

Selected Papers from Impact

Validez de Tres Relojes Inteligentes en la Estimación del Gasto Energético Durante la Marcha y la Carrera al Aire Libre

Validity of Three Smartwatches in Estimating Energy Expenditure During Outdoor Walking and Running

Shenglong Le^{1,2,3,4}, Xiuqiang Wang^{1,3}, Tao Zhang^{1,3}, Si Man Lei^{1,5}, Sulin Cheng^{1,3,4,6}, Wu Yao⁶ y Moritz Schumann⁷

¹Exercise Translational Medicine Center, Shanghai Center for Systems Biomedicine, Shanghai Jiao Tong University, Shanghai, China

²Department of Physical Therapy, Taihe Hospital, Hubei University of Medicine, Shiyan, China

³Key Laboratory of Systems Biomedicine (Ministry of Education), Shanghai Center for Systems Biomedicine, Shanghai Jiao Tong University, Shanghai, China

⁴Faculty of Sport and Health Sciences, University of Jyväskylä, Jyväskylä, Finland

⁵Faculty of Education, University of Macao, Macao, China

⁶Physical Education Department, Shanghai Jiao Tong University, Shanghai, China

⁷Department of Molecular and Cellular Sport Medicine, German Sport University, Cologne, Germany

RESUMEN

Los dispositivos comerciales que se llevan en la muñeca a menudo presentan estimaciones inexactas del gasto de energía (EE), con grandes diferencias entre dispositivos. Nuestro objetivo fue evaluar la validez de Apple Watch Series 6 (AW), Garmin FENIX 6 (GF) y Huawei Watch GT 2e (HW) sobre la estimación del EE durante caminatas y carreras al aire libre. Veinte adultos chinos jóvenes de peso normal usaron simultáneamente tres dispositivos colocados aleatoriamente en ambas muñecas mientras caminaban a 6 km/h y corrían a 10 km/h durante 2 km en una pista de 400 metros. Como criterio se evaluó el EE por calorimetría indirecta (COSMED K5). Para caminar, el EE de AW y GF fue significativamente mayor que el obtenido por el K5 ($p < 0.001$ y 0.002 , respectivamente), pero no para HW ($p = 0.491$). El error porcentual absoluto medio (MAPE) fue del 19.8% para AW, del 32.0% para GF y del 9.9% para HW, respectivamente. Los límites de concordancia (LoA) fueron 44.1, 150.1 y 48.6 kcal para AW, GF y HW, respectivamente. El coeficiente de correlación intraclase (CCI) fue de 0.821, 0.216 y 0.760 para AW, GF y HW, respectivamente. Para la carrera, el EE de AW y de GF fueron significativamente más altos que K5 ($p < 0.001$ y 0.001 , respectivamente), pero no para HW ($p = 0.946$). El MAPE fue de 24.4%, 21.8% y 11.9% para AW, GF y HW, respectivamente. Los LoA fueron 62.8, 89.4 y 65.6 kcal para AW, GF y HW, respectivamente. El ICC fue 0.741, 0.594 y 0.698 para AW, GF y HW, respectivamente. Los resultados indican que los relojes inteligentes probados muestran una validez moderada en las estimaciones del EE para caminar y correr al aire libre.

Palabras Clave: gasto energético, relojes, caminar, correr

ABSTRACT

Commercially wrist-worn devices often present inaccurate estimations of energy expenditure (EE), with large between-device differences. We aimed to assess the validity of the Apple Watch Series 6 (AW), Garmin FENIX 6 (GF) and Huawei Watch GT 2e (HW) in estimating EE during outdoor walking and running. Twenty young normal-weight Chinese adults concurrently wore three index devices randomly positioned at both wrists during walking at 6 km/h and running at 10 km/h for 2 km on a 400- meter track. As a criterion, EE was assessed by indirect calorimetry (COSMED K5). For walking, EE from AW and GF was significantly higher than that obtained by the K5 ($p < 0.001$ and 0.002 , respectively), but not for HW ($p = 0.491$). The mean absolute percentage error (MAPE) was 19.8% for AW, 32.0% for GF, and 9.9% for HW, respectively. The limits of agreement (LoA) were 44.1, 150.1 and 48.6 kcal for AW, GF, and HW respectively. The intraclass correlation coefficient (ICC) was 0.821, 0.216 and 0.760 for AW, GF, and HW, respectively. For running, EE from AW and GF were significantly higher than the K5 ($p < 0.001$ and 0.001 , respectively), but not for HW ($p = 0.946$). The MAPE was 24.4%, 21.8% and 11.9% for AW, GF and HW, respectively. LoA were 62.8, 89.4 and 65.6 kcal for AW, GF and HW, respectively. The ICC was 0.741, 0.594, and 0.698 for AW, GF and HW, respectively. The results indicate that the tested smartwatches show a moderate validity in EE estimations for outdoor walking and running.

Keywords: energy expenditure, watches, walking, running

INTRODUCCIÓN

La tecnología portátil se ha clasificado como la principal tendencia de fitness durante varios años (Thompson, 2021) y en 2021 los envíos mundiales de dispositivos portátiles alcanzaron los 533.6 millones de unidades (Laricchia, 2022). Entre estos, los dispositivos de muñeca son los más comunes y son capaces de monitorear una gran variedad de parámetros vitales, incluidos los pasos dados, la distancia recorrida, la frecuencia cardíaca y el gasto de energía (EE) con una variedad de sensores. Dada su conveniencia, los datos proporcionados por los dispositivos portátiles a menudo se utilizan para monitorear y/o modificar comportamientos de salud tanto para el autocontrol como en entornos de atención médica e investigación (Lyons y cols., 2014). El uso de los dispositivos generalmente se fomenta mediante afirmaciones de marketing establecidas por los fabricantes, mientras que normalmente faltan datos reales sobre la validez de los dispositivos (Evenson y cols., 2015).

La estimación del EE mediante dispositivos portátiles de muñeca parece conveniente, considerando la relación lineal de la frecuencia cardíaca y/o la actividad física con el intercambio gaseoso. En consecuencia, se han realizado numerosos estudios de validación de las estimaciones de EE, pero arrojaron resultados heterogéneos (O'driscoll y cols., 2020; Argent y cols., 2022). En general, parece que los dispositivos de muñeca o de brazo disponibles comercialmente no mostraron una precisión suficiente, mientras que se observaron grandes diferencias entre los dispositivos (O'driscoll y cols., 2020).

Las diferencias en la precisión de los distintos dispositivos pueden depender de varios factores. Lo que es más importante, la precisión de los parámetros de entrada (*input*), como la precisión de la frecuencia cardíaca de los sensores de fotoplethismografía (PPG), afecta las estimaciones del EE (O'driscoll y cols., 2020; Argent y cols., 2022). En este contexto, parece que la precisión de la estimación del EE puede variar con el tipo y la intensidad de las actividades (O'driscoll y cols., 2020). Si bien esto puede estar relacionado con el movimiento que afecta la precisión de la señal PPG, el error observado para diferentes tipos de actividades también puede estar relacionado con los algoritmos que probablemente no toman en cuenta el tipo de actividad física o la postura corporal (Schneller y cols., 2015). Varios estudios encontraron que los ajustes del algoritmo pueden mejorar la validez de la estimación del EE durante el ejercicio (Jakicic y cols., 2004; Van Hoya y cols., 2015). Además, la adición de sensores de frecuencia cardíaca o calor al acelerómetro puede mejorar la precisión de las estimaciones del EE en comparación con la acelerometría sola (O'driscoll y cols., 2020; Kinnunen y cols., 2019). Colectivamente, parece que más avances tecnológicos y algoritmos revisados pueden mejorar la calidad de las estimaciones del EE. Por lo tanto, las validaciones continuas de los dispositivos que se introducen recientemente en el mercado son indispensables.

Para la validación exhaustiva de las estimaciones del EE, se deben considerar una serie de factores. Entre las variables más importantes parece estar el uso previsto del dispositivo, es decir, validar las estimaciones del EE en entornos que representen un uso real por parte de los consumidores (Argent y cols., 2022). Esto se debe a que se pueden usar sensores específicos para actividades específicas. En este sentido, las actividades al aire libre pueden estar respaldadas por la señal GPS, mientras que las actividades en interiores pueden depender únicamente de la acelerometría, lo que probablemente

comprometa la precisión (Charlot y cols., 2014). Por lo tanto, este estudio tiene como objetivo examinar la precisión de las estimaciones del EE de tres nuevas generaciones de relojes inteligentes de los principales fabricantes, a saber, Apple Watch Series 6 (AW), Garmin Fenix 6 (GF) y Huawei GT 2e (HW) durante una caminata y carrera al aire libre en adultos jóvenes chinos.

MATERIALES Y MÉTODOS

Participantes

Veinte sujetos chinos sanos (10 hombres y 10 mujeres) fueron reclutados del campus universitario local (Tabla 1). Sobre la base de un t-test de muestras apareadas, se realizó un análisis *post-hoc* para estimar si el tamaño de la muestra es suficiente con el tamaño del efecto adecuado de tres dispositivos en condiciones de caminar y correr por G*Power versión 3.1.9.6 (Franz Faul, University Kiel, Alemania) utilizando la media y las desviaciones estándar entre el dispositivo de criterio (K5) y el dispositivo de prueba (Passler y cols. 2019). Para AW, GF y HW, el poder estadístico es 1.000, 0.929 y 0.105 para caminar al aire libre, respectivamente, mientras que 0.999, 0.942 y 0.051 para correr al aire libre, respectivamente. Los participantes fueron seleccionados para los criterios de inclusión mediante el cuestionario de estilo de vida y enfermedad y el historial médico, signos y síntomas y factores de riesgo para la estratificación de riesgo del ACSM (Jonas, 2009). Los criterios de inclusión eran ser adultos jóvenes sanos (de 18 a 30 años de edad) con un peso normal (índice de masa corporal de 18.5 a 25 kg/m²) y ejercicio recreativo regular autoinformado (≥ 3 sesiones de ejercicio semanales). Todos los participantes fueron informados sobre los procedimientos del estudio y dieron su consentimiento informado por escrito antes de la prueba. El estudio se realizó de acuerdo con la declaración de Helsinki y fue aprobado por la Junta de Revisión Institucional para la Protección de la Investigación Humana de la Universidad Jiao Tong de Shanghai (número de registro B2020024I). Evaluación de la medida del criterio.

Tabla 1. Características físicas de los participantes.

	Male (n = 10)	Female (n = 10)	All participants (n = 20)
Age (yr)	23.8 ± 1.2	22.1 ± 2.7	23.0 ± 2.2
Height (cm)	178.5 ± 8.3	168.7 ± 8.5	173.6 ± 9.6
Weight (kg)	71.7 ± 10.3	62.6 ± 10.3	67.1 ± 11.1
BMI (kg/m ²)	22.4 ± 2.0	21.8 ± 1.9	22.1 ± 1.9
BF (%)	14.5 ± 4.8	23.8 ± 5.6	19.2 ± 6.9

Data are expressed as mean ± SD. BMI, body mass index; BF, body fat percentage; SD, standard deviation.

Como criterio se utilizó el sistema Cosmed K5 (K5) (Cosmed, Italia). El K5 es un sistema portátil de análisis de gases que utiliza la tecnología *Breath-by-Breath* (respiro a respiro). Esta tecnología permite una determinación precisa del VE, VO₂ y VCO₂ individuales para una amplia gama de tasas metabólicas (Deblois y cols., 2021). La estimación del EE se basa en la relación entre el oxígeno inhalado y el dióxido de carbono exhalado. Se analizan los gases expirados de cada respiración. Antes de cada medición, el dispositivo se calentó durante un mínimo de 15 min y se calibró con gases de calibración de alto grado y una jeringa de calibración de 3 L según las recomendaciones del fabricante. El tamaño de la máscara se ajustó individualmente antes de la prueba y se mantuvo durante toda la medición.

Evaluación del dispositivo de referencia

Este estudio examinó la validez del Apple Watch Series 6 (AW), Garmin Fenix 6 (GF) y Huawei GT 2e (HW), todos los cuales fueron las generaciones más nuevas presentadas en enero de 2021. Cada reloj inteligente tiene la capacidad para caminar y correr al aire libre y registra la distancia, la velocidad y la frecuencia cardíaca en tiempo real. Todos los dispositivos usan fotoplethismografía para estimar la frecuencia cardíaca desde la muñeca y usan el GPS para estimar la distancia y la velocidad al caminar y correr al aire libre. El EE era estimada en tiempo real y era visualizada.

Todos los relojes utilizados en el estudio se compraron comercialmente. Se colocaron tres relojes, dos en una muñeca y uno en la otra muñeca para diferentes participantes de acuerdo con una lista de asignación aleatoria previa para mantener contrabalanceda la cantidad de dispositivos en una muñeca. La posición del reloj se seleccionó con el mayor cuidado posible de acuerdo con las instrucciones del fabricante.

Protocolo

Los datos para este estudio se recopilaron durante una visita. Se instruyó a los sujetos para que no consumieran alimentos, café, té u otros estimulantes, por ej., bebidas energéticas o refrescos excepto agua al menos 6 hs antes de las mediciones y no realizar ninguna actividad física vigorosa ni consumir alcohol durante las 24 hs previas a las mediciones. La altura de los sujetos se midió con una precisión de 0.1 cm utilizando una escala de altura montada en la pared.

La masa corporal y la composición corporal se midieron después de vaciar la vejiga y en ropa interior ligera utilizando un dispositivo de bioimpedancia calibrado InBody 720 (Biospace, Co, Ltd., Seúl, Corea). Se utilizaron la altura, el peso, el sexo y la fecha de nacimiento para inicializar los relojes inteligentes de cada sujeto.

Después de la medición antropométrica, el protocolo incluyó dos sesiones de ejercicio realizadas en la pista de atletismo al aire libre, como se muestra en la Figura 1. Dos sesiones fueron separadas por un período de recuperación, que fue de 10 minutos de descanso para sentarse. En la primera sesión, se aconsejó a los sujetos que caminaran una distancia de 2 km a aproximadamente 6 km/h. En la segunda sesión, se pidió a los sujetos que corrieran una distancia de 2 km a aproximadamente 10 km/h. Se instruyó a los sujetos para que mantuvieran un ritmo constante durante la sesión. La velocidad de marcha fue de 6.15 ± 0.29 km/h y la velocidad de carrera fue de 10.94 ± 0.99 km/h. El examinador confirmaba que la frecuencia cardíaca había vuelto a los niveles de reposo antes de la segunda sesión. El modo de caminar o correr al aire libre correspondiente se seleccionó e inició y finalizó al mismo tiempo para todos los relojes inteligentes. Los investigadores registraron la hora de inicio y finalización de cada sesión en una hoja de datos. La temperatura y humedad relativa fue de $16.5 \pm 2.6^\circ\text{C}$ y $56.7 \pm 14.3\%$ para caminar y $13.9 \pm 2.9^\circ\text{C}$ y $62.6 \pm 21.8\%$ para correr.

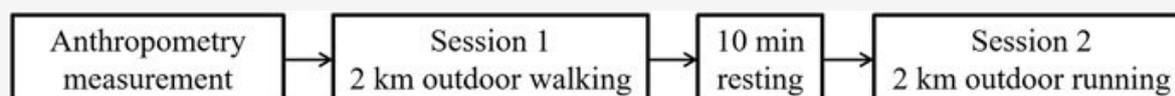


Figura 1. Protocolo experimental. En la primera sesión, se aconsejó a los sujetos que caminaran una distancia de 2 km a aproximadamente 6 km/h. En la segunda sesión, se pidió a los sujetos que corrieran una distancia de 2 km a aproximadamente 10 km/h. Dos sesiones estuvieron separadas por 10 minutos de descanso para sentarse.

Adquisición y procesamiento de datos

Los datos de criterio del K5 se descargaban 'respiro a respiro' e incluyeron medidas de EE. Los valores de EE del K5 se sumaron individualmente para cada sesión de ejercicio. Las estimaciones del EE de los relojes se obtuvieron directamente de capturas de pantalla de las respectivas aplicaciones, porque los relojes informaban principalmente totales acumulativos. Debido a errores técnicos, sólo obtuvimos 19 datos de EE del HW para caminar y correr, respectivamente.

Análisis de datos

Los análisis estadísticos se realizaron utilizando el software IBM *SPSS Statistics* versión 24 (IBM, Armonk, Nueva York). Se utilizaron estadísticas descriptivas para caracterizar la población de la muestra. La validez de los relojes fue determinada por varias pruebas estadísticas. Los datos de los relojes inteligentes se compararon con el criterio mediante t-tests de muestras apareadas. Los errores porcentuales absolutos medios (MAPE) se calcularon como un indicador de error de medición. Los errores porcentuales absolutos (%) se calcularon de la siguiente manera: $\{EE \text{ de relojes inteligentes} - EE \text{ de K5}\} / (EE \text{ de K5}) \times 100$ para cada sujeto. Se utilizó un MAPE de $\leq 10\%$ como valor criterio de validez (Nelson y cols., 2016). Como método comúnmente utilizado para validar dispositivos portátiles, el Coeficiente de correlación intraclase (ICC) definió el acuerdo entre el estándar de oro y los dispositivos evaluados, proporcionando una estimación de la concordancia general entre dos métodos (Fokkema y cols., 2017; Wahl y cols., 2017; Boudreaux y cols., 2018). Los umbrales de acuerdo excelente, bueno, moderado y bajo se definieron como valores de ICC de *ont*, *Good*, *Excelent*, *gooices*, *providin* (Fokkema y cols., 2017). Para investigar el nivel de acuerdo, se prepararon gráficos de Bland-Altman según Bland y Altman (Bland y Altman, 1986). Para ello, se establecieron límites de concordancia al 95%.

RESULTADOS

EE durante la caminata al aire libre

El EE durante la caminata al aire libre se muestra en la Tabla 2. En promedio, los participantes lograron un EE de 108.7 ± 17.4 kcal en el K5 durante la caminata al aire libre. El EE promedio estimado fue de 129.1 ± 20.1 kcal, 139.6 ± 39.6 kcal y 111.2 ± 18.1 kcal de AW, GF y HW, respectivamente. El análisis del t-test apareado mostró que el EE de AW y GF fueron significativamente más altos que los obtenidos por el K5 ($p < 0.001$ y $p < 0.002$, respectivamente), pero no significativamente diferentes para HW ($p = 0.491$). En comparación con el K5, el MAPE de EE fue de 19.8%, 32.0% y 9.9% en AW, GF y HW, respectivamente (Cuadro 2). Los valores individuales del EE de AW (ICC = 0.821) y de HW (ICC = 0.760) también mostraron una buena concordancia con el K5. Sin embargo, la concordancia para GF (ICC = 0.216) fue baja (Fokkema y cols., 2017).

Tabla 2. Examen descriptivo de las diferencias entre el EE estimado (relojes inteligentes) y el EE medido (K5) durante la caminata y carrera al aire libre.

Activity	Device	N	EE (kcal)	Diff (kcal)	MAPE (%)	ICC	t	p	p
Walking	AW	20	129.1 ± 20.1	20.5 ± 11.3	19.8 ± 12.4	0.821	-8.129	0.000	1.000
	GF	20	139.6 ± 39.6	31.0 ± 38.3	32.0 ± 34.1	0.216	-3.615	0.002	0.929
	HW	19	111.2 ± 18.1	2.0 ± 12.4	9.9 ± 8.2	0.760	-0.703	0.491	0.105
	K5	20	108.7 ± 17.4	NA	NA	NA	NA	NA	NA
Running	AW	20	137.8 ± 23.1	25.7 ± 16.0	24.4 ± 16.1	0.741	-7.162	0.000	0.999
	GF	20	137.8 ± 23.1	19 ± 22.8	21.8 ± 17.3	0.594	-3.715	0.001	0.942
	HW	19	111.8 ± 21.1	-0.3 ± 16.7	11.9 ± 9.9	0.698	0.069	0.946	0.051
	K5	20	112.2 ± 21.3	NA	NA	NA	NA	NA	NA

Data are expressed as mean ± SD, for EE, Diff and MAPE., Data from the smartwatches were compared with the criterion using paired sample t-tests. EE, energy expenditure; Diff: the difference of the estimated EE (smartwatches) with the measured EE (K5); K5, the golden standard of EE; assessment; AW, Apple Watch Series 6; GF, Garmin FENIX, 6; HW, Huawei Watch GT, 2e; MAPE, mean absolute percentage error; ICC, intraclass correlation coefficient; SD, standard deviation; p, statistical power.

En los gráficos de Bland-Altman, el porcentaje de los valores dentro de los límites de concordancia fue del 100%, 90% y 95% para AW (Figura 2A), GF (Figura 2C) y HW (Figura 2E), respectivamente. El AW presentó los límites de concordancia más estrechos (44.1 kcal), seguido del HW con 48.6 kcal, mientras que el GF exhibió los límites de concordancia más amplios (150.1 kcal).

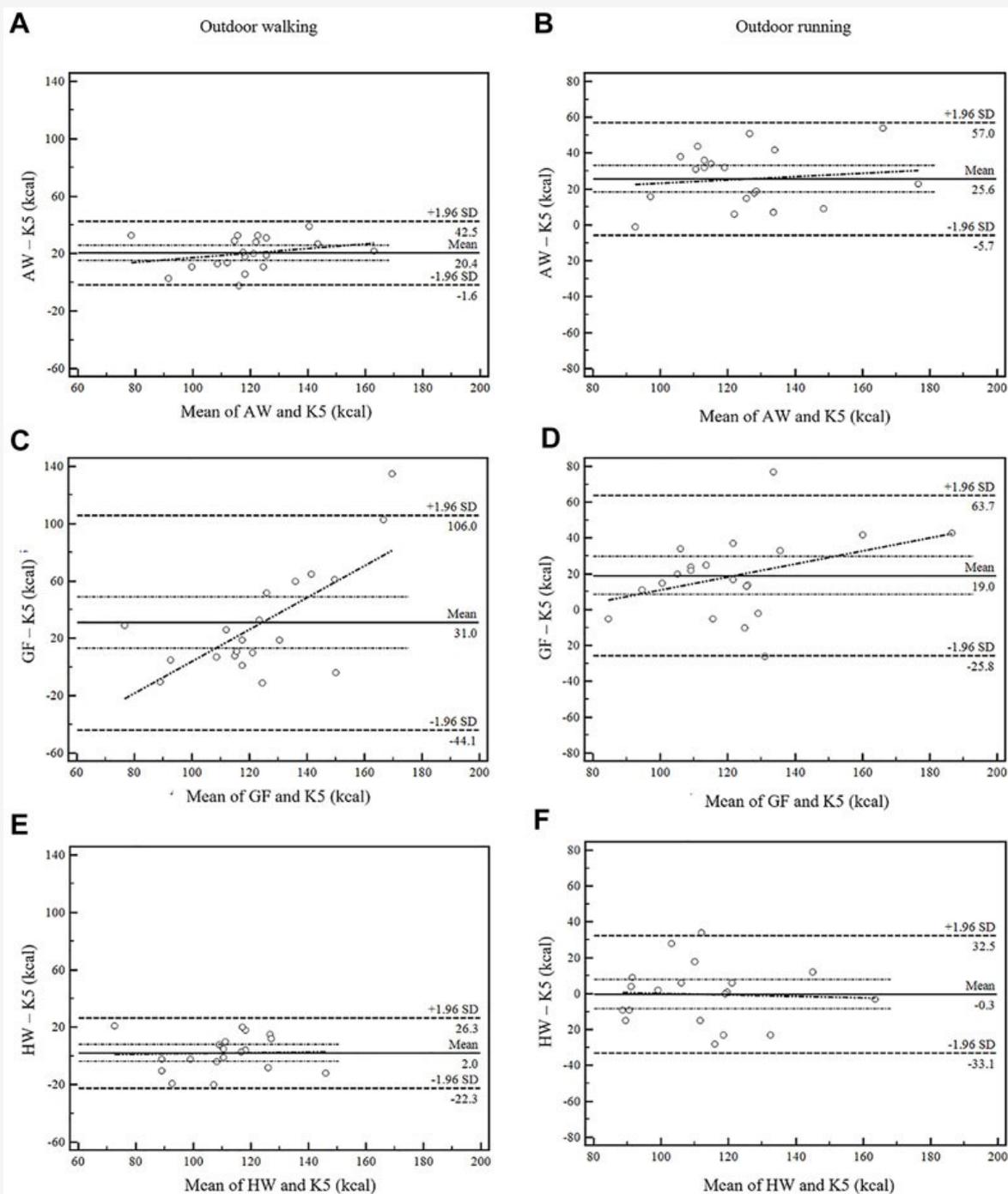


Figura 2. Gráficas de Bland-Altman que comparan las estimaciones del EE por los relojes inteligentes (AW (A,B), GF (C,D), HW (E,F)) y K5. Las diferencias de los valores del EE en el eje y en relación con la media de los dos métodos (reloj inteligente y K5) en el eje x. Las diferencias medias (sesgo) entre el EE estimado y el EE del criterio, los límites superior e inferior de concordancia (ULOa, LLoa) están etiquetados en los gráficos. Los límites de concordancia (LoA) se calcularon como medias ± 1.96 ; K5, el criterio de evaluación de EE; AW, Apple Watch Serie 6; GF, Garmin FENIX 6; HW, reloj Huawei GT 2e; SD, desviación estándar.

EE durante la carrera al aire libre

El EE durante la carrera al aire libre se muestra en la Tabla 2. En promedio, los participantes lograron un EE de $112,2 \pm 21,3$ kcal en K5 durante la carrera al aire libre. El EE promedio estimado fue de $137,8 \pm 23,1$ kcal, $131,1 \pm 28,8$ kcal y $111,8 \pm 21,1$ kcal de AW, GF y HW, respectivamente. El análisis de la prueba t pareada mostró que EE de AW y GF fueron significativamente más altos que K5 ($p < 0.001$ y $p < 0.001$, respectivamente), pero no significativamente diferentes para HW ($p = 0.946$). En comparación con el K5, el MAPE de EE fue de 24.4%, 21.8% y 11.9% en AW, GF y HW, respectivamente. Los valores individuales del EE de AW (ICC = 0.741) y de HW (ICC = 0.698) también mostraron una

concordancia moderada con el K5. Sin embargo, la concordancia para GF (ICC = 0.594) fue baja (Fokkema y cols., 2017). En los gráficos de Bland-Altman, el porcentaje de los valores dentro de los límites de concordancia fue del 95% para AW (Figura 2B) y para GF (Figura 2D) y del 100% para HW (Figura 2F). El AW mostró los límites de concordancia más estrechos (62.8 kcal), seguido por HW de 65.6 kcal, mientras que el GF exhibió los límites de concordancia más amplios (89.4 kcal).

Discusión

Este estudio investigó la validez de tres relojes inteligentes populares (AW, GF, HW) para estimar el EE durante caminatas y carreras al aire libre mediante un sistema portátil de análisis de gases (Cosmed K5) como criterio. Los resultados revelaron que algunos, pero no todos, los relojes inteligentes proporcionan una precisión razonable para la estimación del EE. Las estimaciones del EE proporcionadas por AW y HW demostraron un acuerdo de criterio bueno/moderado, mientras que el EE de GF mostró un acuerdo deficiente.

Aunque el HW arrojó los mejores resultados generales, el AW también se desempeñó bien en este estudio. El HW mostró el MAPE más bajo (9.9% para caminar y 11.9% para correr), mientras que el AW tuvo los límites de concordancia más estrechos (44.1 kcal para caminar y 62.8 kcal para correr). De hecho, un MAPE de <10% puede considerarse una precisión alta (Nelson y cols., 2016). Dado que el fabricante de HW tiene su sede en China y se incluyó una población china en este estudio, es probable que esto explique la alta precisión. En un estudio anterior, se demostró que la precisión de las estimaciones del EE por parte de los monitores de actividad difería entre diferentes etnias (Brazeau y cols., 2014), pero en ese estudio sólo se usaron los datos del acelerómetro para las estimaciones del EE. Desafortunadamente, la mayoría de los fabricantes no brindan detalles sobre los algoritmos que se utilizan para las estimaciones del EE y, como tal, aún se desconoce qué variables se usaron en los dispositivos evaluados en el presente estudio. Sin embargo, es probable que se haya utilizado una combinación de datos derivados de las acelerometría, PPG y GPS. Como tal, sólo podemos evaluar la validez ecológica que es el resultado tanto de la calidad del sensor como del algoritmo, pero se desconoce la fuente real de imprecisión especialmente observada en GF.

Existe una variabilidad considerable en la precisión del dispositivo para predecir el EE de diferentes tipos de actividades (O'driscoll y cols., 2020; Lee y cols., 2014). El presente estudio demostró que las estimaciones generales del error fueron similares entre caminar y correr para AW y HW, pero diferentes para GF. El reloj GF probablemente podría proporcionar mejores estimaciones del EE para caminar al aire libre (MAPE = 32.0%) que para correr (MAPE = 21.8%). Wahl y cols. (Wahl y cols., 2017) examinaron la influencia del ritmo de carrera en el EE. El estudio muestra una influencia significativa del ritmo de carrera en el EE estimado. El gasto de energía tiende a sobreestimarse a un ritmo más bajo y a subestimarse a un ritmo más alto. Roos y cols. (Roos y cols., 2017) encontraron que los relojes deportivos subestimaron significativamente el EE durante la carrera de alta intensidad con un error proporcional que aumentaba a medida que aumentaba la velocidad de carrera. Ambos estudios sugirieron mejorar el algoritmo de estimación del EE en todos los rangos de velocidad (Roos y cols., 2017; Wahl y cols., 2017). Sin embargo, esta diversidad también puede estar relacionada con imprecisiones del sensor (por ejemplo, artefactos de movimiento) a diferentes velocidades. Estas observaciones indicaron que los consumidores/profesionales de la salud deberían prestar más atención para seleccionar el dispositivo más preciso para los distintos tipos de actividad.

El EE del reloj inteligente sobrestimó el EE en comparación con el K5 durante la caminata y la carrera al aire libre. Estas observaciones son congruentes con estudios previos (Bai y cols., 2016; Pope y cols., 2019). El EE durante la marcha al aire libre fue sobrestimado en un 19.8% y un 32.0% por AW y GF, respectivamente, mientras que el coste energético de la carrera fue sobrestimado en un 24.4% y un 21.8%. Teniendo en cuenta que los datos del EE del reloj inteligente pueden usarse con el fin de mantener o incluso perder peso corporal, estas observaciones son notables. Por ejemplo, el individuo puede establecer la meta de aumentar el EE en 250 kcal por día caminando o corriendo para perder 1 kg de peso corporal por mes. Sin embargo, debido a la inexactitud de los datos del EE del reloj inteligente, el EE diario en aumento real puede ser inferior a 200 kcal, lo que resulta en la ineficacia del programa de pérdida de peso. Nuestros hallazgos, por lo tanto, sugieren un enfoque cauteloso al interpretar los datos del EE del reloj inteligente dada la sobreestimación observada del EE.

Al interpretar nuestros hallazgos, es necesario tener en cuenta algunas limitaciones. En particular, nuestro estudio fue diseñado para incluir sólo caminar y correr al aire libre con una velocidad constante autoseleccionada recomendada. Además, es posible que nuestros hallazgos no se traduzcan directamente en otras actividades debido a, por ejemplo, diferencias en los artefactos de movimiento. Además, nuestra muestra se restringió a chinos jóvenes sanos dentro del rango normal de IMC. Por lo tanto, no podemos generalizar estos hallazgos a otros grupos de población diversos, como ancianos u obesos. Además, dado que el lado del cuerpo se asignó al azar, consideramos efectos específicos de la precisión. Los estudios futuros deberían considerar potenciar el estudio para comparar los efectos de diferentes lugares de uso en la precisión de los dispositivos. Finalmente, este estudio no evaluó la confiabilidad de la estimación del EE de los relojes inteligentes. A pesar de los problemas de precisión, la confiabilidad es muy importante para que las personas y los

profesionales de la salud controlen los cambios en el EE a lo largo del tiempo. Los estudios futuros también deberían incorporar pruebas de confiabilidad.

Conclusión

Los hallazgos de este estudio indican que los relojes inteligentes de AW, GF y HW generalmente pueden tener una validez moderada en las estimaciones del EE para caminar y correr al aire libre. El tamaño de muestra pequeño con grandes variables incluso tuvo suficiente poder siendo un factor limitante para generalizar los resultados para diferentes poblaciones. Los consumidores también deben tener cuidado al usar los relojes inteligentes probados para predecir su gasto de energía.

Declaración de Ética

Los estudios con participantes humanos fueron revisados y aprobados por la Junta de Revisión Institucional para la Protección de la Investigación Humana de la Universidad Jiao Tong de Shanghai (número de registro B2020024I, 03.09.2020). Los pacientes/participantes dieron su consentimiento informado por escrito para participar en este estudio.

Financiamiento

Esta investigación fue apoyada por Huawei Terminal Co. LTD. El financiador no participó en el diseño del estudio, la recopilación, el análisis, la interpretación de los datos, la redacción de este artículo o la decisión de enviarlo para su publicación. Todos los autores declaran que no hay otros intereses en competencia.

Conflicto de intereses

Los autores declaran que la investigación se realizó en ausencia de cualquier relación comercial o financiera que pudiera interpretarse como un potencial conflicto de interés.

REFERENCIAS

1. Argent R., Hetherington-Rauth M., Stang J., Tarp J., Ortega F. B., Molina-Garcia P., et al. (2022). Recommendations for determining the validity of consumer wearables and smartphones for the estimation of energy expenditure: Expert statement and checklist of the INTERLIVE network. *Sports Med.* 52, 1817-1832.
2. Bai Y., Welk G. J., Nam Y. H., Lee J. A., Lee J. M., Kim Y., et al. (2016). Comparison of consumer and research monitors under semistructured settings. *Med. Sci. Sports Exerc.* 48, 151-158.
3. Bland J. M., Altman D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1, 307-310.
4. Boudreaux B. D., Hebert E. P., Hollander D. B., Williams B. M., Cormier C. L., Naquin M. R., et al. (2018). Validity of wearable activity monitors during cycling and resistance exercise. *Med. Sci. Sports Exerc.* 50, 624-633.
5. Brazeau A. S., Suppere C., Strychar I., Belisle V., Demers S. P., Rabasa-Lhoret R. (2014). Accuracy of energy expenditure estimation by activity monitors differs with ethnicity. *Int. J. Sports Med.* 35, 847-850.
6. Charlot K., Cornolo J., Borne R., BrugniauxandPichon J. V. A., Richalet J. P., Chapelot D., et al. (2014). Improvement of energy expenditure prediction from heart rate during running. *Physiol. Meas.* 35, 253-266.
7. Deblois J. P., White L. E., Barreira T. V. (2021). Reliability and validity of the COSMED K5 portable metabolic system during walking. *Eur. J. Appl. Physiol.* 121, 209-217.
8. Evenson K. R., Goto M. M., Furberg R. D. (2015). Systematic review of the validity and reliability of consumer-wearable activity trackers. *Int. J. Behav. Nutr. Phys. Act.* 12, 159.
9. Fokkema T., Kooiman T. J., Krijnen W. P., Cees V. D. S., Martijn D. E. G., De Groot M. (2017). Reliability and validity of ten consumer activity trackers depend on walking speed. *Med. Sci. Sports Exerc.* 49, 793-800.
10. Jakicic J. M., Marcus M., Gallagher K. I., Randall C., Thomas E., Goss F. L., et al. (2004). Evaluation of the SenseWear Pro Armband to assess energy expenditure during exercise. *Med. Sci. Sports Exerc.* 36, 897-904.
11. Jonas S. (2009). ACSM's exercise is medicine : a clinician's guide to exercise prescription. Philadelphia: Lippincott, Williams and Wilkins.
12. Kinnunen H., Hakkinen K., Schumann M., Karavirta L., Westerterp K. R., Kyrolainen H. (2019). Training-induced changes in daily energy expenditure: Methodological evaluation using wrist-worn accelerometer, heart rate monitor, and doubly labeled water technique. *PLoS ONE* 14, e0219563.
13. Laricchia F. (2022). Total wearable device unit shipments worldwide 2014-2021 Available at: <https://www.statista.com/statistics/437871/wearables-worldwide-shipments/>[Accessed March 29, 2022].
14. Lee J. M., Kim Y., Welk G. J. (2014). Validity of consumer-based physical activity monitors. *Med. Sci. Sports Exerc.* 46, 1840-1848.
15. Lyons E. J., Lewis Z. H., Mayrsohn B. G., Rowland J. L. (2014). Behavior change techniques implemented in electronic lifestyle

- activity monitors: a systematic content analysis. *J. Med. Internet Res.* 16, e192.
16. Nelson M. B., Kaminsky L. A., Dickin D. C., Montoye A. H. (2016). Validity of consumer-based physical activity monitors for specific activity types. *Med. Sci. Sports Exerc.* 48, 1619-1628.
 17. O'driscoll R., Turicchi J., Beaulieu K., Scott S., Matu J., Deighton K., et al. (2020). How well do activity monitors estimate energy expenditure? A systematic review and meta-analysis of the validity of current technologies. *Br. J. Sports Med.* 54, 332-340.
 18. Passler S., Bohrer J., Blochinger L., Senner V. (2019). Validity of wrist-worn activity trackers for estimating VO2max and energy expenditure. *Int. J. Environ. Res. Public Health* 16.E3037.
 19. Pope Z. C., Lee J. E., Zeng N., Gao Z. (2019). Validation of four smartwatches in energy expenditure and heart rate assessment during exergaming. *Games Health J.* 8, 205-212.
 20. Roos L., Taube W., Beeler N., Wyss T. (2017). Validity of sports watches when estimating energy expenditure during running. *BMC Sports Sci. Med. Rehabil.* 9, 22.
 21. Schneller M. B., Pedersen M. T., Gupta N., Aadahl M., Holtermann A. (2015). Validation of five minimally obstructive methods to estimate physical activity energy expenditure in young adults in semi-standardized settings. *Sensors (Basel)* 15, 6133-6151.
 22. Thompson W. R. (2021). Worldwide survey of fitness trends for 2021. *ACSM's Health Fit. J.* 25, 10-19.
 23. Van Hoya K., Boen F., Lefevre J. (2015). Validation of the SenseWear Armband in different ambient temperatures. *J. Sports Sci.* 33, 1007-1018.
 24. Wahl Y., Duking P., Droszez A., Wahl P., Mester J. (2017). Criterion-validity of commercially available physical activity tracker to estimate step count, covered distance and energy expenditure during sports conditions. *Front. Physiol.* 8, 725.

Cita Original

Le S, Wang X, Zhang T, Lei SM, Cheng S, Yao W and Schumann M (2022) Validity of three smartwatches in estimating energy expenditure during outdoor walking and running. *Front. Physiol.* 13:995575. doi: 10.3389/fphys.2022.995575