



Informe Proyecto 1

Análisis estadístico

Santiago Tabares Osorio, Juan Andres Melo Melo

Bioingeniería, Facultad de Ingeniería, Universidad de Antioquia

Bioseñales

22 Septiembre 2024

Resumen

Este estudio presenta un análisis estadístico comparativo entre muestras de sujetos sanos y aquellos con enfermedades neurodegenerativas, específicamente Parkinson, Huntington y Esclerosis Lateral Amiotrófica (ALS), utilizando datos de la dinámica de la marcha. Las pruebas estadísticas incluyen el test de normalidad de Shapiro-Wilk, la prueba de homogeneidad de varianzas de Levene, y según los resultados, la prueba t de Student o la prueba U de Mann-Whitney para comparar grupos. Se evaluaron varios parámetros relacionados con la marcha para cada grupo de pacientes, limitando la muestra a 13 sujetos por grupo para garantizar la coherencia del análisis. Los resultados permiten observar diferencias significativas en las dinámicas de la marcha entre los grupos, lo que puede contribuir a mejorar la comprensión de la fisiopatología de estas enfermedades y ayudar en el diseño de intervenciones terapéuticas más efectivas.

1. Introducción

Las enfermedades neurodegenerativas, como el Parkinson, la enfermedad de Huntington y la Esclerosis Lateral Amiotrófica (ALS), suelen tener un impacto significativo en la marcha y la movilidad de los pacientes. Cuantificar con precisión las dinámicas de la marcha es clave para mejorar la comprensión de la fisiopatología subyacente a estas condiciones, así como para medir la efectividad de intervenciones terapéuticas. La base de datos utilizada en este estudio, disponible en PhysioNet [1], incluye registros de la marcha de pacientes con Parkinson ($n = 15$), Huntington ($n = 20$), ALS ($n = 13$) y un grupo de control sano ($n = 16$). Los datos se obtuvieron a través de resistencias sensibles a la fuerza bajo el pie, proporcionando información valiosa sobre la distribución y dinámica de las fuerzas durante la marcha. En este estudio, se aplican pruebas estadísticas para comparar estos parámetros entre los grupos, con el objetivo de identificar diferencias significativas que puedan reflejar alteraciones en la dinámica de la marcha relacionadas con estas enfermedades.

2. Materiales y métodos

Para este análisis, se utilizó la base de datos disponible en PhysioNet que contiene registros de la marcha de pacientes con Parkinson ($n = 15$), enfermedad de Huntington ($n = 20$), Esclerosis Lateral Amiotrófica (ALS) ($n = 13$) y un grupo control de sujetos sanos ($n = 16$). Los datos fueron obtenidos utilizando resistencias sensibles a la fuerza, las cuales capturan la presión

ejercida bajo el pie durante la marcha. Estas señales proporcionan una representación precisa de la dinámica de la marcha y su variabilidad entre los grupos.

Para el análisis estadístico, inicialmente se calculan parámetros estadísticos relacionados a la marcha, como entropía de permutación, desviación estándar y coeficiente de variación, luego se grafican mediante gráfico de cajas y bigotes, y gráfico de densidad para tener una idea de como es la dispersión de los datos para continuar con las pruebas estadísticas, es importante mencionar que se limitaron las muestras a 13 sujetos por grupo, seleccionados aleatoriamente para asegurar una comparación equilibrada. Se realizaron las siguientes pruebas estadísticas:

Prueba de normalidad de Shapiro-Wilk, para verificar si los datos seguían una distribución normal en cada grupo.

Prueba de homogeneidad de varianza de Levene, para evaluar si las varianzas entre los grupos eran homogéneas. Dependiendo de los resultados anteriores, se aplicó una prueba t de Student para comparar las medias entre los grupos si los datos eran normales y homogéneos, o una prueba U de Mann-Whitney si no se cumplían los supuestos de normalidad o igualdad de varianzas. Se evaluaron múltiples parámetros en variables relacionados con la marcha como el intervalo y el balanceo de cada pie. Además, se generaron gráficas comparativas para cada parámetro entre los grupos de sujetos sanos y aquellos con enfermedades neurodegenerativas. Los análisis se realizaron utilizando Python

y librerías estadísticas y de visualización como SciPy y Matplotlib.

3. Requerimientos

Enfermedad de Parkinson: El diagnóstico de la enfermedad de Parkinson se basa principalmente en criterios clínicos. Los médicos evalúan los síntomas motores clave, como temblores en reposo, rigidez muscular, bradicinesia (lentitud en el movimiento) y la inestabilidad postural. No existe una prueba de diagnóstico específica para el Parkinson, por lo que el diagnóstico suele depender del historial médico del paciente, un examen neurológico y la exclusión de otras posibles causas de los síntomas. Los síntomas motores del Parkinson generalmente se manifiestan después de una pérdida significativa de neuronas dopaminérgicas en la sustancia negra del cerebro, lo que retrasa el diagnóstico temprano. Además, los síntomas iniciales, como cambios en el olfato, depresión y estreñimiento, son inespecíficos, lo que dificulta aún más el diagnóstico temprano. Herramientas de imagen, como la tomografía por emisión de positrones (PET) y la tomografía por emisión de fotón único (SPECT), pueden ayudar en algunos casos, pero no están ampliamente disponibles para un diagnóstico temprano [4].

Enfermedad de Huntington: El diagnóstico se confirma mediante una prueba genética que detecta la repetición anormal de trinucleótidos (CAG) en el gen HTT. Además de la prueba genética, los médicos evalúan los síntomas motores (como movimientos involuntarios conocidos como corea), síntomas psiquiátricos y cognitivos que suelen aparecer en fases iniciales. A pesar de la claridad de la prueba genética, el desafío en el diagnóstico temprano radica en el inicio tardío de los síntomas motores evidentes. Muchos individuos portadores del gen pueden no mostrar síntomas significativos durante varios años, incluso décadas. Los primeros signos de la enfermedad pueden incluir cambios sutiles en el comportamiento, irritabilidad o depresión, que no siempre se asocia inmediatamente con la enfermedad de Huntington. El uso de biomarcadores y estudios de neuroimagen en investigación puede ayudar a predecir la aparición de la enfermedad, pero no están estandarizados para el diagnóstico clínico temprano [5].

Esclerosis lateral amiotrófica (ELA o ALS por sus siglas en inglés): El diagnóstico de ALS es complejo y se basa en la exclusión de otras enfermedades. Incluye un examen neurológico completo, estudios de electrodiagnóstico (EMG), estudios de conducción nerviosa y pruebas de imagen como resonancia magnética para descartar otras causas. El criterio de diagnóstico más utilizado es el de El Escorial, desarrollado por la World Federation of Neurology. El diagnóstico temprano de ALS es un desafío porque los síntomas iniciales, como la debilidad muscular o el desgaste, suelen ser inespecíficos y pueden confundirse con otras afecciones. No existen biomarcadores fiables y específicos que confirmen el diagnóstico de ALS en etapas tempranas, lo que retrasa el tratamiento. Los esfuerzos actuales se centran en identificar biomarcadores en el líquido cefalorraquídeo y la sangre, así como en mejorar las técnicas de neuroimagen, pero estos métodos aún no están disponibles para su uso rutinario [6].

Funcionamiento de la función Detrend: La función detrend de SciPy, en particular del módulo `scipy.signal`, se utiliza para

eliminar tendencias lineales (y opcionalmente de otros tipos) de los datos. Esto es útil en análisis de señales, donde las tendencias pueden ser causadas por fenómenos externos o ruido y pueden enmascarar las características reales de la señal.

En cuanto al desarrollo matemático:

$$x(t) = mt + b + r(t)$$

Donde, m es la pendiente de la tendencia (la tasa de cambio de la señal), b es el término independiente (el valor inicial o desplazamiento), $r(t)$ es la parte "real" o "detrended" de la señal (sin la tendencia). La función detrend elimina la tendencia lineal $mt + b$, devolviendo solo $r(t)$.

La tendencia lineal de una señal se ajusta usando una regresión lineal. Para un conjunto de datos $x = \{x_1, x_2, \dots, x_n\}$, la tendencia lineal se ajusta con el modelo:

$$x^{\wedge}(t) = mt + b$$

Donde m y b se estiman utilizando el método de mínimos cuadrados.

Dado que la señal tiene n puntos, podemos usar el enfoque matricial de mínimos cuadrados para estimar m y b . Definimos la matriz de diseño A :

$$A = \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \dots & \dots \\ 1 & t_n \end{bmatrix}$$

Donde t_1, t_2, \dots, t_n son los tiempos o índices de las muestras, y el vector de observaciones x es:

$$x = [x_1, x_2, \dots, x_n]$$

El problema de mínimos cuadrados se resuelve con la ecuación normal:

$$(b, m) = (A^t A)^{-1} A^t x$$

Esto proporciona los coeficientes b y m que mejor ajustan la tendencia lineal de los datos. Una vez estimados m y b , la señal sin tendencia se obtiene restando la tendencia ajustada de los datos originales:

$$r(t) = x(t) - (mt + b)$$

Donde $r(t)$ es la señal sin tendencia, que es lo que devuelve `scipy.signal.detrend` [7].

Funcionamiento del filtro Hampel: El filtro Hampel es una técnica de procesamiento de señales para detectar y eliminar los outliers (valores atípicos) en una serie temporal o secuencia de datos. Se basa en el uso de una ventana deslizante y una medida robusta de la dispersión para identificar valores que se desvían significativamente de la tendencia local de los datos.

Desarrollo matemático: Dado un conjunto de datos $x = \{x_1, x_2, \dots, x_n\}$, el filtro Hampel evalúa cada dato x_i dentro de una ventana deslizante de tamaño w centrada en x_i . El algoritmo aplica los siguientes pasos:

Para cada valor en la serie x_i , se considera una ventana de tamaño w alrededor de x_i , es decir, los valores:

$$Ventana(x_i) = \{x_i - w, x_i - w + 1, \dots, x_i + w\}$$

En esta ventana, los índices negativos y mayores a n (extremos) pueden ser tratados de manera especial (por ejemplo, ignorados o rellenados).

Se calcula la mediana de los valores dentro de la ventana:

$$x_{\sim i} = mediana(\{x_i - w, \dots, x_i + w\})$$

La mediana es una medida robusta de la tendencia central, lo que significa que es resistente a valores atípicos dentro de la ventana.

Se calcula la Desviación Absoluta de la Mediana (MAD) dentro de la ventana. Esto proporciona una medida robusta de la dispersión de los valores:

$$MAD = mediana(|x_j - x_{\sim i}|), j \in \{i - w, \dots, i + w\}$$

La MAD es robusta a los valores extremos, a diferencia de la desviación estándar, lo que la hace adecuada para detectar outliers.

Se define un umbral de decisión para identificar valores atípicos. Si el valor absoluto de la diferencia entre x_i y la mediana $x_{\sim i}$ es mayor que un múltiplo k de la MAD, se considera un outlier. La condición es:

$$|x_i - x_{\sim i}| > k \cdot MAD$$

Donde k es un parámetro ajustable que determina cuán estricto es el filtro. Típicamente, $k = 3$ es una elección común, que significa que cualquier valor que se desvíe más de tres veces la MAD será tratado como un valor atípico.

Si x_i es detectado como un outlier (es decir, si cumple la condición anterior), entonces puede ser reemplazado por la mediana local $x_{\sim i}$:

$$x_i \leftarrow x_{\sim i}$$

Esto elimina el valor atípico y lo reemplaza por una estimación basada en los datos vecinos [8].

Entropía de permutación: La entropía de permutación (PE, por sus siglas en inglés: Permutation Entropy) es una medida no lineal que cuantifica la complejidad y el grado de desorden en una serie temporal. Introducida por Bandt y Pompe en 2002 [2], se ha utilizado ampliamente para analizar datos en diversos campos, como la bioingeniería, debido a su simplicidad, bajo costo computacional y capacidad para capturar patrones de orden y desorden en señales fisiológicas.

Desde la perspectiva bioingeniería, la entropía de permutación es útil para analizar la dinámica de señales fisiológicas, como electrocardiogramas (ECG), señales cerebrales (EEG), y patrones de marcha, al proporcionar una

forma de detectar anomalías o cambios patológicos en el comportamiento del sistema biológico.

Desarrollo matemático: Dada una serie temporal $\{x_1, x_2, \dots, x_n\}$ la entropía de permutación mide el grado de aleatoriedad en los patrones de la secuencia considerando el orden relativo entre los elementos de la serie.

Primero, se reconstruye la serie temporal en forma de vectores de dimensión m llamados vectores de retardo embebido. Estos vectores representan subsecuencias locales de la serie. Para una serie x_t , se construyen vectores de longitud m usando un retardo τ , de la siguiente manera:

$$x_t = [x_t, x_{t+\tau}, x_{t+2\tau}, \dots, x_{t+(m-1)\tau}]$$

Aquí, m es el orden de la permutación (número de puntos en cada subsecuencia) y τ es el retardo temporal.

Cada subsecuencia x_t es transformada en una permutación que describe el orden relativo de sus valores. Por ejemplo, si para un vector de tres elementos $[x_2, x_5, x_3]$, el orden relativo de los elementos es $x_2 < x_3 < x_5$, entonces el patrón correspondiente será (0, 2, 1). Hay $m!$ posibles permutaciones para cada subsecuencia de m elementos.

Se calcula la frecuencia de cada permutación en la serie completa. Denotemos por p_i la proporción de veces que la i -ésima permutación ocurre en la serie. Como hay $m!$ permutaciones posibles, la probabilidad asociada a cada permutación es:

$$p_i = (\text{N}^\circ \text{ de veces que aparece la permutación } i) / n - (m - 1)\tau$$

La entropía de permutación se define utilizando la fórmula de la entropía de Shannon aplicada a la distribución de probabilidades de las permutaciones:

$$HPE = \sum_{i=1}^m p_i \cdot \log(p_i)$$

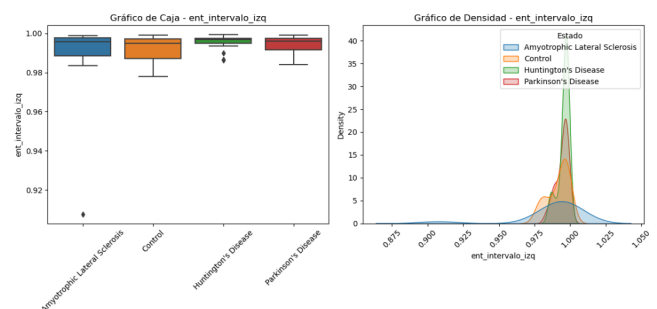
Sumatoria desde $i=1$ hasta $m!$

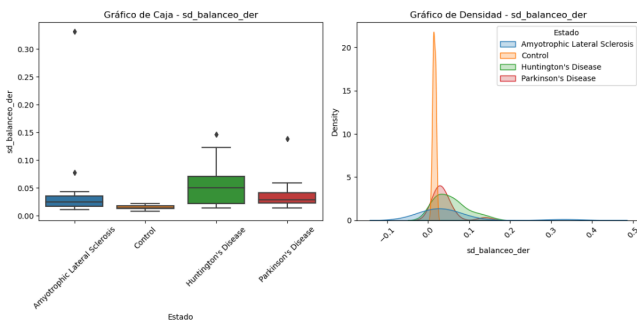
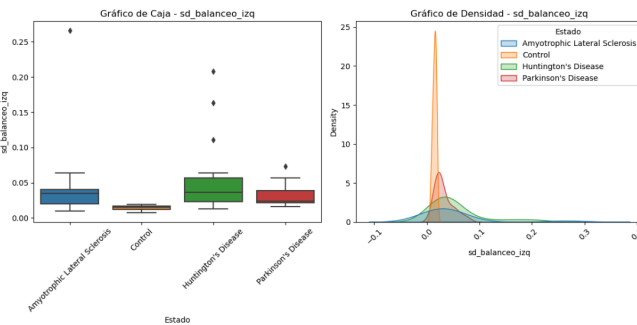
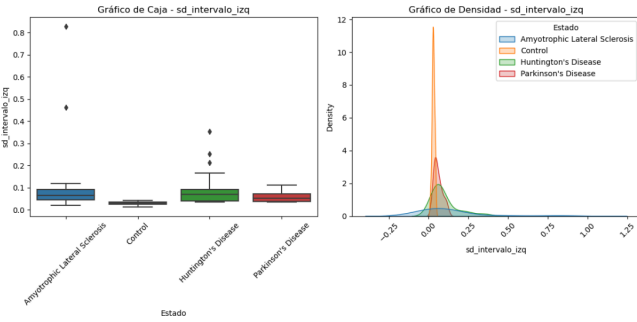
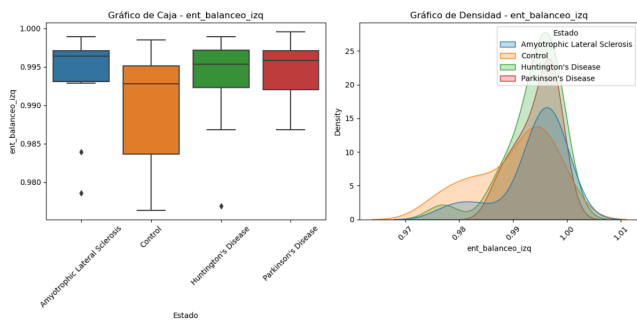
Este valor de entropía está normalizado entre 0 y $\log(m!)$, donde:

- Un valor cercano a 0 indica una secuencia altamente predecible y ordenada.
- Un valor cercano a $\log(m!)$ indica una secuencia completamente aleatoria.

4. Resultados y discusión

a. Análisis estadístico descriptivo





En el presente estudio se analizaron señales extraídas del ciclo de marcha de sujetos con enfermedades neurodegenerativas, estas incluyen Esclerosis lateral amiotrófica, síndrome huntington y síndrome de parkinson, así como un grupo de control compuestos por sujetos sanos. El estudio se centró en analizar 4 momentos clave del ciclo de marcha, el intervalo de zancada izquierda, intervalo de zancada derecha, intervalo de balanceo izquierdo y el intervalo de balanceo derecho. Con el fin de obtener resultados más precisos se aplicaron rutinas de filtrado de trend y hampel con el fin de eliminar tendencias lineales atípicas y valores atípicos a la señal que pueden ser originados por ruido o factores en la adquisición de la señal.

Dentro de los gráficos del coeficiente de variación(CV) y de desviación estándar(SD) de la zancada y el balanceo, es

posible observar una mayor variabilidad presente en los grupos de sujetos con enfermedades neurodegenerativas, especialmente aquellos diagnosticados con parkinson y esclerosis lateral amiotrófica; esta mayor variabilidad sugiere que estas afecciones influyen de manera significativa en el control del movimiento. Además, la presencia de valores atípicos indica que los individuos presentan un control motor mucho más inestable, lo que podría implicar deterioro de la coordinación y estabilidad durante el ciclo de marcha. Estos hallazgos permiten entender en cierta medida el impacto de las enfermedades neurodegenerativas durante procesos considerados simples como lo es caminar.

Las gráficas de entropía revelan que los grupos con síndrome de Huntington y esclerosis lateral amiotrófica presentan una mayor variabilidad en sus patrones de movimiento. Dicha variabilidad podría estar relacionada directamente con el deterioro neuromuscular y las alteraciones en el control motor que caracterizan este tipo de enfermedades. Para el caso del síndrome de Huntington el incremento en la entropía de las señales de marcha podrían explicarse con las disfunciones motoras y cognitivas propias de esta afección. Por otro lado los grupos de control y parkinson muestran una menor variabilidad en la entropía; lo cual podría indicar que aunque el síndrome de parkinson está asociado con síntomas motores evidentes, los mecanismos de control motor podrían estar más regulados o adaptación y compensaciones en el control del movimiento.

Las gráficas de entropía de permutación por grupo, muestran una diferenciación clara en la complejidad de las señales de marcha entre los grupos de estudio. Las entropías más altas se aprecian en los sujetos con esclerosis lateral amiotrófica, lo que indica una mayor variabilidad y pérdida de los patrones de marcha. Por otro lado, los pacientes con síndrome de Huntington y parkinson también presentan niveles de entropía elevados en comparación con el grupo de control, pero menos elevados que el grupo con esclerosis lateral amiotrófica. Esto sugiere que, a pesar de que el control neuromotor también está afectado, aún conservan cierta capacidad de preservación motora.

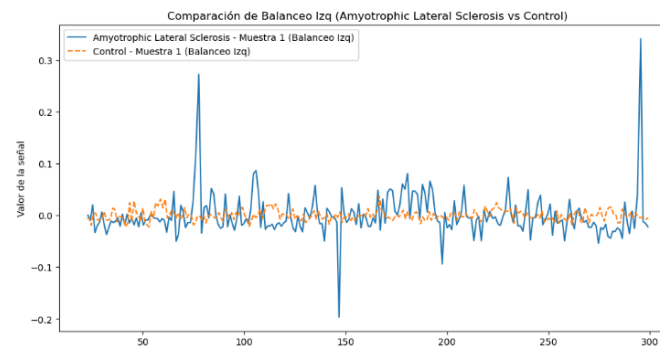


Figura 1: Comparación de sujeto con ALS y un sujeto sano; durante la fase de balanceo izquierdo.

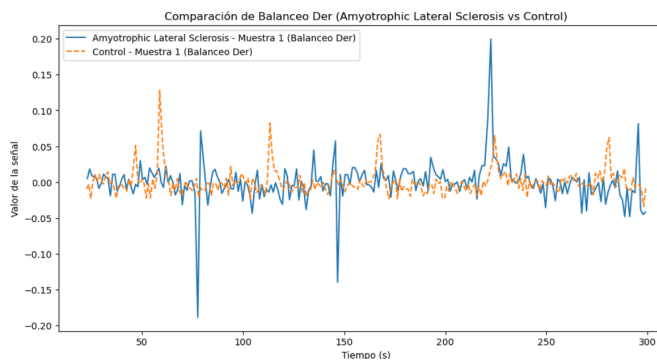


Figura 2: Comparación de sujeto con ALS y un sujeto sano; durante la fase de balanceo derecho.

En las imágenes se puede observar cómo el paciente con Esclerosis Lateral Amiotrófica (ALS) muestra una varianza mayor en los valores de la señal a lo largo del tiempo comparado con un sujeto sano, dicha variabilidad se presenta con picos pronunciados en zonas donde en un sujeto sano no deberían presentarse, cabe aclarar que los picos más pronunciados o con cambios más abruptos pueden ser ruido provocado por interferencia con los sensores o algún movimiento brusco adicional a la marcha. Estos pronunciamientos pueden interpretarse como disfunciones en el control neuromuscular, afección que suele estar asociada con la pérdida de la coordinación motora.

Las imágenes denotan cómo un sujeto sano presenta una mayor estabilidad, respecto al comportamiento y amplitud de los picos y a los patrones que estos presentan en las distintas fases del balanceo. Que las señales presenten valores más cercanos a cero, y además tengan patrones repetitivos y no erráticos; podrían tomarse como un indicador de un control neuromotor más eficiente y predecible durante el balanceo. Las amplitudes de los picos alcanzadas por el sujeto con ALS, sugieren activación extrema del sistema motor durante períodos cortos de tiempo pero repetitivos, lo que podría derivarse en espasmos propios de condiciones neurodegenerativas.

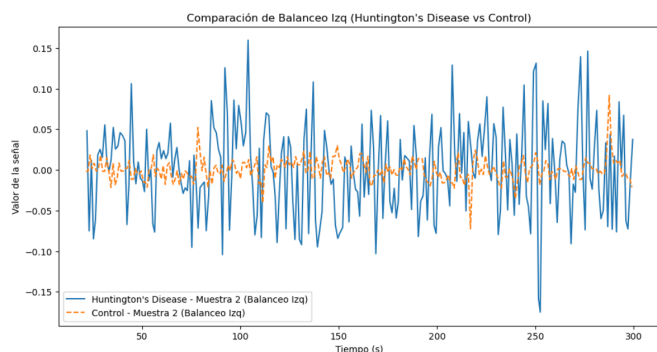


Figura 3: Comparación de sujeto con Huntington 's disease y un sujeto sano; durante la fase de balanceo izquierdo.

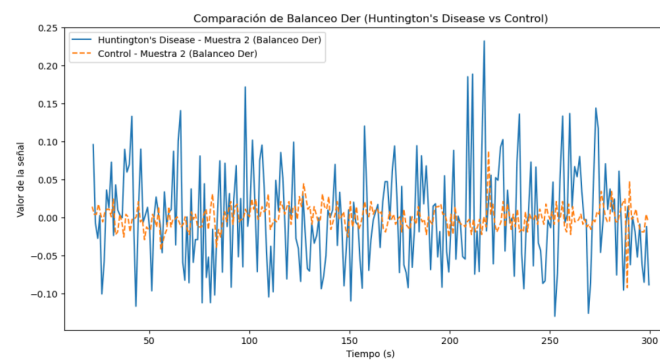


Figura 4: Comparación de sujeto con Huntington 's disease y un sujeto sano; durante la fase de balanceo derecho.

Las imágenes permiten observar cómo el sujeto con síndrome de Huntington presenta una mayor varianza en los valores de su señal a lo largo del tiempo comparado con un sujeto sano. En este caso la variabilidad de los picos oscilan entre -0.15 y 0.15 intermitentemente, estos cambios abruptos e ininterrumpidos indican una activación neuromotora errática; esto podría ser un indicador de pérdida del equilibrio durante las fases de balanceo, una condición propia de dicho síndrome [5].

Las imágenes denotan cómo el sujeto sano posee patrones definidos en la señal por su fisiología durante el ciclo de marcha, donde no se presentan oscilaciones abruptas, mostrando una señal más estable que se mantiene cerca a 0.

Las diferencias en la estabilidad de la señal durante la fase de balanceo reflejan el estado de las características motoras de cada individuo. Permitiendo determinar de forma cualitativa el nivel de deterioro que presenta un sujeto con Huntington frente a un sujeto sano.

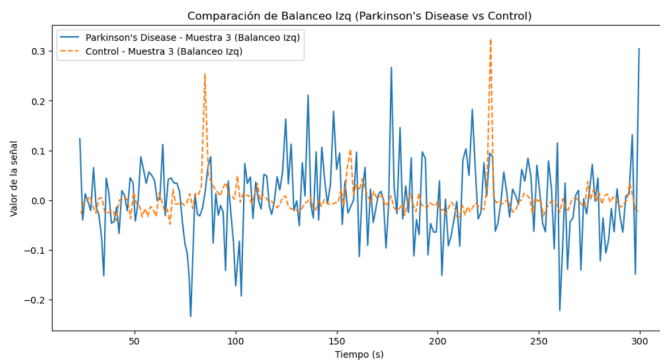


Figura 5: Comparación de sujeto con Parkinson 's disease y un sujeto sano; durante la fase de balanceo izquierdo.

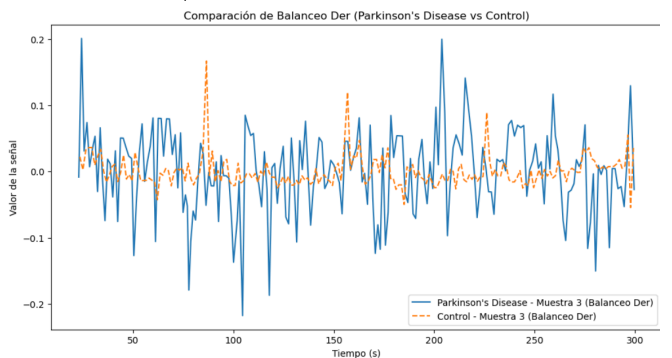


Figura 6: Comparación de sujeto con Parkinson 's disease y un sujeto sano; durante la fase de balanceo izquierdo.

Las imágenes permiten observar cómo el sujeto con síndrome de Parkinson presenta una mayor varianza comparada con el

sujeto sano, en este caso la variabilidad no se hace presente solamente en los cambios abruptos de las oscilaciones sino también en la duración de los picos que en muchos casos se convierten en valles indicando una activación continua en ciertos momentos del balanceo. Si bien hay cambios abruptos en los valores en lo cuales oscila la señal, se puede observar que no llega a ser errática del todo, ya que presenta ciertos patrones de activación motora; estos patrones indican espasmos repetitivos en determinados tramos, un comportamiento esperado para un sujeto con síndrome de Parkinson [4].

Las imágenes indican que el sujeto sano presenta patrones y fluctuaciones en su señal que son característicos de su fisiología durante el ciclo de marcha, y en ciertos casos se hace presente ruido generado por los electros al capturar la señal. En cambio, el patrón del sujeto con síndrome de Parkinson muestra activaciones motoras abruptas que son consecuencia de su afección neuromotora y en ciertos casos se presentan picos que son generados posiblemente por ruido externo al estudio, movimientos adicionales al ciclo de marcha, sudor o algún registro del sensor que no involucra directamente la marcha.

b. Para determinar el tipo de prueba que se realizará, se realizan pruebas estadísticas para determinar qué supuestos cumplen los datos a analizar.

Hipótesis para normalidad

- Hipótesis nula (H_0): Los datos obtenidos de la señal segmentada siguen una distribución normal
- Hipótesis alternativa (H_1): Los datos obtenidos de la señal segmentada no siguen una distribución normal

Hipótesis para varianza

- Hipótesis nula (H_0): Las varianzas de los datos comparados presentan homocedasticidad
- Hipótesis alternativa (H_1): Los varianzas de los datos comparados no presentan homocedasticidad

Hipótesis prueba paramétrica (t-student)

- Hipótesis nula (H_0): No hay diferencia significativa entre la señal del sujeto con una afección neuromotora y un sujeto sano.
- Hipótesis alternativa (H_1): Hay diferencia significativa entre la señal del sujeto con una afección neuromotora y un sujeto sano.

Hipótesis prueba no paramétrica (Prueba U de Mann-Whitney)

- Hipótesis nula (H_0): Las distribuciones de ambos grupos son iguales. En otras palabras, no hay diferencia en la ubicación (mediana) de las distribuciones de los dos grupos.
- Hipótesis alternativa (H_1): Las distribuciones de los dos grupos son diferentes.

Cabe aclarar que el valor p-values debe ser mayor al 5% o 0.05 para aceptar la hipótesis nula, si es menor a este valor se rechaza y se acepta la hipótesis alternativa. Según los resultados de los datos al aplicar las pruebas de normalidad y varianza para determinar qué hipótesis se cumple se procederá a aplicar una prueba paramétrica T-Test o una prueba no paramétrica Mann-Whitney U.

Comparación de parámetros estadísticos entre el grupo ALS y el grupo control:

Se evaluaron diversas métricas relacionadas con la marcha, utilizando como referencia un grupo control y un grupo de pacientes con Esclerosis Lateral Amiotrófica (ALS). Las métricas estudiadas incluyen la entropía de permutación (ent), el coeficiente de variación (CV) y la desviación estándar (SD), aplicadas a los intervalos de tiempo de los pies (izquierdo y derecho) y a las fases de balanceo.

```
Columna: ent_intervalo_izq
Prueba realizada: Mann-Whitney U
Estadístico: 82.0
p-value: 0.9183089345454835
Normalidad Control (p-value): 0.00780416870253563
Normalidad Enfermo (p-value): 7.3873195266429745e-06
Homogeneidad de varianzas (p-value): 0.4740701217800135
-----
Columna: ent_intervalo_der
Prueba realizada: Mann-Whitney U
Estadístico: 103.0
p-value: 0.3559671197337311
Normalidad Control (p-value): 0.15117337507595474
Normalidad Enfermo (p-value): 1.0234799596120573e-05
Homogeneidad de varianzas (p-value): 0.16538463074427698
```

Figura 7: Comparación de la entropía de permutación del grupo ALS y el grupo control en el intervalo de paso izquierdo y derecho.

La elección de la prueba de Mann-Whitney U fue fundamental debido a la falta de normalidad en los datos de ambos grupos, tal como lo indican los valores de normalidad significativamente bajos (p-value < 0.05). La normalidad de los datos es un requisito esencial para aplicar pruebas paramétricas como la t de Student, y en este caso, su incumplimiento justificó el uso de una prueba no paramétrica. En las variables donde se rechazó la hipótesis nula, como la desviación estándar del intervalo de tiempo (SD_intervalo_izq y SD_intervalo_der) y del balanceo (SD_balanceo_izq y SD_balanceo_der), se concluye que existen diferencias significativas en la variabilidad de la marcha entre los grupos. Esto implica que los pacientes con ALS muestran una mayor dispersión en los intervalos de tiempo de los pasos y en las fases de balanceo en comparación con el grupo control, lo cual puede ser indicativo de alteraciones motoras y de un control menos preciso de la marcha.

Por otro lado, en las métricas donde no se rechazó la hipótesis nula, como la entropía de los intervalos de tiempo y balanceo, la prueba sugiere que no hay diferencias estadísticamente significativas entre ambos grupos. Esto implica que, aunque la ALS afecta la variabilidad de los movimientos, la complejidad y la estructura de los patrones temporales de la marcha se mantienen relativamente similares entre los pacientes con ALS y los individuos sanos. En estas situaciones, la prueba no proporciona suficiente evidencia para afirmar que la ALS introduce un cambio sustancial en la regularidad de la marcha en términos de la entropía de permutación.

En términos del análisis estadístico, un valor-p bajo (menor a 0.05) junto con un estadístico bajo en la prueba de Mann-Whitney U sugiere que las diferencias entre los grupos son estadísticamente significativas, indicando que la mediana de las distribuciones en el grupo control y el grupo con ALS difiere notablemente. Esto fue evidente en las variables relacionadas con la desviación estándar y el coeficiente de variación de ciertos aspectos de la marcha. En cambio, valores-p altos sugieren que no hay suficientes diferencias entre los grupos para rechazar la hipótesis nula, como se

observó en las métricas de entropía para los intervalos de tiempo y balanceo. Según los resultados en las comparaciones, es claramente notorio que la marcha de los pacientes con ALS es más variable y menos constante que la de los individuos sanos, especialmente en términos de la fase de balanceo y los intervalos entre pasos. Las diferencias significativas observadas en la desviación estándar y el coeficiente de variación refuerzan la importancia de estas métricas como potenciales indicadores de alteraciones motoras en pacientes con ALS.

Comparación de parámetros estadísticos entre el grupo Huntington y el grupo control: El objetivo es determinar si existen diferencias en las características de la marcha entre ambos grupos utilizando diferentes métricas: entropía de permutación (ent), coeficiente de variación (CV) y desviación estándar (SD). Para esta comparación, se utilizó la prueba de Mann-Whitney U, una prueba estadística no paramétrica adecuada para datos que no siguen una distribución normal.

El valor-p obtenido en cada prueba es crucial para determinar si las diferencias observadas entre ambos grupos son significativas o si podrían ser el resultado de la variabilidad natural de la muestra. En este contexto, un valor p menor a un umbral convencional de 0.05 indica que la diferencia observada es estadísticamente significativa y que es poco probable que haya ocurrido por azar. Esto lleva al rechazo de la hipótesis nula (que sostiene que no hay diferencias entre los grupos) y sugiere que existe una diferencia real entre las características de la marcha en individuos sanos y aquellos con Huntington.

```
-----
Columna: sd_intervalo_izq
Prueba realizada: Mann-Whitney U
Estadístico: 4.0
p-value: 4.085965995779528e-05
Normalidad Control (p-value): 0.99680322816484
Normalidad Enfermo (p-value): 0.0006936579168083647
Homogeneidad de varianzas (p-value): 0.04601522899616375
-----
Columna: sd_intervalo_der
Prueba realizada: Mann-Whitney U
Estadístico: 3.0
p-value: 3.269330273511375e-05
Normalidad Control (p-value): 0.9914919922898477
Normalidad Enfermo (p-value): 0.000181611475973102
Homogeneidad de varianzas (p-value): 0.09128979328273032
```

Figura 8: Comparación de la desviación estándar del grupo Huntington y el grupo control en el intervalo de paso izquierdo y derecho.

Por ejemplo, los valores p obtenidos para la desviación estándar del intervalo de tiempo del pie izquierdo y derecho (SD_intervalo_izq y SD_intervalo_der), así como del balanceo (SD_balanceo_izq y SD_balanceo_der), fueron inferiores a 0.05. Esto indica que la variabilidad en la duración de los pasos y las fases de balanceo es significativamente mayor en los pacientes con Huntington en comparación con los controles. La mayor variabilidad de la marcha en los pacientes sugiere una inestabilidad y dificultad en el control motor, que son características típicas de la progresión de la enfermedad. Esto coincide con lo que se observa clínicamente, ya que la enfermedad de Huntington se asocia con movimientos descoordinados y fluctuantes.

Por otro lado, algunas métricas como la entropía de los intervalos de tiempo y de los periodos de balanceo (ent_intervalo_izq, ent_intervalo_der, ent_balanceo_izq y

ent_balanceo_der) no mostraron diferencias significativas entre los grupos, al tener valores p superiores a 0.05. Esto sugiere que, aunque la enfermedad de Huntington altera la estabilidad de la marcha, no necesariamente modifica la complejidad de los patrones temporales de manera diferente a lo que se observa en individuos sanos. Esto podría indicar que la enfermedad afecta principalmente la regularidad y consistencia de la marcha, más que su estructura temporal subyacente. Las diferencias observadas en la desviación estándar pueden tener implicaciones importantes en el manejo de la enfermedad. La mayor variabilidad en la marcha podría utilizarse como un indicador del grado de deterioro motor en los pacientes con Huntington. En la práctica clínica, esto puede ayudar a monitorear la progresión de la enfermedad y ajustar los tratamientos de rehabilitación para mejorar la estabilidad y el control motor.

Comparación de parámetros estadísticos entre el grupo Parkinson y el grupo control: En cuanto a la comparación de los parámetros estadísticos del grupo Parkinson con el grupo control se obtiene los siguientes resultados:

```
Columna: ent_balanceo_izq
Prueba realizada: t-test
Estadístico: -2.0997533125379224
p-value: 0.04644515667571262
Normalidad Control (p-value): 0.19994492545922515
Normalidad Enfermo (p-value): 0.46285788915512693
Homogeneidad de varianzas (p-value): 0.07056058950226324
-----
Columna: ent_balanceo_der
Prueba realizada: Mann-Whitney U
Estadístico: 56.0
p-value: 0.15103152823115532
Normalidad Control (p-value): 0.009190278974801437
Normalidad Enfermo (p-value): 0.5857985086869787
Homogeneidad de varianzas (p-value): 0.07547915595233946
```

Figura 9: Comparación de la entropía de permutación del grupo Parkinson y el grupo control en el intervalo balanceo izquierdo y derecho.

En la comparación de la entropía de los intervalos de tiempo (ent_intervalo_izq y ent_intervalo_der) entre el grupo control y los pacientes con Parkinson, los valores p obtenidos fueron 0.608 y 0.959, respectivamente, lo cual es considerablemente mayor a 0.05. Esto indica que no hay evidencia estadística para afirmar que la complejidad en la variabilidad temporal de la marcha difiere entre ambos grupos. A pesar de las alteraciones motoras en la enfermedad de Parkinson, los patrones temporales de la marcha no muestran una variación significativa en términos de complejidad respecto al grupo control. Sin embargo, en la entropía del balanceo izquierdo (ent_balanceo_izq), el análisis arrojó un valor p de 0.046, ligeramente inferior al umbral de 0.05, indicando una diferencia significativa entre ambos grupos. Esto sugiere que la fase de balanceo del pie izquierdo es menos regular en los pacientes con Parkinson, lo que podría ser un reflejo de la bradicinesia y la rigidez muscular, síntomas característicos de la enfermedad.

Un hallazgo notable se observa en las métricas de la desviación estándar (SD). Los valores p para la desviación estándar del intervalo de tiempo y del balanceo (SD_intervalo_izq, SD_intervalo_der, SD_balanceo_izq y SD_balanceo_der) fueron significativamente menores a 0.05, indicando diferencias relevantes entre los grupos. Esto demuestra que los pacientes con Parkinson presentan una mayor variabilidad en la duración de sus pasos y fases de

balanceo en comparación con los individuos sanos. Esta mayor variabilidad es característica de la inestabilidad de la marcha asociada con la enfermedad de Parkinson, donde los movimientos pueden ser menos predecibles y consistentes. Una mayor variabilidad en la marcha podría aumentar el riesgo de caídas y accidentes en los pacientes con Parkinson.

c. Comparación de resultados obtenidos en el artículo base contra los resultados obtenidos en este estudio.

Antes de iniciar la comparación, es importante señalar que el estudio original realiza un análisis de los datos de marcha usando un parámetro de filtro diferente al usado en este estudio y que la función detrend es usada de forma diferente en el artículo base. Esto podría explicar alguna de las diferencias observadas entre ambos estudios.

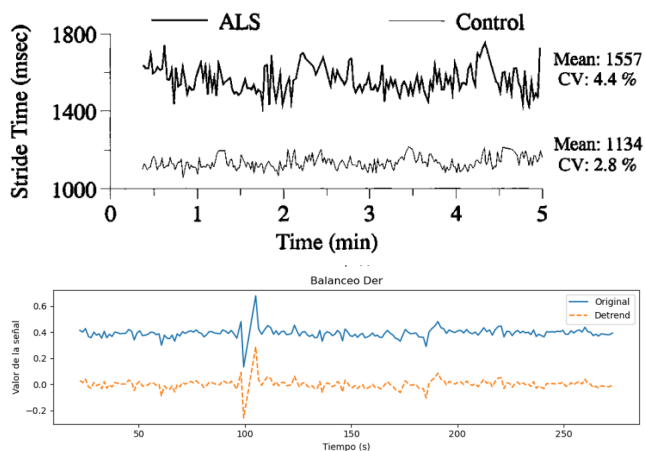


Figura 10: Comparación de la señal usada en el estudio base contra la señal filtrada usada en este estudio [10].

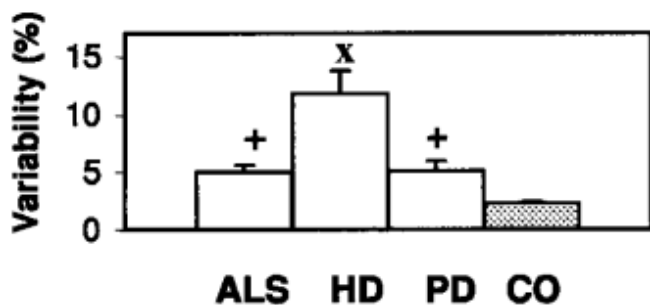


Figura 11: Variabilidad obtenida en el estudio base [10]

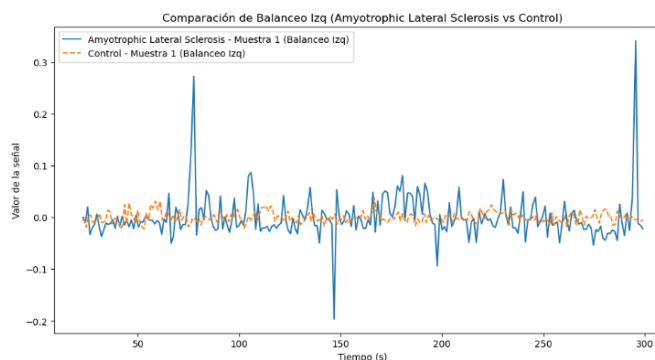


Figura 12: Variabilidad obtenida de un sujeto con ALS.

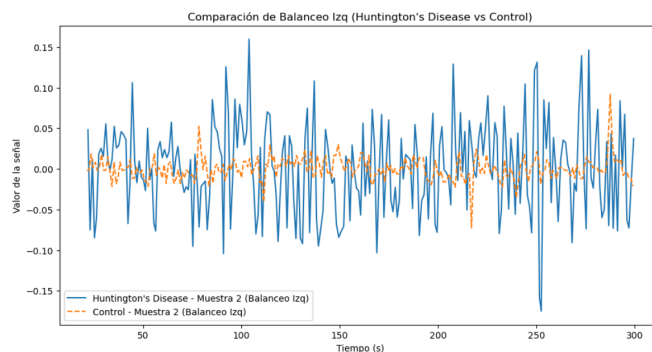


Figura 13: Variabilidad obtenida de un sujeto con síndrome de Huntington.

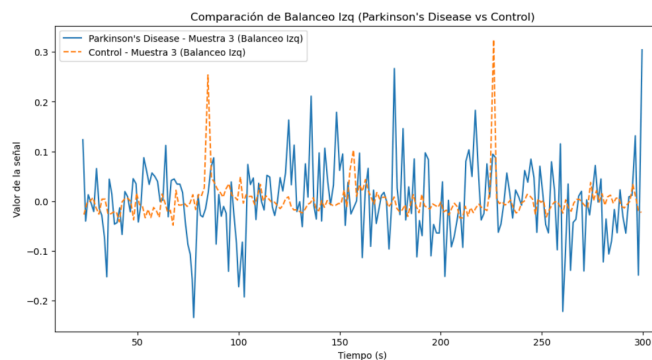


Figura 14: Variabilidad obtenida de un sujeto con síndrome de Parkinson.

Las imágenes permiten observar cómo los resultados obtenidos en el estudio base y los resultados obtenidos en este estudio tienen cierto grado de semejanza. Ya que en ambos estudios se observa cómo los sujetos con síndrome de Huntington son los que presentan impulsos motores más erráticos denotado así en la variabilidad de los resultados que estos pueden tener. Si bien en el estudio base los pacientes con ALS y síndrome de Parkinson tienen una variabilidad casi igual; el estudio actual permite denotar mejor que los sujetos con síndrome de parkinson denotan una variabilidad mayor en la activación neuromotora que los sujetos con ALS. Además que se logra observar un patrón no errático en los sujetos con síndrome de Parkinson.

Al comparar los resultados de las pruebas realizadas entre el grupo control y los grupos de Parkinson, ALS y Huntington con los hallazgos del artículo base, se destacan varias diferencias significativas en los patrones de marcha. En el artículo base, los sujetos con ALS mostraron un tiempo de stride más largo y una velocidad de marcha significativamente menor en comparación con los controles, lo que no se corroboró en los resultados de la comparación entre el grupo control y los pacientes con Parkinson o Huntington, donde los valores de p en las pruebas t para las variables como "ent_intervalo_izq" ($p = 0.213$), "ent_balanceo_der" ($p = 0.243$) y "sd_balanceo_der" ($p = 0.179$) indicaron diferencias no significativas. Por otro lado, en la comparación entre el grupo control y el grupo ALS, la prueba para "sd_intervalo_izq" mostró una diferencia significativa ($p = 0.015$), algo que no se observó en las comparaciones con Parkinson. Además, mientras que el artículo base reportó un incremento en la variabilidad del stride en pacientes con ALS, en las comparaciones con el grupo control se encontró que las diferencias en las variables "cv_intervalo_izq" y "cv_balanceo_der" no fueron significativas ($p = 0.506$ y $p = 0.4$, respectivamente), lo que sugiere que la irregularidad en la marcha observada en el artículo podría no replicarse en los nuevos análisis. Esto resalta la complejidad y variabilidad en los déficits de marcha entre las distintas

patologías neurodegenerativas y sugiere que las características de la marcha pueden ser influenciadas por factores específicos de cada enfermedad, lo que no se reflejó en los análisis de comparación de los grupos en esta ocasión, posiblemente debido a las limitaciones del análisis comentadas más adelante en el apartado de conclusiones.

d. Para analizar las diferencias entre grupos de poblaciones con padecimientos crónicos y grupos de control conformados por individuos sanos, es esencial utilizar herramientas estadísticas que permitan interpretar los datos de manera precisa. Las medidas como la media, desviación estándar, y medidas de sensibilidad y especificidad son fundamentales para evaluar la distribución que presentan los datos inicialmente, pero también para evaluar la efectividad de las pruebas paramétricas y diagnósticas, con el fin de determinar la relación entre las variables de estudio de los grupos de estudio.[11]

Teniendo en cuenta la relevancia de ciertas medidas en el estudio realizado se permite determinar que las medidas más significativas para discriminar poblaciones de controles clínicos son:

- **Desviación estándar:** Esta medida evalúa la variabilidad de los datos, indicando cuánto se desvían los valores del promedio. Una mayor desviación estándar en las poblaciones clínicas puede reflejar un comportamiento más errático en la marcha o en otros parámetros relevantes, lo que podría ser indicativo de la gravedad de la enfermedad [12].
- **Media:** Como medida de tendencia central, la media permite identificar el comportamiento general de una población. Comparar las medias entre grupos clínicos y controles puede evidenciar diferencias significativas que ayuden a entender los efectos de las patologías en las características funcionales [12].
- **Análisis de varianza (ANOVA):** Este método permite comparar las medias de tres o más grupos y determinar si hay diferencias significativas entre ellos. Es especialmente útil en estudios donde se evalúan múltiples grupos clínicos [13].
- **Medidas de sensibilidad y especificidad:** Estas métricas son esenciales para evaluar la precisión de las pruebas diagnósticas y determinar la capacidad de los parámetros analizados para identificar correctamente a los individuos con y sin la condición clínica [14].
- **D de Cohen:** Esta medida de tamaño del efecto permite cuantificar la magnitud de la diferencia entre grupos, lo que ayuda a entender no solo si hay una diferencia significativa, sino también cuán relevante es en términos prácticos [15].
- **Análisis de regresión:** Permite evaluar la relación entre variables, controlando otros factores que pueden influir en los resultados, y así identificar predictores significativos de los resultados clínicos [16].

5. Conclusiones

Los resultados sugieren que las diferencias más pronunciadas en los parámetros de marcha se observan en los pacientes con ALS, con mayores niveles de variabilidad en la marcha, lo cual es consistente con la naturaleza progresiva de la enfermedad y la afectación a las neuronas motoras. En el grupo de Parkinson, aunque no se detectaron diferencias significativas en la mayoría de los parámetros, la mayor variabilidad en el balanceo derecho podría ser indicativa de problemas de control motor que afectan de manera más sutil la estabilidad y la coordinación. El hecho de que se haya rechazado la hipótesis nula para ciertas métricas de dispersión como la desviación estándar y el coeficiente de variación, pero no para otras como la entropía, sugiere que la ALS afecta principalmente la consistencia temporal de la marcha, más que la complejidad de los patrones temporales. Estos hallazgos podrían ser de gran utilidad para el diagnóstico y seguimiento de la enfermedad, ya que permiten identificar patrones de marcha característicos de la ALS y podrían contribuir a la evaluación de la progresión de la enfermedad.

En cuanto a Huntington, los resultados son menos concluyentes, pero los indicios de mayor variabilidad en el balanceo podrían estar relacionados con la ataxia y los movimientos involuntarios propios de la enfermedad. Sin embargo, estos efectos no fueron significativos en este análisis. El análisis de la marcha entre el grupo control y los pacientes con Huntington revela que, si bien la complejidad de la marcha medida por la entropía no presenta diferencias significativas, la variabilidad de los movimientos sí lo hace. Esto subraya el impacto de la enfermedad en la estabilidad del caminar. Los resultados no solo permiten entender mejor cómo la enfermedad de Huntington afecta la movilidad de los pacientes, sino que también pueden guiar estrategias terapéuticas y de intervención que apunten a reducir la inestabilidad durante la marcha.

También, es importante señalar algunas limitaciones de este análisis que podrían haber afectado los resultados e interpretaciones. En primer lugar, el tamaño reducido de la muestra puede haber disminuido la potencia estadística, limitando la capacidad de detectar diferencias significativas entre los grupos. Un tamaño de muestra mayor (aunque a un mayor costo operacional) podría haber permitido identificar con mayor claridad patrones y relaciones significativas que, en este análisis, pueden haberse pasado por alto. Además, aunque la base de datos incluye información sobre el grado de desarrollo o la gravedad de la enfermedad en los pacientes, estos factores no fueron considerados en este estudio. La inclusión de estos datos podría haber proporcionado una visión más detallada y precisa sobre la variabilidad en los resultados dentro de los grupos de pacientes. Al no considerar la gravedad de la enfermedad, es posible que se hayan agrupado individuos con diferentes grados de afectación, lo que podría haber contribuido a una mayor variabilidad en las mediciones y haber enmascarado diferencias relevantes entre los grupos.

Asimismo, el uso de pruebas como el t-test y la prueba de Mann-Whitney, aunque adecuadas para los tipos de datos y distribuciones observadas, tienen sus propias limitaciones. El supuesto de normalidad en los datos fue evaluado y, aunque en muchos casos se cumplió, en otros no se observó una

distribución normal, lo que motivó el uso de pruebas no paramétricas. Sin embargo, la dependencia de estas pruebas puede haber influido en la capacidad para detectar diferencias finas en los datos.

En futuras investigaciones, sería valioso considerar la posibilidad de analizar subgrupos basados en la gravedad de la enfermedad y ampliar el tamaño de la muestra para mejorar la robustez y precisión de los resultados. También sería recomendable incorporar enfoques estadísticos más avanzados, como modelos mixtos o análisis multivariados, que permitan captar mejor la complejidad de los datos y las interacciones entre diferentes factores.

Finalmente, aunque el presente análisis ha permitido identificar algunas diferencias y similitudes en las mediciones de los diferentes grupos, se sugiere profundizar en estos resultados con estudios complementarios que consideren otras variables relevantes, como la medicación, el tiempo de diagnóstico y el progreso de la enfermedad, así como también agrandar la base de datos, aumentando la cantidad de la población de todos los sujetos para tener gran variedad de muestras y analizar detalles o patrones que en muestras pequeñas no sea posible observar, esto en pro de obtener una comprensión más completa de las diferencias entre los grupos control y las enfermedades neurodegenerativas analizadas.

6. Referencias

- [1] "Gait in Neurodegenerative Disease Database". PhysioNet. [En línea]. Disponible: <https://physionet.org/content/gaitndd/1.0.0/>
- [2] "Permutation Entropy | Aptech". GAUSS Software - Fastest Platform for Data Analytics | Aptech. [En línea]. Disponible: <https://www.aptech.com/blog/permutation-entropy/>
- [3] "Entropy analysis in gait research – methodological considerations and recommendations". PubMed Central (PMC). [En línea]. Disponible: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8051436/>
- [4] Postuma RB, Berg D, Stern M, et al. "MDS clinical diagnostic criteria for Parkinson's disease." *Movement Disorders*. 2015;30(12):1591-1601. DOI:10.1002/mds.26424
- [5] Walker FO. "Huntington's disease." *The Lancet*. 2007;369(9557):218-228. DOI:10.1016/S0140-6736(07)60111-1
- [6] Turner MR, Talbot K. "ALS: diagnosis and management in the clinic." *Practical Neurology*. 2013;13(3):145-151. DOI:10.1136/practneurol-2013-000557
- [7] Virtanen, P. et al. (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17, 261-272. <https://doi.org/10.1038/s41592-019-0686-2>
- [8] Pearson, R. K. (2002). Outliers in Process Modeling and Identification. *IEEE Transactions on Control Systems Technology*, 10(1), 55-63.
- [9] M, M. A., & M, M. A. (n.d.). ¿Qué significa realmente el valor de p? https://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1139-76322017000500014
- [10] J. M. Hausdorff, A. Lertratanakul, M. E. Cudkowicz, A. L. Peterson, D. Kaliton y A. L. Goldberger, "Dynamic markers of altered gait rhythm in amyotrophic lateral sclerosis", *J. Appl. Physiol.*, vol. 88, n.º 6, pp. 2045-2053, junio de 2000. <https://doi.org/10.1152/jappl.2000.88.6.2045>
- [11] Lazcano-Ponce, E., Salazar-Martínez, E., & Hernández-Avila, M. (n.d.). *Estudios epidemiológicos de casos y controles. Fundamento teórico, variantes y aplicaciones*. https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0036-36342001000200009
- [12] A. Field, *Discovering Statistics Using IBM SPSS Statistics*, 4th ed. SAGE Publications, 2013.
- [13] D. C. Howell, *Statistical Methods for Psychology*, 8th ed. Cengage Learning, 2013.
- [14] M. S. Pepe, *The Statistical Evaluation of Medical Tests for Classification and Prediction*. Oxford University Press, 2003.
- [15] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Lawrence Erlbaum Associates, 1988.
- [16] A. Field, *Discovering Statistics Using IBM SPSS Statistics*, 4th ed. SAGE Publications, 2013.