges, A. P. O., Carneiro, J. A. O., Zaia, J. E., Carneiro, A. A. O., & Takayanagui, O. M. (2016). Evaluation of postural balance in mild cognitive impairment through a three-dimensional electromagnetic system. *Brazilian Journal of Otorhinolaryngology, 82*(4), 433-441.

35. Bose, I., & Mahapatra, R. K. (2001). Business data mining—a machine learning perspective. *Information & Management, 39*(3), 211-225.

36. Bruniera, J. R. Z., Camiloti, J. F., Penha, O. d. M., Franco, P. P. R., Junior, S., & Marchiori, Luciana Lozza de Moraes. (2015). Comparative analysis of postural balance by posturography in patients with isolated vertigo or associated with hearing loss. *Audiology-Communication Research, 20*(4), 321-326.

37. Bujang, M. A., & Baharum, N. (2017). A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: A review. *Archives of Orofacial Science, 12*(1)

38. Caña-Pino, A., Apolo-Arenas, M. D., Moral-Blanco, J., Álvaro-de Diego, J., & Fernández Gutiérrez, C. (2015). Valoración del equilibrio postural en bipedestación-sedestación en sujetos sanos mediante

acelerometría. estudio piloto. *Fisioterapia, 37*(6), 271-278. doi:http://dx.doi.org.roble.unizar.es:9090/10.1016/j.ft.2014.12.004

39. Cantor, A. B. (1996). Sample-size calculations for cohen's kappa. *Psychological Methods, 1*(2), 150.
40. Carlos Tellería, Sergio Ilarri, Carlos Sánchez. (2020). Text mining of medical documents in spanish: Semantic annotation and detection of recommendations. Paper presented at the *, 16th International Conference on Web Information Systems and Technologies (WEBIST 2020)*(accepted, to appear)
41. Cassidy, J. D., Carroll, L. J., Cote, P., Lemstra, M., Berglund, A., & Nygren, Å. (2000). Effect of eliminating compensation for pain and suffering on the outcome of insurance claims for whiplash injury. *New England Journal of Medicine, 342*(16), 1179-1186.
42. Catley, C., Smith, K., McGregor, C., & Tracy, M. (2009). Extending CRISP-DM to incorporate temporal data mining of multidimensional medical data streams: A neonatal intensive care unit case study., 1-5.
43. Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). CRISP-DM 1.0 step-by-step data mining guide.
44. Charan, J., & Biswas, T. (2013). How to calculate sample size for different study designs in medical research? *Indian Journal of Psychological Medicine, 35*(2), 121-126. doi:10.4103/0253-7176.116232 [doi]
45. Choi, E., Bahadori, M. T., Schuetz, A., Stewart, W. F., & Sun, J. (2016). Doctor ai: Predicting clinical events via recurrent neural networks. Paper presented at the *Machine Learning for Healthcare Conference,* 301-318.
46. Choi, H., & Lee, J. (2020). Immediate effect of balance taping using kinesiology tape on dynamic and static balance after ankle muscle fatigue. Paper presented at the *Healthcare, , 8*(2) 162.
47. Christodoulou, E., Ma, J., Collins, G. S., Steyerberg, E. W., Verbakel, J. Y., & Van Calster, B. (2019). A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *Journal of Clinical Epidemiology, 110*, 12-22.
48. Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*(4), 284.
49. Clavel, D., Mahulea, C., Albareda, J., & Silva, M. (2020). A decision support system for elective surgery scheduling under uncertain durations. *Applied Sciences, 10*(6), 1937.
50. Climent Barbera JM. (2003). Posturografía ¿algo se mueve? *Revista De Biomecánica Instituto De Biomecánica De Valencia, Febrero: 21*
51. Corazza, S. T., Laux, R. C., Foesch, M. L. S., Santos, D. L. d., Machado, R. R., Macedo, T. L., . . . Mezzomo, S. P. (2016). Benefits of functional training for the balance and propioception in visually impaired persons. *Revista Brasileira De Medicina do Esporte, 22*(6), 471-475.
52. Cruz, J. A., & Wishart, D. S. (2006). Applications of machine learning in cancer prediction and prognosis. *Cancer Informatics, 2*
53. Cuozzo Lemos, L. F., de David, A. C., & Bolli Mota, C. (2016). Development of postural balance in brazilian children aged 4-10 years compared to young adults. *Brazilian Journal of Kineanthropometry & Human Performance, 18*(4), 419.
54. Cvecka, J., Cacek, J., & Ondracek, J. (2014). *Agings and the parameters in static postural way* Universidad de Alicante. Área de Educación Física y Deporte. doi:10.14198/jhse.2014.9.Proc1.43
55. da Costa Barbosa, R., & Vieira, M. F. (2017). Postural control of elderly adults on inclined surfaces. *Annals of Biomedical Engineering, 45*(3), 726-738.

56. Dab W, Ségala C, Dor F, Festy B, Lameloise P, Moullec YL, Tertre AL. (2001). Pollution and health: Correlation or causality? the case of the relationship between exposure to particles and cardiopulmonary mortality. *Journal of the Air & Waste Management Association, 51(2)*, 220-235. doi:10.1080/10473289.2001.10464267

57. Dankar, F. K., & El Emam, K. (2012). The application of differential privacy to health data. Paper presented at the *Proceedings of the 2012 Joint EDBT/ICDT Workshops,* 158-166.

58. Davranche, K., Hall, B., & McMorris, T. (2009). Effect of acute exercise on cognitive control required during an eriksen flanker task. *Journal of Sport and Exercise Psychology, 31*(5), 628-639.

59. De la Torre, J., Bonnet, V., Mauti, R., Chabaud, P., & Robert, T. (2020). Implementing a test to assess reaction, attention and inhibition capacity in elderly. *Computer Methods in Biomechanics and Biomedical Engineering, 23*(sup1), S294-S296.

60. De la Torre, J., Marin, J., Marin, J. J., Auria, J. M., & Sanchez-Valverde, M. B. (2017). Balance study in asymptomatic subjects: Determination of significant variables and reference patterns to improve clinical application. *Journal of Biomechanics, 65*, 161-168.

61. de la Torre, J., Marin, J., Marin, J. J., Auria, J. M., & Sanchez-Valverde, M. B. (2017b). Balance study in asymptomatic subjects: Determination of significant variables and reference patterns to improve clinical application. *Journal of Biomechanics,*

62. de la Torre, J., Marin, J., Marin, J. J., Auria, J. M., & Sanchez-Valverde, M. B. (2017c). Balance study in asymptomatic subjects: Determination of significant variables and reference patterns to improve clinical application. *Journal of Biomechanics, 65*, 161-168.

63. De la Torre, J., Marin, J., Polo, M., & Marín, J. J. (2020). Applying the minimal detectable change of a static and dynamic balance test using a portable stabilometric platform to individually assess patients with balance disorders. Paper presented at the *Healthcare, , 8*(4) 402.

64. de Oliveira Silva, D., Briani, R. V., Pazzinatto, M. F., Ferrari, D., Aragão, F. A., de Albuquerque, C. E., . . . de Azevedo, F. M. (2015). Reliability and differentiation capability of dynamic and static kinematic measurements of rearfoot eversion in patellofemoral pain. *Clinical Biomechanics, 30*(2), 144-148.

65. de Sá Ferreira, A., & Baracat, P. J. F. (2014a). Test–retest reliability for assessment of postural stability using center of pressure spatial patterns of three-dimensional statokinesigrams in young health participants. *Journal of Biomechanics, 47*(12), 2919-2924.

66. de Sá Ferreira, A., & Baracat, P. J. F. (2014b). Test–retest reliability for assessment of postural stability using center of pressure spatial patterns of three-dimensional statokinesigrams in young health participants. *Journal of Biomechanics, 47*(12), 2919-2924.

67. de Vet, H. C., & Terwee, C. B. (2010). The minimal detectable change should not replace the minimal important difference. *Journal of Clinical Epidemiology, 63*(7), 804.

68. Degani, A. M., Leonard, C. T., & Danna-dos-Santos, A. (2017). The effects of early stages of aging on postural sway: A multiple domain balance assessment using a force platform. *Journal of Biomechanics, 64*, 8-15.

69. Derebery, M. J. (2000). Diagnóstico y tratamiento del vértigo. *Revista Cubana De Medicina, 39*(4), 238-253.

70. Destrebecqz, A., Perruchet, P., Cleeremans, A., Laureys, S., Maquet, P., & Peigneux, P. (2010). The influence of temporal factors on automatic priming and conscious expectancy in a simple reaction time task. *The Quarterly Journal of Experimental Psychology, 63*(2), 291-309.

71. Di Fabio, R. P. (1996). Meta-analysis of the sensitivity and specificity of platform posturography. *Archives of Otolaryngology–Head & Neck Surgery, 122*(2), 150-156.

72. Dieterich, M., & Eckhardt-Henn, A. (2004). Neurologische und somatoforme schwindelsyndrome. *Der Nervenarzt, 75*(3), 281-302.

73. Dolley, S. (2015). *Big Data Solution to Harnessing Unstructured Data in Healthcare.,*

74. Donders, A. R. T., Van Der Heijden, Geert JMG, Stijnen, T., & Moons, K. G. (2006). A gentle introduction to imputation of missing values. *Journal of Clinical Epidemiology, 59*(10), 1087-1091.

75. Donoghue, D., & Stokes, E. K. (2009). How much change is true change? the minimum detectable change of the berg balance scale in elderly people. *Journal of Rehabilitation Medicine, 41*(5), 343-346.

76. Dorneles, P. P., Meereis, E. C. W., Pranke, G. I., & Mota, C. B. (2014). RELAÇÃO DO ÍNDICE DO ARCO PLANTAR COM O EQUILÍBRIO POSTURAL-DOI: Http://Dx. doi. org/10.18511/0103-1716/rbcm. v22n2p114-120. *Revista Brasileira De Ciência E Movimento, 22*(2), 114-120.

77. Dörre, J., Gerstl, P., & Seiffert, R. (1999). Text mining: Finding nuggets in mountains of textual data. Paper presented at the *Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,* 398-401.

78. Dounskaia, N., Peterson, D., & Bruhns, R. P. (2018). Destabilization of the upright posture through elevation of the center of mass. *Annals of Biomedical Engineering, 46*(2), 318-323.

79. Doyle, T. L., Newton, R. U., & Burnett, A. F. (2005). *Reliability of traditional and fractal dimension measures of quiet stance center of pressure in young, healthy people* doi:https://doi.org/10.1016/j.apmr.2005.05.014

80. Duarte, M., & Freitas, S. M. (2010). Revision of posturography based on force plate for balance evaluation. *Brazilian Journal of Physical Therapy, 14*(3), 183-192.

81. E.A.P.A. Batista G, C. Prati G, and Monard MC. (2004). A study of the behavior of several methods for balancing machine learning training data. *SIGKDD Explor. Newsl, 6(1)*, 20–29. doi:10.1145/1007730.1007735

82. E.A.P.A. Batista G, C. Prati G, and Monard MC. (2005). Balancing strategies and class overlapping. *Lecture Notes in Computer Science, 3646*, 24–35. doi:10.1007/11552253_3

83. El Emam, K., Dankar, F. K., Issa, R., Jonker, E., Amyot, D., Cogo, E., . . . Vaillancourt, R. (2009). A globally optimal k-anonymity method for the de-identification of health data. *Journal of the American Medical Informatics Association, 16*(5), 670-682.

84. El-Kashlan, H. K., Shepard, N. T., Asher, A. M., Smith-Wheelock, M., & Telian, S. A. (1998). Evaluation of clinical measures of equilibrium. *The Laryngoscope, 108*(3), 311-319.

85. Essery, R., Geraghty, A. W., Kirby, S., & Yardley, L. (2017). Predictors of adherence to home-based physical therapies: A systematic review. *Disability and Rehabilitation, 39*(6), 519-534.

86. Evans, R., Lloyd, J., & Pierce, L. (2012). Clinical use of an enterprise data warehouse. *AMIA Annu Symp Proc,* , 189–198.

87. Farhangfar, A., Kurgan, L., & Dy, J. (2008). Impact of imputation of missing values on classification error for discrete data. *Pattern Recognition, 41*(12), 3692-3705.

88. Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P., & Uthurusamy, R. (1996). *Advances in knowledge discovery and data mining* AAAI press Menlo Park.

89. Ferber, R., Osis, S. T., Hicks, J. L., & Delp, S. L. (2016). Gait biomechanics in the era of data science. *Journal of Biomechanics, 49*(16), 3759-3761.

90. Furlan, L., & Sterr, A. (2018). The applicability of standard error of measurement and minimal detectable change to motor learning research—a behavioral study. *Frontiers in Human Neuroscience, 12*, 95.

91. García, R. B., Corresa, S. P., Bertomeu, J. M. B., & Suárez-Varela, M. M. M. (2012). Posturografía estática con pruebas dinámicas. utilidad de los parámetros biomecánicos en la valoración del paciente vestibular. *Acta Otorrinolaringológica Española, 63*(5), 332-338.

92. Gavrilov, L. A., & Heuveline, P. (2003). Aging of population. *The Encyclopedia of Population, 1*, 32-37.

93. Geldhof, E., Cardon, G., De Bourdeaudhuij, I., Danneels, L., Coorevits, P., Vanderstraeten, G., & De Clercq, D. (2006). Static and dynamic standing balance: Test-retest reliability and reference values in 9 to 10Â year old children. *European Journal of Pediatrics, 165*(11), 779-786. doi:10.1007/s00431-006-0173-5

94. General Assembly of the World Medical Association. (2014). World medical association declaration of helsinki: Ethical principles for medical research involving human subjects. *The Journal of the American College of Dentists, 81*(3), 14-18.

95. Gil-Agudo, A., Baydal-Bertomeu, J., Fernández-Bravo, C., Peydro, F., García-Ruisánchez, M., Zubizarreta, C., & Chamarro, E. L. (2006). Determinación de parámetros cinéticos en las pruebas de equilibrio y marcha de pacientes con latigazo cervical. *Rehabilitación, 40*(3), 141-149.

96. Goebel, J. A., Sataloff, R. T., Hanson, J. M., Nashner, L. M., Hirshout, D. S., & Sokolow, C. C. (1997). Posturographic evidence of nonorganic sway patterns in normal subjects, patients, and suspected malingerers. *Otolaryngology-Head and Neck Surgery, 117*(4), 293-302.

97. González, M. Á M., Villegas, A. S., Atucha, E. T., & Fajardo, F. J. F. (2014a). *Bioestadística amigable* Elsevier España.

98. González, M. Á M., Villegas, A. S., Atucha, E. T., & Fajardo, F. J. F. (2014b). *Bioestadística amigable* Elsevier España.

99. Gonzalo Utrilla, J. Javier Marin, M. Belen Sanchez-Valverde, Virginia Gomez , J. Manuel Auria, Maria J.Bone, A .Cristina Royo. (Universidad de Zaragoza. Zaragoza (España). 2017.). Cervical mobility testing in flexion-extension and protraction-retraction to evaluate whiplash syndrome through motion capture.

100. Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black box models. *ACM Computing Surveys, 51(5)*, 1–42. doi:10.1145/3236009

101. Hall, P., & Gill, N. (2018). An introduction to machine learning interpretability.*O'Reilly Media, Inc*

102. Hamburg, M. A., & Collins, F. S. (2010). The path to personalized medicine. *New England Journal of Medicine, 363*(4), 301-304.

103. Hamet, P., & Tremblay, J. (2017). Artificial intelligence in medicine. *Metabolism, 69*, S36-S40.

104. Hanes, D. A., & McCollum, G. (2006). Cognitive-vestibular interactions: A review of patient difficulties and possible mechanisms. *Journal of Vestibular Research, 16*(3), 75-91.

105. Hansson, E. E., Persson, L., & Malmström, E. M. (2013). Influence of vestibular rehabilitation on neck pain and cervical range of motion among patients with whiplash-associated disorder: A randomized controlled trial. *Journal of Rehabilitation Medicine, 45*(9), 906-910.

106. Heinrich, S., Rapp, K., Rissmann, U., Becker, C., & König, H. (2010). Cost of falls in old age: A systematic review. *Osteoporosis International, 21*(6), 891-902.

107. Henri Laude. (2017). *Data scientist y lenguaje R guía de autoformación para el uso de big data* (Francisco Javier Piqueres Juan ed.). Cornellá de Llobregat (Barcelona): Colecciones Epsilon.

108. Hickey, S., Ford, G., Buckley, J., & O'connor, A. F. (1990). Unterberger stepping test: A useful indicator of peripheral vestibular dysfunction? *The Journal of Laryngology & Otology, 104*(8), 599-602.

109. Hopkins, W. G. (2000). Measures of reliability in sports medicine and science. *Sports Medicine, 30*(1), 1-15.

110. Hopkins, W. G. (2017). A spreadsheet for monitoring an individual's changes and trend. *Sportscience, 21*(5-9), 10.

111. Hopkins, W. G., & Batterham, A. M. (2016). Error rates, decisive outcomes and publication bias with several inferential methods. *Sports Medicine, 46*(10), 1563-1573.

112. Hopkins, W. G., Marshall, S. W., Batterham, A. M., & Hanin, J. (2009). Progressive statistics for studies in sports medicine and exercise science. *Medicine and Science in Sports and Exercise, 41*(1), 3-13. doi:10.1249/MSS.0b013e31818cb278 [doi]

113. Hopkins, W. G. (2019). Rebranding MBI as magnitude-based decisions (MBD). *Sportscience, 23*(i-iii)

114. Hoving, J. L., Pool, J. J., van Mameren, H., Devillé, W. J., Assendelft, W. J., de Vet, H. C., . . . Bouter, L. M. (2005). Reproducibility of cervical range of motion in patients with neck pain. *BMC Musculoskeletal Disorders, 6*(1), 59.

115. Huang, Z., Dong, W., Bath, P., Ji, L., & Duan, H. (2015). On mining latent treatment patterns from electronic medical records. *Data Mining and Knowledge Discovery, 29*(4), 914-949.

116. Hughes, J. M. (2016). *Arduino: A technical reference: A handbook for technicians, engineers, and makers* " O'Reilly Media, Inc.".

117. Huurnink, A., Fransz, D. P., Kingma, I., & van Dieën, J. H. (2013). Comparison of a laboratory grade force platform with a nintendo wii balance board on measurement of postural control in single-leg stance balance tasks. *Journal of Biomechanics, 46*(7), 1392-1395.

118. IBM. (Accessed February 16, 2020.). SPSS software [computer software]

119. IBM Corp. (2017). IBM SPSS statistics for windows, version 25.0. armonk, NY: IBM corp.[computer software]

120. Ito, T., Sakai, Y., Ito, Y., Yamazaki, K., & Morita, Y. (2020). Association between back muscle strength and proprioception or mechanoreceptor control strategy in postural balance in elderly adults with lumbar spondylosis. Paper presented at the *Healthcare, , 8*(1) 58.

121. Jacobs, J. V., Horak, F. B., Tran, V. K., & Nutt, J. G. (2006). Multiple balance tests improve the assessment of postural stability in subjects with parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry, 77*(3), 322-326. doi:77/3/322 [pii]

122. James McCaffrey.Neural network train-validate-test stopping. Retrieved from https://visualstudiomagazine.com/articles/2015/05/01/train-validate-test-stopping.aspx

123. Janasik, N., Honkela, T., & Bruun, H. (2009). Text mining in qualitative research: Application of an unsupervised learning method. *Organizational Research Methods, 12*(3), 436-460.

124. Jannot, A., Zapletal, E., Avillach, P., Mamzer, M., Burgun, A., & Degoulet, P. (2017). The georges pompidou university hospital clinical data warehouse: A 8-years follow-up experience. *International Journal of Medical Informatics, 102*, 21-28.

125. Joyner, M. J., & Paneth, N. (2015). Seven questions for personalized medicine. *Jama, 314*(10), 999-1000.

126. Jull, G. A. (2004). Grieve's modern manual therapy: The vertebral column.

127. Kapteyn, T., & Wit, G. d. (1972). Posturography as an auxiliary in vestibular investigation. *Acta Oto-Laryngologica, 73*(2-6), 104-111.

128. Kapteyn, T., Bles, W., Njiokiktjien, C., Kodde, L., Massen, C., & Mol, J. (1983). Standardization in platform stabilometry being a part of posturography. *Agressologie: Revue Internationale De Physio-Biologie Et De Pharmacologie Appliquees Aux Effets De L'Agression, 24*(7), 321â€”326.

129. Kargupta, H., Datta, S., Wang, Q., & Sivakumar, K. (2003). On the privacy preserving properties of random data perturbation techniques. Paper presented at the *Third IEEE International Conference on Data Mining,* 99-106.

130. Karlsson, A., & Frykberg, G. (2000). Correlations between force plate measures for assessment of balance. *Clinical Biomechanics, 15*(5), 365-369.

131. Kayyali, B., Knott, D., & Van Kuiken, S. (2013). The big-data revolution in US health care: Accelerating value and innovation. *Mc Kinsey & Company, 2*(8), 1-13.

132. Khalaj, N., Osman, N. A. A., Mokhtar, A. H., Mehdikhani, M., & Abas, Wan Abu Bakar Wan. (2014). Balance and risk of fall in individuals with bilateral mild and moderate knee osteoarthritis. *PloS One, 9*(3)

133. Khasnis, A., & Gokula, R. M. (2003). Romberg's test. *Journal of Postgraduate Medicine, 49*(2), 169-172.

134. Kim, S. K., Kim, Y. B., Park, I. S., Hong, S. J., Kim, H., & Hong, S. M. (2016). Clinical analysis of dizzy patients with high levels of depression and anxiety. *Journal of Audiology & Otology, 20*(3), 174-178. doi:10.7874/jao.2016.20.3.174 [doi]

135. Klotz, S. (2019). *Tapis equilibre: Évaluation d'un dispositif bas coût pour tester les mécanismes de régulation de l'équilibre*

136. Koh, H. C., & Tan, G. (2011). Data mining applications in healthcare. *Journal of Healthcare Information Management, 19*(2), 65.

137. Korhonen, M. T., Suominen, H., Viitasalo, J. T., Liikavainio, T., Alen, M., & Mero, A. A. (2010). Variability and symmetry of force platform variables in maximum-speed running in young and older athletes. *Journal of Applied Biomechanics, 26*(3), 357-366.

138. Kovacs, F. M., Abraira, V., Royuela, A., Corcoll, J., Alegre, L., Tomás, M., . . . Zamora, J. (2008a). Minimum detectable and minimal clinically important changes for pain in patients with nonspecific neck pain. *BMC Musculoskeletal Disorders, 9*(1), 43.

139. Kovacs, F. M., Abraira, V., Royuela, A., Corcoll, J., Alegre, L., Tomás, M., . . . Zamora, J. (2008b). Minimum detectable and minimal clinically important changes for pain in patients with nonspecific neck pain. *BMC Musculoskeletal Disorders, 9*(1), 43.

140. Ku, P., Osman, N. A., & Abas, W. W. (2016). The limits of stability and muscle activity in middle-aged adults during static and dynamic stance. *Journal of Biomechanics, 49*(16), 3943-3948.

141. Ku, P. X., Abu Osman, N. A., & Wan Abas, W. A. B. (2016). The limits of stability and muscle activity in middle-aged adults during static and dynamic stance. *Journal of Biomechanics, 49*(16), 3943-3948. doi:http://doi.org/10.1016/j.jbiomech.2016.11.006

142. Lafond, D., Corriveau, H., Hébert, R., & Prince, F. (2004). Intrasession reliability of center of pressure measures of postural steadiness in healthy elderly people. *Archives of Physical Medicine and Rehabilitation, 85*(6), 896-901.

143. Lee, H., & Granata, K. P. (2008). Process stationarity and reliability of trunk postural stability. *Clinical Biomechanics, 23*(6), 735-742.

144. Lee, P., Liu, C., Fan, C., Lu, C., Lu, W., & Hsieh, C. (2013a). The test–retest reliability and the minimal detectable change of the purdue pegboard test in schizophrenia. *Journal of the Formosan Medical Association, 112*(6), 332-337.

145. Lee, P., Liu, C., Fan, C., Lu, C., Lu, W., & Hsieh, C. (2013b). The test–retest reliability and the minimal detectable change of the purdue pegboard test in schizophrenia. *Journal of the Formosan Medical Association, 112*(6), 332-337.

146. Lin, D., Seol, H., Nussbaum, M. A., & Madigan, M. L. (2008). Reliability of COP-based postural sway measures and age-related differences. *Gait & Posture, 28*(2), 337-342.

147. Liu, C., Talaei-Khoei, A., Zowghi, D., & Daniel, J. (2017). Data completeness in healthcare: A literature survey. *Pacific Asia Journal of the Association for Information Systems, 9*(2)

148. López, L. F. C., & Calidonio, J. D. L. (2009). "CgMed" design and construction of a platform to determine the position of the center of gravity in standing. *Revista Ingeniería Biomédica, 3*(6), 26-36.

149. Lord, S. R., & Fitzpatrick, R. C. (2001). Choice stepping reaction time: A composite measure of falls risk in older people. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences, 56*(10), M627-M632.

150. Ma, C. Z., Wong, D. W., Lam, W. K., Wan, A. H., & Lee, W. C. (2016). Balance improvement effects of biofeedback systems with state-of-the-art wearable sensors: A systematic review. *Sensors, 16*(4), 434.

151. Maity, N., & Das, S. (2017). Machine learning for improved diagnosis and prognosis in healthcare., 1-9.

152. Mannini, A., & Sabatini, A. M. (2010). Machine learning methods for classifying human physical activity from on-body accelerometers. *Sensors, 10*(2), 1154-1175.

153. Mannini, A., Trojaniello, D., Cereatti, A., & Sabatini, A. M. (2016). A machine learning framework for gait classification using inertial sensors: Application to elderly, post-stroke and huntington's disease patients. *Sensors, 16*(1), 134.

154. Marchetti, G. F., Lin, C. C., Alghadir, A., & Whitney, S. L. (2014). Responsiveness and minimal detectable change of the dynamic gait index and functional gait index in persons with balance and vestibular disorders. *Journal of Neurologic Physical Therapy : JNPT, 38*(2), 119-124. doi:10.1097/NPT.0000000000000015 [doi]

155. Marín Zurdo, J. J., Boné Pina, M. J., & Benito Gil, C. (2013). Evaluación de riesgos de manipulación repetitiva a alta frecuencia basada en análisis de esfuerzos dinámicos en las articulaciones sobre modelos humanos digitales. *Ciencia & Trabajo, 15*(47), 86-93.

156. Marin, J., Blanco, T., & Marin, J. J. (2017). Research lines to improve access to health instrumentation design. *Procedia Computer Science, 113*, 641-646.

157. Marín, J., Boné, M., Ros, R., & Martínez, M. (2008). Move-human sensors: Sistema portátil de captura de movimiento humano basado en sensores inerciales, para el análisis de lesiones musculoesqueléticas y utilizable en entornos reales. Paper presented at the *6º Congreso Internacional De Prevención De Riesgos Laborales; La Coruña,*

158. Martínez Carrasco, Á. (2016). Análisis del riesgo de caídas en ancianos institucionalizados mediante escalas de marcha y equilibrio. *Proyecto De Investigación:,*

159. Maze, F., Blázquez-Tejada, M. T., & Rojas, F. J. (2016). Exploring body sway to disclose changes in postural control strategy associated with proprioceptive training. *European Journal of Human Movement, 37*, 1-20.

160. McGregor, C., Catley, C., & James, A. (2011). A process mining driven framework for clinical guideline improvement in critical care. Paper presented at the *Proceedings of the Learning from Medical Data Streams Workshop. Bled, Slovenia (July 2011),*

161. Medical Subject Headings (MeSH). (2020). Vertigo MeSH descriptor. Retrieved from https://meshb.nlm.nih.gov/record/ui?ui=D014717

162. Meshkati, Z., Namazizadeh, M., Salavati, M., & Mazaheri, M. (2011). Reliability of force-platform measures of postural sway and expertise-related differences. *Journal of Sport Rehabilitation, 20*(4), 442-456.

163. Mohler, M. J., Wendel, C. S., Taylor-Piliae, R. E., Toosizadeh, N., & Najafi, B. (2016). Motor performance and physical activity as predictors of prospective falls in community-dwelling older adults by frailty level: Application of wearable technology. *Gerontology, 62*(6), 654-664. doi:000445889 [pii]

164. Moons, K. G., Kengne, A. P., Grobbee, D. E., Royston, P., Vergouwe, Y., Altman, D. G., & Woodward, M. (2012). Risk prediction models: II. external validation, model updating, and impact assessment. *Heart (British Cardiac Society), 98*(9), 691-698. doi:10.1136/heartjnl-2011-301247 [doi]

165. Moons, K. G., Kengne, A. P., Woodward, M., Royston, P., Vergouwe, Y., Altman, D. G., & Grobbee, D. E. (2012). Risk prediction models: I. development, internal validation, and assessing the incremental value of a new (bio)marker. *Heart (British Cardiac Society), 98*(9), 683-690. doi:10.1136/heartjnl-2011-301246 [doi]

166. Muir, J. W., Kiel, D. P., Hannan, M., Magaziner, J., & Rubin, C. T. (2013). Dynamic parameters of balance which correlate to elderly persons with a history of falls. *PloS One, 8*(8), e70566. doi:10.1371/journal.pone.0070566 [doi]

167. Murphy, C. K. (2001). Identifying diagnostic errors with induced decision trees. *Medical Decision Making, 21*(5), 368-375.

168. Naumann, F. (2014). Data profiling revisited. *SIGMOD Rec, 42(4)*, 40–49. doi:10.1145/2590989.2590995

169. Neuhauser, H. (2016). The epidemiology of dizziness and vertigo. *Handbook of clinical neurology* (pp. 67-82) Elsevier.

170. Obenshain, M. K. (2004). Application of data mining techniques to healthcare data. *Infection Control & Hospital Epidemiology, 25*(8), 690-695.

171. Orejas, J. I. B., Varea, J. A., Rodrigo, J. V., & Navas, A. C. (2020). Resultados y seguimiento de la rehabilitación vestibular. *Revista ORL, 11*(1), 107-114.

172. Padgett, P. K., Jacobs, J. V., & Kasser, S. L. (2012). Is the BESTest at its best? A suggested brief version based on interrater reliability, validity, internal consistency, and theoretical construct. *Physical Therapy, 92*(9), 1197-1207.

173. Pagnacco, G., Carrick, F. R., Wright, C. H. G., & Oggero, E. (2015). *Between-subjects differences of within-subject variability in repeated balance measures: Consequences on the minimum detectable change* doi:https://doi.org/10.1016/j.gaitpost.2014.09.016

174. Parreira, R. B., Boer, M. C., Rabello, L., Costa, Viviane de Souza P, de Oliveira Jr, E., & da Silva, R. A. (2013). Age-related differences in center of pressure measures during one-leg stance are time dependent. *Journal of Applied Biomechanics, 29*(3), 312-316.

175. Patrícia Paludette, D., Fabrício Santana da, S., & Carlos Bolli, M. (2015). Comparação do equilíbrio postural entre grupos de mulheres com diferentes faixas etárias / comparison of postural balance among groups of women with different age ranges / comparación del balance postural en un grupo de mujeres de franjas etarias distintas. *Fisioterapia E Pesquisa, 22*(4), 392.

176. Pawar, P. K., & Dadhich, A. (2012). Study of correlation between human height and hand length in residents of mumbai. *Int J Biol Med Res, 3*(3), 2072-2075.

177. Penger, M., Strobl, R., & Grill, E. (2017). Country-specific and individual determinants of dizziness in europe: Results from the survey of health ageing and retirement in europe (SHARE). *Public Health, 149*, 1-10.

178. Penny, W., & Frost, D. (1996). Neural networks in clinical medicine. *Medical Decision Making, 16*(4), 386-398.

179. Peterson, C. L., Ferrara, M. S., Mrazik, M., Piland, S., & Elliott, R. (2003). Evaluation of neuropsychological domain scores and postural stability following cerebral concussion in sports. *Clinical Journal of Sport Medicine, 13*(4), 230-237.

180. Peydro de Moya, M. F., Baydal Bertomeu, J. M., & Vivas Broseta, M. J. (2005). *Evaluación y rehabilitación del equilibrio mediante posturografía* doi:http://dx.doi.org/10.1016/S0048-7120(05)74365-6

181. phidgets. (2017). *Load cells*

182. Pijnappels, M., Delbaere, K., Sturnieks, D. L., & Lord, S. R. (2010). The association between choice stepping reaction time and falls in older adults—a path analysis model. *Age and Ageing, 39*(1), 99-104.

183. Pinsault, N., & Vuillerme, N. (2009a). Test–retest reliability of centre of foot pressure measures to assess postural control during unperturbed stance. *Medical Engineering & Physics, 31*(2), 276-286.

184. Pinsault, N., & Vuillerme, N. (2009b). Test–retest reliability of centre of foot pressure measures to assess postural control during unperturbed stance. *Medical Engineering & Physics, 31*(2), 276-286.

185. Porta, M. (2014). *A dictionary of epidemiology* Oxford university press.

186. Postolache, O. A., & Postolache, G. B. (2017). Development and selection of balance sensing devices. *IEEE Instrumentation & Measurement Magazine, 20*(1), 38-48.

187. Prashant Gupta.Cross validation in machine learning. Retrieved from https://towardsdatascience.com/cross-validation-in-machine-learning-72924a69872f

188. Rahm E, D. H. (2000). Data cleaning: Problems and current approaches. *Bulletin of the Technical Committee on Data Engineering, 23(4)*, 3-13.

189. Ramírez, C., Ordi, G., Fernández, S., & MI, C. M. (2014). Detección de exageración de síntomas en esguince cervical: Pacientes clínicos versus sujetos análogos. *Trauma, 25*(1), 4-10.

190. Raschka, S., & Mirjalili, V. (2017). *Python machine learning* Packt Publishing Ltd.

191. Raymakers, J., Samson, M., & Verhaar, H. (2005). The assessment of body sway and the choice of the stability parameter (s). *Gait & Posture, 21*(1), 48-58.

192. Redick, T. S., Calvo, A., Gay, C. E., & Engle, R. W. (2011). Working memory capacity and go/no-go task performance: Selective effects of updating, maintenance, and inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(2), 308.

193. Reed, R., & MarksII, R. J. (1999). *Neural smithing: Supervised learning in feedforward artificial neural networks* Mit Press.

194. Refaeilzadeh, P., Tang, L., & Liu, H. (2009). Cross-validation. *Encyclopedia of Database Systems, 5*, 532-538.

195. Reiley, A. S., Vickory, F. M., Funderburg, S. E., Cesario, R. A., & Clendaniel, R. A. (2017). How to diagnose cervicogenic dizziness. *Archives of Physiotherapy, 7*(1), 12.

196. Rosales, R. E., & Rao, R. B. (2010). Guest editorial: Special issue on impacting patient care by mining medical data. *Data Mining and Knowledge Discovery, 20*(3), 325.

197. Rubenstein, L. Z., & Josephson, K. R. (2002). The epidemiology of falls and syncope. *Clinics in Geriatric Medicine, 18*(2), 141-158.

198. Ruhe, A., Fejer, R., & Walker, B. (2010). The test–retest reliability of centre of pressure measures in bipedal static task conditions–a systematic review of the literature. *Gait & Posture, 32*(4), 436-445.

199. Salavati, M., Hadian, M. R., Mazaheri, M., Negahban, H., Ebrahimi, I., Talebian, S., . . . Parnianpour, M. (2009). Test–retest reliabty of center of pressure measures of postural stability during quiet standing in a group with musculoskeletal disorders consisting of low back pain, anterior cruciate ligament injury and functional ankle instability. *Gait & Posture, 29*(3), 460-464.

200. Salehi, R., Ebrahimi-Takamjani, I., Esteki, A., Maroufi, N., & Parnianpour, M. (2010). Test-retest reliability and minimal detectable change for center of pressure measures of postural stability in elderly subjects. *Medical Journal of the Islamic Republic of Iran (MJIRI), 23*(4), 224-232.

201. Sample, R. B., Jackson, K., Kinney, A. L., Diestelkamp, W. S., Reinert, S. S., & Bigelow, K. E. (2016). Manual and cognitive dual tasks contribute to fall-risk differentiation in posturography measures. *Journal of Applied Biomechanics, 32*(6), 541-547.

202. Sánchez-Sánchez, M. L., Belda-Lois, J., Mena-del Horno, S., Viosca-Herrero, E., Igual-Camacho, C., & Gisbert-Morant, B. (2018). A new methodology based on functional principal component analysis to study postural stability post-stroke. *Clinical Biomechanics, 56*, 18-26.

203. Saxena, A., & Prabhakar, M. C. (2013). Performance of DHI score as a predictor of benign paroxysmal positional vertigo in geriatric patients with dizziness/vertigo: A cross-sectional study. *PloS One, 8*(3), e58106. doi:10.1371/journal.pone.0058106 [doi]

204. Schneiders, A. G., Sullivan, S. J., O'Malley, K. J., Clarke, S. V., Knappstein, S. A., & Taylor, L. J. (2010). A valid and reliable clinical determination of footedness. *Pm&r, 2*(9), 835-841.

205. Schuck, P., & Zwingmann, C. (2003). The 'smallest real difference'as a measure of sensitivity to change: A critical analysis. *International Journal of Rehabilitation Research, 26*(2), 85-91.

206. Schuller, E., Eisenmenger, W., & Beier, G. (2000). Whiplash injury in low speed car accidents: Assessment of biomechanical cervical spine loading and injury prevention in a forensic sample. *Journal of Musculoskeletal Pain, 8*(1-2), 55-67.

207. Scoppa, F., Capra, R., Gallamini, M., & Shiffer, R. (2013). Clinical stabilometry standardization: Basic definitions–acquisition interval–sampling frequency. *Gait & Posture, 37*(2), 290-292.

208. Sen, I., & Khandelwal, K. (2018). *Data mining in healthcare* doi:10.13140/RG.2.2.22189.38887

209. Shumway-Cook, A., Woollacott, M., Shumway-Cook, A., & Woollacott, M. (2001). Mobility function. *Motor Control: Theory and Practical Applications.Philadelphia: Lippincott Williams & Wilkins,* , 305-337.

210. Shumway-Cook, A., Brauer, S., & Woollacott, M. (2000). Predicting the probability for falls in community-dwelling older adults using the timed up & go test. *Physical Therapy, 80*(9), 896-903.

211. Solomon, D. (2000). Distinguishing and treating causes of central vertigo. *Otolaryngologic Clinics of North America, 33*(3), 579-601. doi:S0030-6665(05)70228-0 [pii]

212. Solovykh, E. A., Maksimovskaya, L. N., Bugrovetskaya, O. G., & Bugrovetskaya, E. A. (2011). Comparative analysis of methods for evaluation of stabilometry parameters. *Bulletin of Experimental Biology and Medicine, 152*(2), 266-272. doi:10.1007/s10517-011-1504-8

213. Stanton M, F. A. (2020 Apr 28). Vertigo. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK482356/

214. Stavrianou, A., Andritsos, P., & Nicoloyannis, N. (2007). Overview and semantic issues of text mining. *ACM Sigmod Record, 36*(3), 23-34.

215. Steffen, T., & Seney, M. (2008a). 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. *Physical Therapy, 88*(6), 733-746.

216. Steffen, T., & Seney, M. (2008b). 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. *Physical Therapy, 88*(6), 733-746.

217. Steffen, T., & Seney, M. (2008c). 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. *Physical Therapy, 88*(6), 733-746.

218. Stelmach, G. E., & Worringham, C. J. (1985). Sensorimotor deficits related to postural stability: Implications for falling in the elderly. *Clinics in Geriatric Medicine, 1*(3), 679-694.

219. Stewart, M. G., Chen, A. Y., Wyatt, J. R., Favrot, S., Beinart, S., Coker, N. J., & Jenkins, H. A. (1999). Cost-effectiveness of the diagnostic evaluation of vertigo. *The Laryngoscope, 109*(4), 600-605.

220. Steyerberg, E. W. (2009). *Clinical prediction models* Springer.

221. Strupp, M., Dieterich, M., & Brandt, T. (2013). The treatment and natural course of peripheral and central vertigo. *Deutsches Ärzteblatt International, 110*(29-30), 505.

222. Swanenburg, J., de Bruin, E. D., Favero, K., Uebelhart, D., & Mulder, T. (2008). The reliability of postural balance measures in single and dual tasking in elderly fallers and non-fallers. *BMC Musculoskeletal Disorders, 9*(1), 162.

223. Swanenburg, J., de Bruin, E. D., Stauffacher, M., Mulder, T., & Uebelhart, D. (2007). Effects of exercise and nutrition on postural balance and risk of falling in elderly people with decreased bone mineral density: Randomized controlled trial pilot study. *Clinical Rehabilitation, 21*(6), 523-534.

224. Tao, W., Liu, T., Zheng, R., & Feng, H. (2012). Gait analysis using wearable sensors. *Sensors, 12*(2), 2255-2283.

225. Tapias, L. M. (2014). Factores individuales que influyen en la adherencia al tratamiento fisioterapéutico. una revisión del tema.(determining individual aspects treatment adherence in population with disabilities in medellin city: A literature review). *CES Movimiento Y Salud, 2*(1), 23-30.

226. Taylor, M. R., Sutton, E. E., Diestelkamp, W. S., & Bigelow, K. E. (2015). Subtle differences during posturography testing can influence postural sway results: The effects of talking, time before data acquisition, and visual fixation. *Journal of Applied Biomechanics, 31*(5), 324-329.

227. Terrier, P., & Reynard, F. (2014). To what extent does not wearing shoes affect the local dynamic stability of walking?: Effect size and intrasession repeatability. *Journal of Applied Biomechanics, 30*(2), 305-309.

228. Tesio, L., Rota, V., Longo, S., & Grzeda, M. T. (2013). Measuring standing balance in adults: Reliability and minimal real difference of 14 instrumental measures. *International Journal of Rehabilitation Research, 36*(4), 362-374.

229. Thanh Noi, P., & Kappas, M. (2018). Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using sentinel-2 imagery. *Sensors, 18*(1), 18.

230. Timothy C. Hain, M. (May 5, 2019). MOVING PLATFORM POSTUROGRAPHY TESTING computerized dynamic posturography (CDP). Retrieved from https://www.dizziness-and-balance.com/testing/posturography.html

231. Tinetti, M. E. (2003). Preventing falls in elderly persons. *New England Journal of Medicine, 348*(1), 42-49.

232. Toledo, E., & Olivera, R.Desarrollo de una plataforma dinamométrica para el cálculo del centro de presiones utilizando galgas extensiométricas.

233. Tomar, D., & Agarwal, S. (2013). A survey on data mining approaches for healthcare. *International Journal of Bio-Science and Bio-Technology, 5*(5), 241-266.

234. Torre, J. d. l., Marin, J., Ilarri, S., & Marin, J. J. (2020). Applying machine learning for healthcare: A case study on cervical pain assessment with motion capture. *Applied Sciences, 10*(17), 5942.

235. Tsukamoto, H. F., Costa, Viviane de Souza Pinho, Silva Junior, Rubens Alexandre da, Pelosi, G. G., Marchiori, Luciana Lozza de Moraes, Vaz, C. R. S., & Fernandes, K. B. P. (2015). Effectiveness of a vestibular rehabilitation protocol to improve the health-related quality of life and postural balance in patients with vertigo. *International Archives of Otorhinolaryngology, 19*(3), 238-247.

236. Uebbing, T. J. (2016). *User Experience in Smart Environments: Design and Prototyping,*

237. Uyar, A., Bener, A., & Ciray, H. N. (2015). Predictive modeling of implantation outcome in an in vitro fertilization setting: An application of machine learning methods. *Medical Decision Making, 35*(6), 714-725.

238. Uzefovsky, F., Allison, C., Smith, P., & Baron-Cohen, S. (2016). Brief report: The go/no-go task online: Inhibitory control deficits in autism in a large sample. *Journal of Autism and Developmental Disorders, 46*(8), 2774-2779.

239. Vanmeerhaeghe, A. F., Rodriguez, D. R., Tutusaus, L. C., Calafat, C. B., Riera, M. L., & Vidal, A. M. (2009). Diferencias en la estabilidad postural estática y dinámica según sexo y pierna dominante. *Apunts.Medicina De L'Esport, 44*(162), 74-81.

240. Vaupel, J. W., & Loichinger, E. (2006). Redistributing work in aging europe. *Science (New York, N.Y.), 312*(5782), 1911-1913. doi:312/5782/1911 [pii]

241. Vellas, B. J., Wayne, S. J., Romero, L., Baumgartner, R. N., Rubenstein, L. Z., & Garry, P. J. (1997). One-leg balance is an important predictor of injurious falls in older persons. *Journal of the American Geriatrics Society, 45*(6), 735-738.

242. Visser, J. E., Carpenter, M. G., van der Kooij, H., & Bloem, B. R. (2008). The clinical utility of posturography. *Clinical Neurophysiology, 119*(11), 2424-2436.

243. Von Lubitz, D., & Wickramasinghe, N. (2006). Healthcare and technology: The doctrine of networkcentric healthcare. *International Journal of Electronic Healthcare, 2*(4), 322-344.

244. Wang, F., Stiglic, G., Obradovic, Z., & Davidson, I. (2015). *Guest Editorial: Special Issue on Data Mining for Medicine and Healthcare,*

245. Weaver, T. B., Ma, C., & Laing, A. C. (2016). Use of the nintendo wii balance board for studying standing static balance control: Technical considerations, force-plate congruency, and the effect of battery life. *Journal of Applied Biomechanics,* , 1-22.

246. Weiss, J. C., Natarajan, S., Peissig, P. L., McCarty, C. A., & Page, D. (2012). Machine learning for personalized medicine: Predicting primary myocardial infarction from electronic health records. *Ai Magazine, 33*(4), 33-33.

247. Wilkowska, W., & Ziefle, M. (2012). Privacy and data security in e-health: Requirements from the user's perspective. *Health Informatics Journal, 18*(3), 191-201.

248. Williams, Q. I., Gunn, A. H., Beaulieu, J. E., Benas, B. C., Buley, B., Callahan, L. F., . . . Goode, A. P. (2015). Physical therapy vs. internet-based exercise training (PATH-IN) for patients with knee osteoarthritis: Study protocol of a randomized controlled trial. *BMC Musculoskeletal Disorders, 16*(1), 264.

249. Wipperman, J. (2014). Dizziness and vertigo. *Primary Care: Clinics in Office Practice, 41*(1), 115-131.

250. Witten, I., Frank, E., Hall, M., & Pal, C. (2016). Data mining: Practical machine learning tools and techniques. appendix B: The WEKA workbench. *Fourth Edition,* doi:ISBN 978-0128042915

251. Wolf, F., Krebs, R. J., Detânico, R. C., Keulen, G., & Braga, R. (2008). Estudo do equilíbrio plantar do iniciante de tiro com arco recurvo. *Rev Educ Fís/UEM, 19*(1), 1-9.

252. Wolf, S. L., Barnhart, H. X., Kutner, N. G., McNeely, E., Coogler, C., Xu, T., & Atlanta FICSIT Group. (1996a). Reducing frailty and falls in older persons: An investigation of tai chi and computerized balance training. *Journal of the American Geriatrics Society, 44*(5), 489-497.

253. Wolf, S. L., Barnhart, H. X., Kutner, N. G., McNeely, E., Coogler, C., Xu, T., & Atlanta FICSIT Group. (1996b). Reducing frailty and falls in older persons: An investigation of tai chi and computerized balance training. *Journal of the American Geriatrics Society, 44*(5), 489-497.

254. Wrisley, D. (2007). Vision of balance rehabilitation. *A Paradigm Shift: Technology Based Intervention Improving Outcomes for Dysequilibrium, Dizziness, Mobility, Balance, and Falls, Celebration, Florida, USA,*

255. Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., . . . Philip, S. Y. (2008). Top 10 algorithms in data mining. *Knowledge and Information Systems, 14*(1), 1-37.

256. Xu, Y., & Goodacre, R. (2018). On splitting training and validation set: A comparative study of cross-validation, bootstrap and systematic sampling for estimating the generalization performance of supervised learning. *Journal of Analysis and Testing, 2*(3), 249-262.

257. Yamamoto, M., Ishikawa, K., Aoki, M., Mizuta, K., Ito, Y., Asai, M., . . . Murofushi, T. (2017). Japanese standard for clinical stabilometry assessment: Current status and future directions. *Auris Nasus Larynx,*

258. Yardley, L., & Redfern, M. S. (2001). Psychological factors influencing recovery from balance disorders. *Journal of Anxiety Disorders, 15*(1-2), 107-119.

259. Yechiam, E., Goodnight, J., Bates, J. E., Busemeyer, J. R., Dodge, K. A., Pettit, G. S., & Newman, J. P. (2006). A formal cognitive model of the go/no-go discrimination task: Evaluation and implications. *Psychological Assessment, 18*(3), 239.

260. Yin, M., Ishikawa, K., Wong, W. H., & Shibata, Y. (2009). A clinical epidemiological study in 2169 patients with vertigo. *Auris Nasus Larynx, 36*(1), 30-35.

261. Yoo, I., Alafaireet, P., Marinov, M., Pena-Hernandez, K., Gopidi, R., Chang, J., & Hua, L. (2012). Data mining in healthcare and biomedicine: A survey of the literature. *Journal of Medical Systems, 36*(4), 2431-2448.

262. Yuan, Q., Yu, L., Shi, D., Ke, X., & Zhang, H. (2015). Anxiety and depression among patients with different types of vestibular peripheral vertigo. *Medicine, 94*(5)

263. Zhang, P., Wang, F., Hu, J., & Sorrentino, R. (2014). Towards personalized medicine: Leveraging patient similarity and drug similarity analytics. *AMIA Summits on Translational Science Proceedings, 2014*, 132.

264. Zhang, Z., Zhao, Y., Canes, A., Steinberg, D., Lyashevska, O., & written on behalf of AME Big-Data Clinical Trial Collaborative Group. (2019). Predictive analytics with gradient boosting in clinical medicine. *Annals of Translational Medicine, 7*(7), 152. doi:10.21037/atm.2019.03.29 [doi]

265. Zhao, L. (2013). The gut microbiota and obesity: From correlation to causality. *Nat Rev Microbiol, 11*, 639–647. doi:10.1038/nrmicro3089

266. Zhu, Y. (2017). Design and validation of a low-cost portable device to quantify postural stability. *Sensors, 17*(3), 619.

8. APÉNDICES

8.1. FACTOR DE IMPACTO DE LAS PUBLICACIONES

Artículo 1

De la Torre, J.; Marin, J.; Marin, J. J.; Auria, J. M.; Sanchez-Valverde, M. B. **Balance study in asymptomatic subjects: Determination of significant variables and reference patterns to improve clinical application**. *J. Biomech*. **2017**, *65*, 161–168, doi:10.1016/j.jbiomech.2017.10.013.

JOURNAL OF BIOMECHANICS

ISSN:0021-9290

Factor de Impacto JCR (2019):

2.320

Áreas Temáticas:

Biomedical Engineering: Q3 (48/87)

Biophysics: Q3 (38/71)

Artículo 2

De la Torre, J.; Marin, J.; Polo, M.; Marín, J.J. **Applying the Minimal Detectable Change of a Static and Dynamic Balance Test Using a Portable Stabilometric Platform to Individually Assess Patients with Balance Disorders**. *Healthcare (S.I. Ageing Effects on Kinematics, Kinetics and Balance)* **2020**, *8*, 402, https://doi.org/10.3390/healthcare8040402.

HEALTHCARE

ISSN: 2227-9032

Factor de Impacto JCR (2019):

1.916

Áreas Temáticas:

Health Care Sciences & Services: Q3 (62/102)

Artículo 3

De la Torre J, Marin J, Polo M, Gómez-Trullén EM, Marin JJ. **2021. MCQ-Balance: a method to monitor patients with balance disorders and improve clinical interpretation of posturography.** PeerJ 9:e10916 https://doi.org/10.7717/peerj.10916

PEERJ

ISSN: 2167-8359

Factor de Impacto JCR (2019):

2.379

Áreas Temáticas:

Multidisciplinary Sciences: Q2 (32/71)

Artículo 4

De la Torre, J.; Marin, J.; Ilarri, S.; Marin, J.J. **Applying Machine Learning for Healthcare: A Case Study on Cervical Pain Assessment with Motion Capture**. *Appl. Sci (S.I. Medical Informatics and Data Analysis)* **2020**, *10*, 5942, doi: 10.3390/app10175942.

APPLIED SCIENCES

ISSN: 2076-3417

Factor de Impacto JCR (2019):

2.474

Áreas Temáticas:

Multidisciplinary Engineering: Q2 (32/91)

Applied Physics: Q2 (63/155)

Multidisciplinary Materials Science: Q3 (161/314)

Multidisciplinary Chemistry: Q2 (88/177)

Artículo 5

De la Torre, J.; Bonnet, V.; Mauti, R.; Chabaud, P.; Robert, T. **Implementing a test to assess reaction, attention and inhibition capacity in elderly**. *Computer Methods in Biomechanics and Biomedical Engineering*, 2020, vol. 23 (sup1), S294-S296, doi: 10.1080/10255842.2020.1815313.

COMPUTER METHODS IN BIOMECHANICS AND BIOMEDICAL ENGINEERING

ISSN: 1476-8259; 1025-5842

Factor de Impacto SJR (2019):

0.45

Áreas Temáticas:

Bioengineering: Q3

Biomedical Engineering: Q3

Computer Sciences Applications: Q2

Human-Computer Interaction: Q2

Medicine (miscellaneous): Q2

8.2. CONTRIBUCIÓN COAUTORIA

Artículo 1

De la Torre, J.; Marin, J.; Marin, J. J.; Auria, J. M.; Sanchez-Valverde, M. B. **Balance study in asymptomatic subjects: Determination of significant variables and reference patterns to improve clinical application**. *J. Biomech*. **2017**, *65*, 161–168, doi:10.1016/j.jbiomech.2017.10.013.

- Revisión de la literatura y conceptualización.
- Diseño del procedimiento experimental.
- Realización del experimento y recogida de datos.
- Tratamiento de los datos.
- Análisis estadístico de los datos.
- Discusión de los resultados.
- Redacción del artículo.
- Revisión y mejora del documento siguiendo los comentarios de los correspondientes revisores.

Artículo 2

De la Torre, J.; Marin, J.; Polo, M.; Marín, J.J. **Applying the Minimal Detectable Change of a Static and Dynamic Balance Test Using a Portable Stabilometric Platform to Individually Assess Patients with Balance Disorders**. *Healthcare (S.I. Ageing Effects on Kinematics, Kinetics and Balance)* **2020**, *8*, 402, https://doi.org/10.3390/healthcare8040402.

- Revisión de la literatura y conceptualización.
- Diseño del procedimiento experimental.
- Realización del experimento y recogida de datos.
- Tratamiento de los datos.
- Análisis estadístico de los datos.
- Discusión de los resultados.
- Redacción del artículo.
- Revisión y mejora del documento siguiendo los comentarios de los correspondientes revisores.

Artículo 3

De la Torre J, Marin J, Polo M, Gómez-Trullén EM, Marin JJ. **2021. MCQ-Balance: a method to monitor patients with balance disorders and improve clinical interpretation of posturography.** PeerJ 9:e10916 https://doi.org/10.7717/peerj.10916

- Revisión de la literatura y conceptualización.
- Diseño del procedimiento experimental.

- Desarrollo metodología.
- Desarrollo software.
- Realización del experimento y recogida de datos.
- Tratamiento de los datos.
- Análisis estadístico de los datos.
- Discusión de los resultados.
- Redacción del artículo.
- Revisión y mejora del documento siguiendo los comentarios de los correspondientes revisores.
- Supervisión.

Artículo 4

De la Torre, J.; Marin, J.; Ilarri, S.; Marin, J.J. **Applying Machine Learning for Healthcare: A Case Study on Cervical Pain Assessment with Motion Capture**. *Appl. Sci (S.I. Medical Informatics and Data Analysis)* **2020**, *10*, 5942, doi: 10.3390/app10175942.

- Revisión de la literatura y conceptualización.
- Diseño del procedimiento experimental.
- Desarrollo metodología.
- Desarrollo software.
- Realización del experimento y recogida de datos.
- Tratamiento de los datos.
- Análisis estadístico de los datos.
- Discusión de los resultados.
- Redacción del artículo.
- Revisión y mejora del documento siguiendo los comentarios de los correspondientes revisores.
- Supervisión.

Artículo 5

De la Torre, J.; Bonnet, V.; Mauti, R.; Chabaud, P.; Robert, T. **Implementing a test to assess reaction, attention and inhibition capacity in elderly**. *Computer Methods in Biomechanics and Biomedical Engineering*, 2020, vol. 23 (sup1), S294-S296, doi: 10.1080/10255842.2020.1815313.

- Revisión de la literatura y conceptualización.
- Diseño del procedimiento experimental.
- Desarrollo software.
- Realización del experimento y recogida de datos.

- Tratamiento de los datos.
- Análisis estadístico de los datos.
- Discusión de los resultados.
- Redacción del artículo.
- Revisión y mejora del documento siguiendo los comentarios de los correspondientes revisores.

8.3. ACTIVIDADES DE INVESTIGACIÓN RELACIONADAS

Artículo 6

Marin, J.; Marin, J.J.; Blanco, T.; **De la Torre, J.**; Salcedo, I.; Martitegui, E.
Is My Patient Improving? Individualized Gait Analysis in Rehabilitation. Appl. Sci. 2020, 10, 8558, doi: 10.3390/app10238558.

APPLIED SCIENCES

ISSN: 2076-3417

Factor de Impacto JCR (2019):

2.474

Áreas Temáticas:

Multidisciplinary Engineering: Q2 (32/91)

Applied Physics: Q2 (63/155)

Multidisciplinary Materials Science: Q3 (161/314)

Multidisciplinary Chemistry: Q2 (88/177)

Artículo 7

Marín, J.; Blanco, T.; **De la Torre, J.**; Marín, J.J. **Gait Analysis in a Box: A System Based on Magnetometer-Free IMUs or Clusters of Optical Markers with Automatic Event Detection**. Sensors 2020, 20, 3338, doi: 10.3390/s20123338.

SENSORS

ISSN: 1424-8220

Factor de Impacto JCR (2019):

3.275

Áreas Temáticas:

Computer Science: Information Systems: Q1 (70/300)

Physics and Astronomy: Instrumentation: Q1 (17/129)

Electrical and Electronic Engineering: Q1 (147/670)