

Accuracy of wrist-worn activity trackers for measuring steps in patients after major abdominal surgery: A validation study

Zhi Li¹, Weiyan Feng², Lili Zhou² and Shu Gong³ 

Abstract

Background: Wearable activity trackers provide a simple and objective measurement of postoperative mobilization. However, few have validated the accuracy of trackers in patients after major abdominal surgery.

Objective: To examine the accuracy of wrist-worn activity trackers to measure steps of patients in early mobilization after major abdominal surgery, and to explore the influence of clinical variables and gait parameters on the accuracy of trackers.

Methods: Forty-five patients after major abdominal surgery were recruited to participate in modified six-minute walk tests wearing three trackers simultaneously, the Fitbit Inspire HR, Xiaomi MI 4, and HONOR 5. The differences in displayed steps before and after the walking test were considered as the step counts measured by the trackers; the actual steps taken were determined as the average of the values manually counted by two researchers. The intraclass correlation coefficient, Bland–Altman method, mean percentage error, and mean absolute percentage error were used to assess the accuracy of trackers with reference to manual step counts.

Results: The three trackers undercounted postoperative steps by –65.5% to –23.5%. Analysis showed low-to-good agreement between step counts recorded by trackers and actual steps (ICC = 0.35–0.75); the mean absolute percentage errors ranged from 24.5% to 65.7%. For all trackers, mean absolute percentage errors correlated negatively with postoperative days ($r = -0.626$ to -0.744), walking speed ($r = -0.714$ to -0.854), step length ($r = -0.466$ to -0.615), and cadence ($r = -0.681$ to -0.790), while there were positive correlations between mean absolute percentage errors and the number of abdominal drains ($r = 0.450$ – 0.514).

Conclusions: The specific activity trackers used in this study might not be reliable tools for measuring steps counts during the walking test in the early postoperative period for patients undergoing major abdominal surgery.

Keywords

Wearable technology, activity tracker, early mobilization, accuracy, major abdominal surgery

Submission date: 6 June 2024; Acceptance date: 17 October 2024

¹Department of Cardiovascular Surgery, West China Hospital, Sichuan University/West China School of Nursing, Sichuan University, Chengdu, Sichuan Province, China

²Division of Pancreatic Surgery, Department of General Surgery, West China Hospital, Sichuan University/West China School of Nursing, Sichuan University, Chengdu, Sichuan Province, China

³Department of Nursing, West China Hospital, Sichuan University/West China School of Nursing, Sichuan University, Chengdu, Sichuan Province, China

Corresponding author:

Shu Gong, Department of Nursing, West China Hospital, Sichuan University/West China School of Nursing, Sichuan University, 37, Guoxue Lane, Wuhou District, Chengdu 610041, Sichuan Province, China.
Email: gdy-29@163.com



Introduction

Despite a significant decrease in major abdominal surgery-related mortality over the past few decades, the morbidity continues to remain high. Reports indicate that complications occur in up to 40%–60% of patients following major abdominal surgery;^{1–3} notably, the morbidities mostly involve nonsurgical complications, including those of pulmonary and cardiac origin.⁴ Early mobilization, which is one of the important components of enhanced recovery after surgery (ERAS), may minimize these complications and is perceived to be the key for successful enhanced recovery.^{5,6}

Although the central role of early postoperative mobilization is widely known, the quality of measurement and monitoring is poor, making it one of the most difficult ERAS interventions to enforce.^{7,8} Postoperative mobilization assessment methods can be classified into three categories. The first involves scales focusing on mobilization levels, such as the Brown rating of mobility,⁹ Johns Hopkins Highest Level of Mobility,¹⁰ and cumulative ambulation score,¹¹ among others; however, information gained by these scales is limited and lacks adequate specificity for postoperative patients. The second involves measurement of the ability to perform the activity; this includes the generally known two-minute test or six-minute walk test (6MWT),^{12,13} which place more emphasis on functional status, require trained professionals, are relatively more time-consuming compared to other mobilization assessment methods mentioned above, and can only be performed intermittently. The third involves quantification of daily activity, including subjective and objective methods. Activity diaries that rely on patient reports are one of the common subjective activity measurement tools¹⁴; they are easy to use, but may be inaccurate due to recall bias. Objective methods usually use motion sensors, including pedometers and accelerometers, to monitor activity.^{14,15} Among these tools, wearable activity trackers based on accelerometers enable the direct, consecutive, and accurate real-time monitoring of activity¹⁶; they therefore potentially provide an opportunity to compensate for deficiencies in postoperative mobilization monitoring methods mentioned above. Consumer-based trackers are preferred over research-based activity trackers, because they are considerably cheaper and easier to use on a daily basis.¹⁷ In particular, wrist-worn activity trackers have increasingly gained popularity and are associated with higher compliance due to their user-friendliness and the greater comfort offered.^{18,19}

The use of commercially available wrist-worn activity trackers in the postoperative period is therefore increasing. Various consumer activity trackers, such as Fitbit,^{20–22} Vivofit,^{23,24} Polar Loop,²⁵ Jawbone,²⁶ and MI,²⁷ have been reported to monitor postoperative step counts. Studies indicate that trackers providing immediate step

count feedback may increase postoperative mobilization;²⁵ they also show that higher postoperative step counts, as measured via trackers, reduce the probability of prolonged length of stay and readmission after major surgery.^{21,22} This makes it possible to evaluate postoperative recovery trajectories and identify patients at risk of poor clinical outcomes.^{23,24} However, there are few validation studies in this population. Compared to the general population (i.e. healthy adults, children, older adults, and patients with chronic diseases), patients who undergo major abdominal surgery (such as hepatopancreatobiliary surgery) often have altered gait. For these population, factors including suboptimal pain management, the presence of multiple tubes and lines, fatigue caused by prolonged operative duration and high intraoperative blood loss, and large open wounds in the abdomen (especially in the early postoperative period)²⁸ can lead to a slower walking speed, shorter stride length, or less arm movement; these factors are likely to affect the step count accuracy of trackers.²⁹

This study therefore aimed to assess the accuracy of three wrist-worn activity trackers, namely, the Fitbit Inspire HR, Xiaomi MI 4, and HONOR 5, in measuring steps of patients during early mobilization after major abdominal surgery. We hypothesized that some clinical variables and gait parameters may affect the step count accuracy of trackers in this population.

Methods

Design and participants

This cross-sectional study is part of a large research program aiming to promote early postoperative mobilization in pancreatic surgery patients. We recruited patients who underwent major abdominal surgery at a tertiary teaching hospital in western China, between March 8 and May 28 in 2021. Patients were included if they underwent elective liver, biliary tract, or pancreatic surgery lasting for more than 2 hours; were aged 18 or older; having no neurological or musculoskeletal diseases that could preclude normal walking; were able to follow commands necessary for participating in the study; and provided informed consent.

Instruments

Three activity trackers were used: the Fitbit Inspire HR (Fitbit Inc., USA), Xiaomi MI Band 4 (Beijing Xiaomi Technology Co., Ltd, China), and HONOR Band 5 (Honor Device Co., Ltd, China). It should be noted that the selection of those three trackers was based on the widespread use of Fitbit in the literature, while MI and HONOR were chosen due to their common presence in the Chinese market. All three trackers need to be worn on the wrist and have been reported (by manufacturers) to sense motion in three dimensions and convert these readings

into activity information. The trackers measure the steps taken, distance walked, calories burned, and the heart rate; the data can be shown on the touch displays and can be transferred wirelessly by connecting to the corresponding app via Bluetooth.

Procedures

To ensure that the three trackers used in this study had no inherent errors, such as manufacturing defects, we recruited a group of healthy individuals for comparison, though not as a standard control group. Prior to the main experiment, a preliminary test was conducted on 10 staff members. They were asked to walk 120 m while wearing three trackers and were accompanied by two researchers who separately counted the steps silently with hand-held counters.

In the main experiment, all patients underwent a modified 6MWT while simultaneously wearing the trackers. Each tracker was programmed with the patient's age, gender, height, and weight and was then worn randomly on the nondominant wrist (or on the other wrist if an intravenous catheter was present). "Randomly" refers to the random order in which the three trackers were worn on the wrist. The modified 6MWT was conducted along an unobstructed 30-m long predetermined flat route in the corridor, with visible landmarks placed every 5 m. Unlike the standard 6MWT, which aims to assess submaximal exercise capacity by having patients walk as fast and as far as possible, the modified 6MWT was primarily intended to simulate postoperative walking. Patients were encouraged to walk back and forth within this distance at their normal walking pace; they were also instructed to stand still for 10 seconds prior to starting and upon completion of the tests. The number of steps displayed on the trackers was recorded during this time and the pain was assessed. The distance walked was also measured using floor markings. A researcher walked at half a meter behind the patient during the entire course of the walking test for supervision and safety. Patients could stop at any time if they felt any discomfort including pain, fatigue, or unbearable dyspnea; the reason for termination and the actual walking time and distance covered was recorded. Two other researchers, blinded to tracker step counts, separately and simultaneously counted the steps with hand-held counters. To prevent any potential cross-infection between patients, after each patient completed the modified 6MWT, the trackers were wiped and disinfected according to our institutional infection control guidelines, using 500mg/L chlorine-containing disinfectant for surface disinfection.

Measurements

Accuracy was defined as the degree to which the tracker-recorded step counts could accurately reflect the actual steps taken. The average of the two values manually

counted by each researcher using a counter was considered as the gold standard measure of actual steps taken. The number of steps recorded by the trackers was calculated as the difference between the number of steps displayed on the tracker at the beginning and end of each walking test.

The following demographic and clinical data were collected: age, sex, height, weight, type of surgery, surgical approach, postoperative diagnosis, postoperative day (POD), the number of drains placed in the abdomen, first time for postoperative walking, pain scores before and after the modified 6MWT, completion of the modified 6MWT, and total walking time and distance. In this context, the types of drains used included abdominal drainage tubes, subcutaneous drainage tubes, biliary T-tubes, external pancreatic duct stents, and percutaneous transhepatic cholangial drainage tubes. Pain was assessed immediately prior to starting and after completion of the modified 6MWT; it was measured using the numeric rating scale, with scores of 0–10, where 0 denoted "no pain" and 10 denoted "the worst pain imaginable." Several gait parameters were also calculated, including walking speed (distance/time), step length (distance/steps), and cadence (steps/time).

Data analysis

Data were analyzed using SPSS v.24.0 (IBM Inc., Armonk, NY, USA). Accuracy was evaluated from two aspects: agreement and error analysis. In particular, a two-way random absolute agreement intraclass correlation coefficient (ICC) with 95% confidence intervals was used to examine the relative agreement between the tracker-recorded steps and actual steps taken. An ICC value of > 0.90, 0.75–0.90, 0.60–0.75, and < 0.60 was interpreted as excellent, good, moderate, and low, respectively.³⁰ The Bland–Altman method, which plotted the mean differences (bias) and 95% limits of agreement between step counts from different trackers and manual counts, was used to examine the absolute agreement between tracker-identified steps and actual steps taken. Broader ranges between the lower and upper limits reflected greater differences between the two methods, and therefore, lower agreement. Error analysis involved the mean percentage error (MPE) and mean absolute percentage error (MAPE). The MPE between steps obtained via each tracker and the actual steps taken was calculated as follows: $(\text{[tracker step count} - \text{manual step count]}/\text{manual step count}) \times 100$; this measure indicated the direction of error in measurements obtained by each tracker. Positive and negative values represented overestimation and underestimation of tracker-measured steps, respectively, relative to the actual steps taken. To assess the magnitude of error, MAPE was used to represent the specific values for accuracy and calculated using the formula: $(\text{abs. [tracker step count} - \text{manual step count]}/$

manual step count) $\times 100$. A MAPE of below 5% was considered excellent, while that above 10% was considered poor.³¹

The Mann-Whitney *U*, Kruskal-Wallis *H*, and Spearman's rank correlation tests were used to assess the influence of demographic, clinical, and gait parameters on step count accuracy. For factors which were considered statistically significant ($P \leq .05$), further visual assessment was performed by graphing the MAPEs of each tracker against the potential variables. To examine our hypothesis that patients who underwent major abdominal surgery may have altered gait in the early postoperative period, we also performed a simple analysis by graphing PODs against the other potential variables. This analysis was conducted exclusively on the patient group.

Results

We invited 52 patients to participate in this study, seven of them had refused, and a total of 45 patients were included. Most patients underwent open pancreas and liver surgery (80%) and participated in the walking test on median POD 4 (Table 1). Overall, eight (17.8%) patients did not complete the modified 6MWT, mainly because of pain ($n=5$) and fatigue ($n=3$). During the walking test, patients took a median of 399 steps and covered a median distance of 170 m.

Regarding the actual steps, these were determined by averaging the two values manually counted by each researcher. The ICC between the two researchers was 1.00. In 40 out of 45 cases, the step counts recorded by the two researchers were identical. In the remaining five cases, the differences were as follows: two instances with a difference of one step, two instances with a difference of two steps, and one instance with a difference of three steps.

Validity and accuracy of activity trackers

Table 2 shows the ICCs, mean difference in steps, MPE, and MAPE for the three activity trackers. Step counts recorded by the three trackers showed excellent or nearly excellent agreement with the actual steps taken by the 10 healthy staff members; the ICC ranged from 0.85 to 0.91. They also demonstrated acceptable accuracy, with all MAPE values at $< 10\%$.

All three trackers tended to underestimate the actual steps taken by patients after major abdominal surgery. On considering both percentage errors and ICCs, the HONOR appeared to be less inaccurate. On average, the HONOR underestimated the number of steps taken by 80 steps (23.3% of the actual steps, ICC = 0.75, MAPE = 24.5%); this was followed the Fitbit, which underestimated the steps taken by 205 steps (56.4% of the actual steps, ICC = 0.39, MAPE = 57.3%), and the MI, which underestimated the steps taken by 240 steps (65.5% of the actual steps, ICC = 0.35, MAPE = 65.7%). Notably, the MI and Fitbit trackers recorded 0 steps in 15 (33.3%) and 14

(31.1%) patients, respectively, leading to a percentage error of up to 100% in these cases. A sensitivity analysis, excluding patients with zero step recordings, revealed that the median MAPE for the MI was 45.6% (IQR: 18.9%–81.1%, $n=30$), and for the Fitbit was 31.4% (IQR: 6.1%–31.4%, $n=31$).

The Bland–Altman plots supported these findings. Negative values of the mean difference between the tracker readings and manual step counts indicated underestimation by the trackers. Additionally, the plots revealed an upwards trend, where the bias clearly decreased with increasing step count, highlighting that the trackers' agreement improved as the number of steps increased. Regarding the range between the upper and lower limits of agreement, the MI and Fitbit had more widely scattered limits than the HONOR, indicating that these devices had less agreement in detecting step counts of patients after major abdominal surgery (Figure 1).

Influence of demographic, clinical, and gait parameters on the accuracy of activity trackers

For all trackers, the MAPEs correlated negatively with POD ($r = -0.626$ to -0.744), walking speed ($r = -0.714$ to -0.854), step length ($r = -0.466$ to -0.615), and cadence ($r = -0.681$ to -0.790), and positively with the number of drains ($r = 0.450$ – 0.514). The difference and association between step count accuracy and other demographic and clinical variables were not statistically significant (Table 3).

As shown in Figure 2(a), a higher level of MAPEs was observed in the early stages after surgery (POD 1–3), which then reduced to a relatively low and steady level after POD 4. In particular, the MAPE for the HONOR decreased to below 10% on POD 4. The HONOR and Fitbit showed excellent accuracy in patients who had no abdominal drains, with MAPEs of $< 5\%$ and $\sim 10\%$, respectively. Notably, only three participants were without drains, which should be considered when interpreting these low MAPE values (Figure 2(b)). For patients who walked at speeds of below 0.44 m/s, with step lengths of < 0.34 m or cadences of fewer than 60 steps per minute (SPM), all trackers showed low accuracy, with MAPEs above 50%. For the HONOR, the MAPEs decreased to nearly 10% when speeds, strides, or cadences exceeded 0.45 m/s, 0.45 m, or 75 SPM, respectively (Figure 2(c) to (e)).

On average, patients walked at low speeds (from 0.37 to 0.45 m/s) with short step lengths (of below 0.41 m) and had at least two abdominal drains from POD 1 to 3 (Figure 3).

Discussion

We examined the step count accuracy of wrist-worn activity trackers in patients after major abdominal surgery. The results showed that all of the tested trackers underestimated

Table 1. Characteristic of patient subjects.

	Median (IQR)	Mean (SD)	Range	N (%)
Age	–	52.2 (13.6)	23–77	45 (100)
Male	–	–	–	27 (60)
Height (cm)	–	162.5 (7.1)	148–179	45 (100)
Weight (kg)	–	59.8 (10.5)	39.0–79.5	45 (100)
BMI	–	22.5 (2.8)	17.3–29.4	45 (100)
Postoperative diagnosis				
Tumor	–	–	–	26 (57.8)
Non-tumor	–	–	–	19 (42.2)
Surgical approach				
Open	–	–	–	37 (82.2)
Laparoscopy	–	–	–	8 (17.8)
Surgical group				
Pancreas	–	–	–	21 (46.7)
Liver	–	–	–	15 (33.3)
Biliary tract	–	–	–	9 (20)
Postoperative days	4 (2.5, 5.5)	–	1–12	45 (100)
Number of drains	2 (1, 2)	–	0–6	45 (100)
First time of postoperative walking	–	–	–	9 (20)
Pain scores before the modified 6MWT	2 (0, 3)	–	0–5	45 (100)
Pain scores after the modified 6MWT	2 (0, 4)	–	0–6	45 (100)
Completed the modified 6MWT	–	–	–	37 (82.2)
Total time walked (seconds)	360 (360, 360)	–	152–360	45 (100)
Total distance walked (m)	170 (127.5, 210)	–	80–340	45 (100)
Total steps observed	399 (328, 454)	–	201–590	45 (100)
Average gait speed (m/s)		0.50 (0.14)	0.28–0.94	45 (100)
Average step length (m)		0.43 (0.07)	0.32–0.58	45 (100)
Average cadence (steps/minutes)		69.6 (11.1)	48.3–98.3	45 (100)

IQR: interquartile range; SD: standard deviation; BMI: body mass index; 6MWT: 6-minute walk test.

Table 2. The intraclass correlation coefficient (ICC), mean difference, mean percentage error (MPE), and mean absolute percentage error (MAPE) for three activity trackers.

	ICC (95% CI)	Mean difference (SD)	MPE (%; SD)	Range (%)	MAPE (%; SD)	Range (%)
Staff (n = 10)						
MI 4	0.87 (0.44–0.97)	−0.1 (14.3)	0.2 (7.6)	−15.2–10.8	5.6 (4.8)	1.2–15.2
HONOR 5	0.91 (0.41–0.98)	7.8 (8.4)	4.5 (4.6)	−1.8–10.8	5.1 (3.9)	1.1–10.8
Fitbit Inspire HR	0.85 (0.44–0.96)	6.6 (18.0)	3.5 (9.7)	−12.5 to 14.9	9.1 (3.8)	3.1–14.9
Patient (n = 45)						
MI 4	0.35 (−0.19 to 0.70)	−240.1 (129.4)	−65.5 (36.8)	−100.0–2.8	65.7 (36.4)	1.4–100.0
HONOR 5	0.75 (0.19–0.89)	−80.2 (92.3)	−23.5 (27.7)	−85.0–8.9	24.5 (26.7)	0–85.0
Fitbit Inspire HR	0.39 (−0.28 to 0.71)	−205.1 (145.2)	−56.4 (40.6)	−100.0–9.2	57.3 (39.2)	0.6–100.0

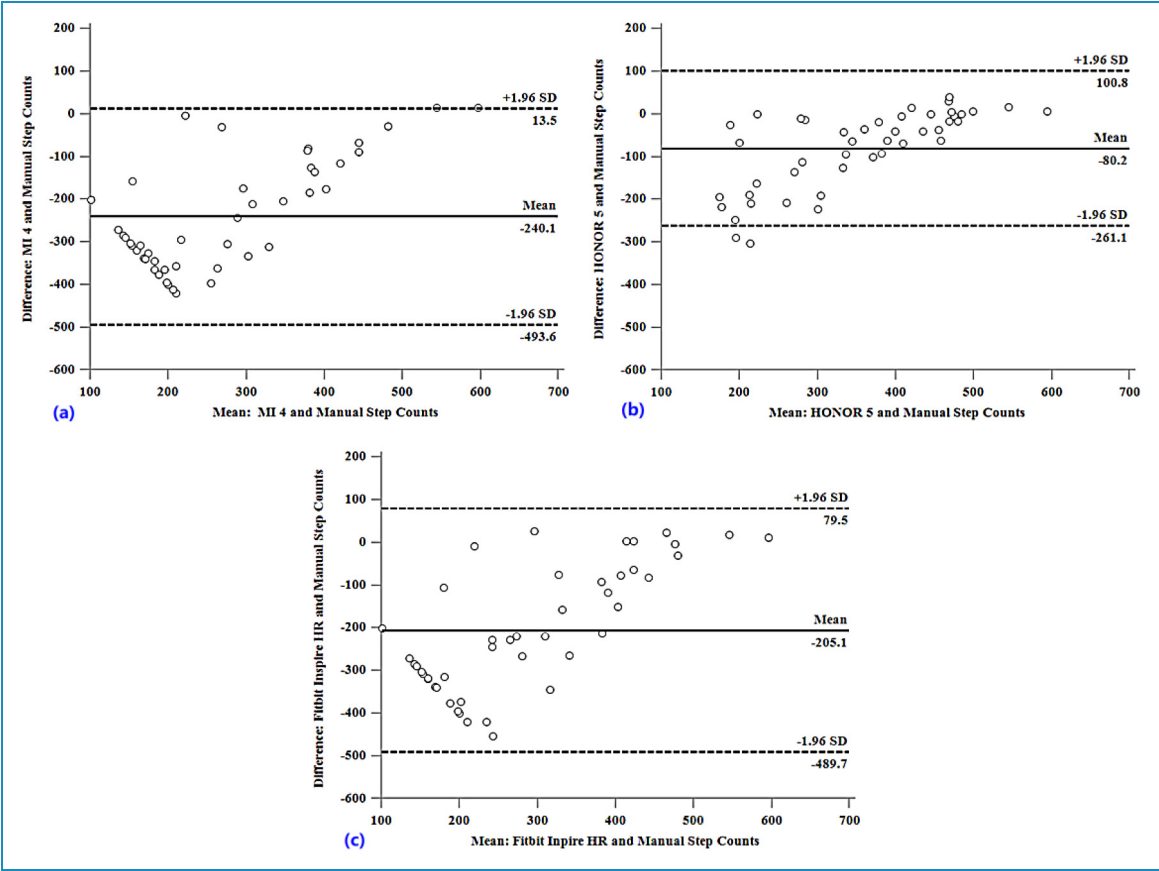


Figure 1. Bland–Altman plots. The middle solid line indicates the mean difference between tracker-recorded steps and manual step counts, and the dashed lines depict the 95% confidence intervals of the limits of agreement. MI 4 versus manual step counts (a); HONOR 5 versus manual step counts (b); Fitbit Inspire HR versus manual step counts (c).

actual steps to a different extent. Among them, the HONOR appeared to outperform the MI and Fitbit. However, it displayed a high percentage error (MAPE = 24.5%) and exceeded the 10% accuracy threshold that is considered clinically acceptable. Furthermore, all trackers demonstrated greater error in detecting the steps of patients

Table 3. Influence of demographic, clinical and gait parameters on step count accuracy.

	Mean absolute percentage error of trackers		
	MI 4	HONOR 5	Fitbit Inspire HR
Age	−0.197 ^a	−0.130 ^a	−0.148 ^a
Gender	−0.047 ^b	−0.012 ^b	−0.353 ^b
Height	0.142 ^a	0.135 ^a	0.135 ^a
Weight	0.137 ^a	0.117 ^a	0.171 ^a
BMI	0.100 ^a	0.048 ^a	0.135 ^a
Postoperative diagnosis	−1.101 ^b	−0.758 ^b	−0.618 ^b
Surgical approach	−1.634 ^b	−0.713 ^b	−1.703 ^b
Surgical groups	0.537 ^c	2.147 ^c	0.893 ^c
Postoperative days	−0.744 ^{a***}	−0.702 ^{a***}	−0.626 ^{a***}
Number of drains	0.514 ^{a***}	0.477 ^{a***}	0.450 ^{a**}
First postoperative walking	−0.116 ^b	−1.334 ^b	−0.893 ^b
Pain scores before the modified 6MWT	0.147 ^a	0.069 ^a	0.063 ^a
Pain scores after the modified 6MWT	0.170 ^a	0.163 ^a	0.144 ^a
Completed the modified 6MWT	−0.968 ^b	−1.470 ^b	−0.512 ^b
Average walking speed	−0.854 ^{a***}	−0.714 ^{a***}	−0.736 ^{a***}
Average step length	−0.615 ^{a***}	−0.466 ^{a***}	−0.475 ^{a***}
Average cadence	−0.790 ^{a***}	−0.681 ^{a***}	−0.739 ^{a***}

BMI: body mass index; 6MWT: 6-minute walk test.

^aSpearman's rank correlation (*r*).

^bMann-Whitney *U* test (*Z*).

^cKruskal-Wallis *H* test (*H*).

P* ≤ .05, *P* ≤ .001.

who walked at slower speeds, had shorter step lengths, had more abdominal drains, and participated in the walking tests on earlier PODs.

Wrist-worn activity trackers are easy to use and allow self-monitoring and goal setting owing to immediate

feedback by the device. Notably, goal-directed mobilization can encourage and increase early postoperative mobilization, thereby enhancing postoperative recovery.^{21,25} However, wide step count variations between trackers and actual steps could induce frustration and disappointment.³² This is because any over- or underestimation may lead to ambiguity regarding the achievement of daily mobilization goals; estimation errors may also lead to inaccuracies in the assessment of patient recovery based on the relationship between step counts and postoperative outcomes. Our results indicated that all three trackers tended to undercount steps by 23.5%–65.5%. Notably, the MI and Fitbit did not record any of the steps taken during the walking tests in nearly a third of the patients. This high percentage of zero recordings could be an issue when these trackers are used for motivation.³³ Given that the trackers were inaccurate in this structured walking test; their performance is likely to be even worse in non-controlled conditions. Furthermore, the results of the preliminary test demonstrated that the same three trackers could accurately measure steps in healthy adults, thus excluding inherent tracker-related factors.

Consistent evidence^{34–36} has indicated that Fitbit devices, one of the most popular consumer wearable activity trackers in health promotion research, generally provide acceptable step count accuracy for healthy adults and those with chronic disease. However, our findings are similar to a prior study by Appelboom et al.,³⁷ which is one of the few assessments of step count accuracy of trackers in postoperative patients. They found that the Fitbit Zip, worn on the hip and ankle, underestimated steps by 81.4% (ICC = 0.326) and 26.1% (ICC = 0.837), respectively. Notably, it did not test wrist-worn trackers, and patients in that study walked at much lower speeds (interquartile range [IQR]: 0.156–0.357 m/s) than in ours (IQR: 0.40–0.60 m/s). This finding was actually expected; because other studies have also noted that wrist-worn activity trackers tend to underestimate steps in other populations with slow gait speeds.^{38,39} Nevertheless, it is necessary to increase the understanding of such inaccuracies in different populations.

Gait parameters such as slower walking speeds, shorter step lengths, or lower cadences are known to contribute to step count inaccuracies in commercially available activity trackers.^{31,33,39} These findings were confirmed in our study. For patients who walked at speeds, step lengths, or cadences exceeding 0.45 m/s, 0.45 m, or 75 SPM, respectively, the MAPEs of the HONOR device decreased from over 50% to nearly 10%. These results may be explained by the device algorithms. The algorithm for step identification is developed based on normal walking in healthy adults; the devices, therefore, have a higher threshold for the acceleration signal, which determines whether the movement detected represents a step; however, smaller accelerations generated during slower walking or by

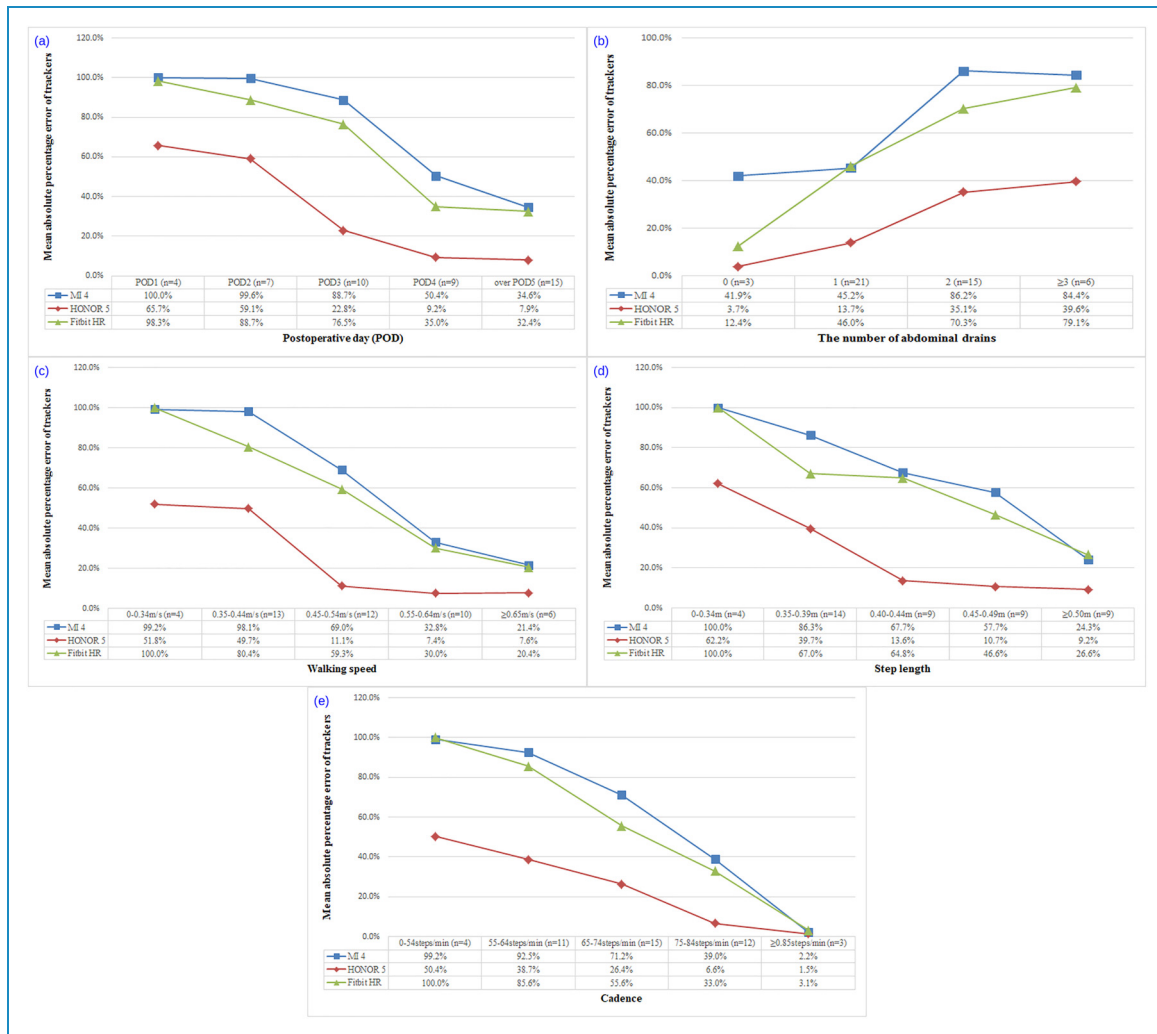


Figure 2. Graphs depicting the mean absolute percentage error of each tracker against the postoperative day (a), the number of abdominal drains (b), walking speed (c), step length (d), and cadence (e).

shorter strides are unlikely to trigger identification. Previous studies suggested that ankle placement may provide a solution for sufficient sensitivity in step detection among slow-walking populations.^{40,41} However, we did not select ankle placement for two reasons. Firstly, wrist-worn devices may not function optimally on the ankle, as manufacturers only recommend wrist placement for these trackers; secondly, postoperative patients may prefer to use trackers placed on the wrist rather than on the ankle. This is because visual feedback from wrist trackers is more convenient to read immediately and at any time; this aids motivation for improving early mobilization.

The presence of multiple drains is a common feature after major abdominal surgery. Most patients in this study underwent open pancreas and liver surgery and had a median of two drains in situ during the walking test. The results showed that the step count error increased with an increase in the number of drains; this finding provides

new insights into the postoperative step count inaccuracy of trackers. Patients with abdominal drains may have an altered walking pattern, because they tend to hold the drains while walking to ease the discomfort caused by traction. This is like walking with a walker or while pushing a shopping cart, with artificially constrained arm swing motion^{31,42,43}; this may explain the major performance drop in step detection.

Our findings suggest that PODs could also affect the step count accuracy of trackers; the MAPEs of the trackers decreased with increasing PODs in this cohort. Notably, the MAPEs of the HONOR decreased to below 10% on POD 4; there was also a sharp decrease between POD 3 and 4 with the other trackers. In line with our hypothesis, we found that patients in the early postoperative period, particularly between POD 1 and 3, walked at slower speeds, had shorter strides and lower cadences, and had a greater number of drains; this may explain why trackers showed

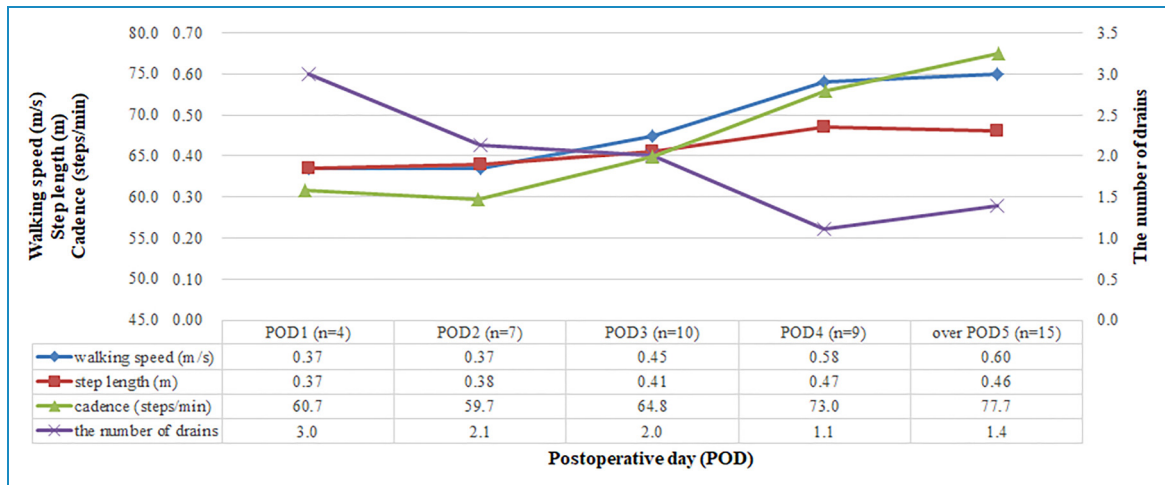


Figure 3. Graph depicting the relationship between walking speed, step length, and the number of abdominal drains with respect to postoperative days. The X-axis represents the postoperative days, and the Y-axis represents walking speed, step length, cadence, and the number of drains, with different colors, respectively.

greater errors in step detection in the earlier PODs. Also, the proportion of patients in POD 1 or 2 was low in this study ($n = 11$; 24%), and it seemed to be that the higher proportion of those in the early postoperative period might lead to less step count accuracy of trackers. Thus, we recommend caution in interpreting and using step count data obtained by trackers in the early postoperative period. More advanced and flexible algorithm models need to be developed for step recognition, particularly for identifying early irregular walking patterns after surgery.

It was worth to mention that pain seemed to have no effect on step count accuracy. Patients were recruited in this validation study when they were actually walking or preparing to walk after surgery, usually in good conditions, and taking a structured walking test would not likely to be a burden. So, most of them had a relatively low level of pain severity before the modified 6MWT and there was no significant change in scores after that. In other words, patients seldom participated in the walking test when they felt pain; it was difficult to directly assess the influence of pain on trackers accuracy in the scheduled walking test.

Limitations

This study had several limitations. First, the majority of patients underwent open pancreas and liver surgery, and their walking patterns may have been particularly affected by the presence of multiple drains and open wounds. Our findings may not therefore be generalizable to all patients following abdominal surgery. Second, this study examined the accuracy of activity trackers in detecting steps during the walking test only; although walking is the main type of early postoperative mobilization, the specific corridor walk used for validation may not fully reflect the

complexity of the postoperative clinical environment. Future studies will need to examine whether trackers can reliably measure activities in the real-world clinical setting, including both pre- and post-surgery patients over a longer period. Third, this is the only study to date to examine the influence of clinical factors on the step count accuracy of trackers following abdominal surgery; however, the sample size for some factors was small. Only four patients walked on POD 1 and only three walked without any drains; this made it difficult to determine the accuracy of trackers for those factors. Furthermore, pain or fatigue may affect postoperative activity; however, it was difficult to directly assess their effects in the scheduled walking test, which need to be confirmed in future studies. Finally, we used multiple approaches including the ICC, Bland–Altman method, and MAPEs, to examine step count accuracies of the trackers; although the findings provided novel insights into the inaccuracies in the postoperative population, they were only limited to accuracy assessment and could not thoroughly evaluate the utility of wearable activity trackers during early postoperative mobilization.

Conclusions

In summary, our preliminary results indicate that wrist-worn activity trackers tend to undercount actual steps during the walking test in patients after major abdominal surgery. Although this is not unexpected, the accuracy of these trackers in a postoperative patient population and the extent of inaccuracies warrant further investigation. Additionally, this is the only study to explore the influence of clinical factors on the accuracy of multiple trackers following abdominal surgery; the results suggested that the

step count data obtained by trackers during the walking test on POD 1–3 should be considered with caution, and these trackers might not be reliable tools for measuring step counts in the early postoperative period. Future research with a larger sample is needed to comprehensively evaluate the accuracy, effectiveness, and feasibility of using an activity tracker over a longer duration and in free-living environments during the early postoperative period. Further work is also needed for developing new movement detection algorithms based on patient features in the early postoperative period and the clinical environment.

Acknowledgements: We would like to thank all patients for their voluntary participation in this study. We also appreciate the methodology experts for their insightful guidance on data analysis.

Authors' contributions: Zhi Li: conceptualization, methodology, and writing—original draft. Weiyan Feng: investigation, formal analysis, and visualization. Lili Zhou: investigation and validation. Shu Gong: writing—review and editing, and supervision.

Data availability: All data included in this study are available upon request by contact with the corresponding author.

Declaration of conflicting interests: The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical approval: This study was approved by the Biomedical Research Ethics Committee of the West China Hospital (IRB No. 2020204). Participants provided written and oral informed consent before participating. All study data were anonymous to ensure privacy and confidentiality. No compensation of any kind was provided to the participants enrolled in this study.

Funding: The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Science and Technology Department Foundation of Sichuan Province, West China Nursing Discipline Development Special Fund Project, Sichuan University (grant numbers 2019YFS0384 and HXHL19048).

Guarantor: GS

ORCID iD: Shu Gong  <https://orcid.org/0000-0003-4862-5125>

References

- Hughes MJ, Hackney RJ, Lamb PJ, et al. Prehabilitation before major abdominal surgery: a systematic review and meta-analysis. *World J Surg* 2019; 43: 1661–1668. PMID: 30788536.
- Staiger RD, Gerns E, Castrejón Subirà M, et al. Can early postoperative complications predict high morbidity and decrease failure to rescue following major abdominal surgery? *Ann Surg* 2020; 272: 834–839. PMID: 32925252.
- Veličković J, Feng C, Palibrk I, et al. The assessment of complications after major abdominal surgery: a comparison of two scales. *J Surg Res* 2020; 247: 397–405. PMID: 31676144.
- Schwab M, Brindl N, Studier-Fischer A, et al. Postoperative complications and mobilisation following major abdominal surgery with vs. without fitness tracker-based feedback (EXPELLIARMUS): study protocol for a student-led multi-centre randomised controlled trial (CHIR-Net SIGMA study group). *Trials* 2020; 21: 293. PMID: 32293519. PMCID: PMC7092422.
- Ljungqvist O, Scott M and Fearon KC. Enhanced recovery after surgery: a review. *JAMA Surg* 2017; 152: 292–298. PMID: 28097305.
- Yip VS, Dunne DFJ, Samuels S, et al. Adherence to early mobilisation: key for successful enhanced recovery after liver resection. *Eur J Surg Oncol* 2016; 42: 1561–1567. PMID: 27528466.
- van Zelm R, Coeckelberghs E, Sermeus W, et al. Variation in care for surgical patients with colorectal cancer: protocol adherence in 12 European hospitals. *Int J Colorectal Dis* 2017; 32: 1471–1478. PMID: 28717841.
- Wolk S, Distler M, Müsle B, et al. Adherence to ERAS elements in major visceral surgery – an observational pilot study. *Langenbecks Arch Surg* 2016; 401: 349–356. PMID: 27013325.
- Brown CJ, Friedkin RJ and Inouye SK. Prevalence and outcomes of low mobility in hospitalized older patients. *J Am Geriatr Soc* 2004; 52: 1263–1270. PMID: 15271112.
- Hoyer EH, Friedman M, Lavezza A, et al. Promoting mobility and reducing length of stay in hospitalized general medicine patients: a quality-improvement project. *J Hosp Med* 2016; 11: 341–347. PMID: 26849277.
- Foss NB, Kristensen MT and Kehlet H. Prediction of post-operative morbidity, mortality and rehabilitation in hip fracture patients: the cumulated ambulation score. *Clin Rehabil* 2006; 20: 701–708. PMID: 16944827.
- Brooks D, Hunter JP, Parsons J, et al. Reliability of the two-minute walk test in individuals with transtibial amputation. *Arch Phys Med Rehabil* 2002; 83: 1562–1565. PMID: 12422326.
- Moriello C, Mayo NE, Feldman L, et al. Validating the six-minute walk test as a measure of recovery after elective colon resection surgery. *Arch Phys Med Rehabil* 2008; 89: 1083–1089. PMID: 18503803.
- Vanhees L, Lefevre J, Philippaerts R, et al. How to assess physical activity? How to assess physical fitness? *Eur J Cardiovasc Prev Rehabil* 2005; 12: 102–114. PMID: 15785295.
- Ainsworth BE. How do I measure physical activity in my patients? Questionnaires and objective methods. *Br J Sports Med* 2009; 43: 6–9. PMID: 18718977.
- Tang MSS, Moore K, McGavigan A, et al. Effectiveness of wearable trackers on physical activity in healthy adults: systematic review and meta-analysis of randomized controlled trials. *JMIR Mhealth Uhealth* 2020; 8: e15576. PMID: 32706685. PMCID: PMC7407266.
- Evenson KR, Goto MM and Furberg RD. Systematic review of the validity and reliability of consumer-wearable activity

- trackers. *Int J Behav Nutr Phys Act* 2015; 12: 159. PMID: 26684758. PMCID: PMC4683756.
18. Höchsmann C, Knaier R, Eymann J, et al. Validity of activity trackers, smartphones, and phone applications to measure steps in various walking conditions. *Scand J Med Sci Sports* 2018; 28: 1818–1827. PMID: 29460319.
 19. Freedson PS and John D. Comment on “estimating activity and sedentary behavior from an accelerometer on the hip and wrist”. *Med Sci Sports Exerc* 2013; 45: 962–963. PMID: 23594509.
 20. Cook DJ, Thompson JE, Prinsen SK, et al. Functional recovery in the elderly after major surgery: assessment of mobility recovery using wireless technology. *Ann Thorac Surg* 2013; 96: 1057–1061. PMID: 23992697.
 21. Daskivich TJ, Houman J, Lopez M, et al. Association of wearable activity monitors with assessment of daily ambulation and length of stay among patients undergoing major surgery. *JAMA Netw Open* 2019; 2: e187673. PMID: 30707226. PMCID: PMC6484591.
 22. Low CA, Bovbjerg DH, Ahrendt S, et al. Fitbit step counts during inpatient recovery from cancer surgery as a predictor of readmission. *Ann Behav Med* 2018; 52: 88–92. PMID: 29538623; PMCID: PMC6051699.
 23. Sun V, Dumitra S, Ruel N, et al. Wireless monitoring program of patient-centered outcomes and recovery before and after major abdominal cancer surgery. *JAMA Surg* 2017; 152: 852–859. PMID: 28593266. PMCID: PMC5607084.
 24. Carmichael H, Overbey DM, Hosokawa P, et al. Wearable technology – a pilot study to define “normal” postoperative recovery trajectories. *J Surg Res* 2019; 244: 368–373. PMID: 31323392.
 25. Wolk S, Linke S, Bogner A, et al. Use of activity tracking in major visceral surgery – the enhanced perioperative mobilization trial: a randomized controlled trial. *J Gastrointest Surg* 2019; 23: 1218–1226. PMID: 30298422.
 26. Massouh F, Martin R, Chan B, et al. Is activity tracker-measured ambulation an accurate and reliable determinant of postoperative quality of recovery? A prospective cohort validation study. *Anesth Analg* 2019; 129: 1144–1152. PMID: 30379677.
 27. Yin T-C, Huang C-W, Tsai H-L, et al. Smartband use during enhanced recovery after surgery facilitates inpatient recuperation following minimally invasive colorectal surgery. *Front Surg* 2021; 7: 608950. PMID: 33585547; PMCID: PMC7874082.
 28. Tang JH, Wang B, Chow JLJ, et al. Improving postoperative mobilisation rates in patients undergoing elective major hepatopancreatobiliary surgery. *Postgrad Med J* 2021; 97: 239–247. PMID: 33184138.
 29. Walker RK, Hickey AM and Freedson P. Advantages and limitations of wearable activity trackers: considerations for patients and clinicians. *Clin J Oncol Nurs* 2016; 20: 606–610. PMID: 27857250.
 30. Wahl Y, Düking P, Droszez A, et al. Criterion-validity of commercially available physical activity tracker to estimate step count, covered distance and energy expenditure during sports conditions. *Front Physiol* 2017; 8: 725. PMID: 29018355. PMCID: PMC5615304.
 31. Schaffer SD, Holzapfel SD, Fulk G, et al. Step count accuracy and reliability of two activity tracking devices in people after stroke. *Physiother Theory Pract* 2017; 33: 788–796. PMID: 28777710.
 32. Rosenberg D, Kadokura EA, Bouldin ED, et al. Acceptability of Fitbit for physical activity tracking within clinical care among men with prostate cancer. *AMIA Annu Symp Proc* 2017; 2016: 1050–1059. PMID: 28269902. PMCID: PMC5333209.
 33. Wong CK, Mentis HM and Kuber R. The bit doesn’t fit: evaluation of a commercial activity-tracker at slower walking speeds. *Gait Posture* 2018; 59: 177–181. PMID: 29049964.
 34. Chevance G, Golaszewski NM, Tipton E, et al. Accuracy and precision of energy expenditure, heart rate, and steps measured by combined-sensing Fitbits against reference measures: systematic review and meta-analysis. *JMIR Mhealth Uhealth* 2022; 10: e35626. PMID: 35416777; PMCID: PMC9047731.
 35. Fuller D, Colwell E, Low J, et al. Reliability and validity of commercially available wearable devices for measuring steps, energy expenditure, and heart rate: systematic review. *JMIR Mhealth Uhealth* 2020; 8: e18694. PMID: 32897239; PMCID: PMC7509623.
 36. Germini F, Noronha N, Borg Debono V, et al. Accuracy and acceptability of wrist-wearable activity-tracking devices: systematic review of the literature. *J Med Internet Res* 2022; 24: e30791. PMID: 35060915; PMCID: PMC8817215.
 37. Appelboom G, Taylor BE, Bruce E, et al. Mobile phone-connected wearable motion sensors to assess postoperative mobilization. *JMIR Mhealth Uhealth* 2015; 3: e78. PMID: 26220691. PMCID: PMC4705357.
 38. Hartung V, Sarshar M, Karle V, et al. Validity of consumer activity monitors and an algorithm using smartphone data for measuring steps during different activity types. *Int J Environ Res Public Health* 2020; 17: 9314. PMID: 33322833. PMCID: PMC7764011.
 39. Chow JJ, Thom JM, Wewege MA, et al. Accuracy of step count measured by physical activity monitors: the effect of gait speed and anatomical placement site. *Gait Posture* 2017; 57: 199–203. PMID: 28666177.
 40. Klassen TD, Simpson LA, Lim SB, et al. “Stepping up” activity poststroke: ankle-positioned accelerometer can accurately record steps during slow walking. *Phys Ther* 2016; 96: 355–360. PMID: 26251478. PMCID: PMC4774387.
 41. Simpson L, Eng J, Klassen T, et al. Capturing step counts at slow walking speeds in older adults: comparison of ankle and waist placement of measuring device. *J Rehabil Med* 2015; 47: 830–835. PMID: 26181670.
 42. Alinia P, Cain C, Fallahzadeh R, et al. How accurate is your activity tracker? A comparative study of step counts in low-intensity physical activities. *JMIR Mhealth Uhealth* 2017; 5: e106. PMID: 28801304. PMCID: PMC5572056.
 43. Bruijn SM, Meijer OG, Beek PJ, et al. The effects of arm swing on human gait stability. *J Exp Biol* 2010; 213: 3945–3952. PMID: 21075935.