Danielsson et al. BMC Digital Health (2024) 2:45
https://doi.org/10.1186/s44247-024-00101-z

BMC Digital Health

RESEARCH

Open Access

Potential sources of inaccuracy in the Apple watch series 4 energy expenditure estimation algorithm during wheelchair propulsion

Marius Lyng Danielsson[1*], Roya Doshmanziari[2], Berit Brurok[3,4], Matthijs Ferdinand Wouda[5,6] and Julia Kathrin Baumgart[1]

Abstract

Background The Apple Watch (AW) was the first smartwatch to provide wheelchair user (WCU) specific information on energy expenditure (EE), but was found to be inaccurate (i.e., it underestimated) and imprecise (i.e., the underestimation was variable). Insight is therefore needed into where these inaccuracies/imprecisions originate. Accordingly, the aim of this study was to investigate how much of the variation in AW EE is explained by heart rate (HR), in addition to other factors such as body mass and height, sex, age, physical activity level and disability.

Methods Forty participants (20 WCU, 20 non-disabled) performed three 4-min treadmill wheelchair propulsion stages at different speed-incline combinations, on three separate days, while wearing an AW series 4 (setting: "outdoor push walking pace"). Linear mixed model analyses investigated how much of the variation in AW EE (kcal·min^{-1}) is explained by the fixed effects AW HR (beats·min^{-1}), body mass and height, sex, age, physical activity level and disability. Participant-ID was included as random-intercept effect. The same mixed model analyses were conducted for criterion EE and HR. Marginal R^2 (R^2m; fixed effects only) and conditional R^2 (R^2c; fixed and random effects) values were computed. An R^2m close to zero indicates that the fixed effects alone do not explain much variation.

Results Although criterion HR explained a significant amount of variation in criterion EE (R^2m: 0.44, R^2c: 0.92, $p < 0.001$), AW HR explained little variation in AW EE (R^2m: 0.06, R^2c: 0.86, $p < 0.001$). In contrast, body mass and sex explained a significant amount of variation in AW EE (R^2m: 0.74, R^2c: 0.79, $p < 0.001$). No further improvements in fit were achieved by adding body height, age, physical activity level or disability to the AW EE model (R^2m: 0.75, R^2c: 0.79, $p = 0.659$).

Conclusion Our results remain inconclusive on whether AW heart rate is used as factor to adjust for exercise intensity in the black box AW EE estimation algorithms. In contrast, body mass explained much of the variation in AW EE, indicating that the AW EE estimation algorithm is very reliant on this factor. Future investigations should explore better individualization of EE estimation algorithms.

Keywords Disability, Upper-body exercise, Digital health, Accuracy, Smart watch, Fitness trackers

*Correspondence:
Marius Lyng Danielsson
mlyngd@gmail.com
Full list of author information is available at the end of the article

Background

Smartwatches, besides functioning as digital clocks, are increasingly equipped with sensors for tracking a variety of parameters related to physical activity and health [1, 2]. One of these parameters is energy expenditure (EE). Energy expenditure, often measured in kilocalories (kcal), refers to the amount of energy an individual uses to maintain essential body functions (e.g., respiration, circulation, digestion) and during physical activity [3]. Feedback on EE, especially when combined with information on one's energy intake, can serve as a tool for maintaining/reducing body mass. This is important as obesity is a growing health problem globally [4], and especially for wheelchair users (WCU) since the prevalence is 2.5 times greater in this group compared to non-disabled individuals (ND) [5]. The higher prevalence in WCU is related to impairment-related lower daily EE of ~ 5–40% compared to ND [6–8], and higher inactivity levels [9–11]. In this context, the meaningfulness of the feedback provided by smartwatches depends heavily on the accuracy and precision of the estimated parameter.

The accuracy of EE estimated by current devices is inadequate even in ND [12, 13], and in WCU, Moreno et. al. [14] reported that the WCU-adapted algorithm of the Apple Watch (AW) series 1 underestimated EE with a mean absolute percentage error of ~ 29%. Danielsson et. al. [15] corroborated those findings using a more recent version of the AW (i.e., series 4), where no improvement in this underestimation was found (mean absolute percentage error of ~ 30%). Furthermore, the underestimation was variable for the individual WCU (i.e., the imprecision was high), and the underestimation of the AW EE increased at higher exercise intensities [15]. This increased underestimation indicates that the AW EE estimation algorithm(s) do not satisfactorily adjust for exercise intensity.

Heart rate (HR) is commonly used to track intensity during daily physical activity and exercise [16, 17], and the AW has been promoted as one of the most accurate smartwatches for estimating HR in diverse ND cohorts and people with cardiovascular disease [18–22]. However, the HR-EE relationship varies with exercise modalities/intensities, cohorts, and between and within individuals [23–27]. HR alone will therefore not account for all variation in EE. The most important factor that influences total EE is fat free mass [28–31], which is typically obtained from a body composition scan. Information on fat free mass is often not available and therefore not considered in the AW EE estimation algorithms. However, other factors that relate to fat free mass, such as body mass (BM), body height, sex and age may be used as substitutes. Individuals with higher BM typically have more fat free mass [32], and a larger total energy cost to maintain bodily functions [33–36]. Furthermore, for a given BM, females [37], older individuals [38], individuals with lower physical activity levels (PAL) [39] and/or with a disability [40] commonly have lower fat free mass. It has not yet been investigated to what extent HR, and the aforementioned factors account for variation in the AW EE. Such investigations may give indirect indications as to which factors the EE estimation black box algorithms take or do not take into account.

Therefore, the primary aim of this study was to investigate how much of the variation in AW EE is explained by HR and to evaluate the extent to which other factors such as BM and body height, sex, age, PAL and having a disability (or not) explain additional variation in AW EE.

Methods

Overall design

Participants performed several stages of treadmill wheelchair propulsion at different speed-incline combinations, while wearing an AW series 4 and two criterion devices (a Vyntus ergospirometer for recording energy expenditure, and a Polar H10 monitor for recording HR). Analyses were performed to investigate how much of the variation in AW EE is explained by AW HR, BM and body height, sex, age, PAL and disability. The same analyses were conducted for criterion EE and HR.

Participants

The data included in the current study is collected from the same participants as described in Lyng Danielsson et al. [15]. Forty participants (20 WCU and 20 ND, 11 males and 9 females within each group) were included. Participant characteristics are summarized in Table 1. The inclusion criteria for participation were as follows: age 18 – 60 years, manual wheelchair used for mobility on a daily basis (WCU only) and no health- or injury related issues that could be aggravated by physical exertion and/or wheelchair propulsion. The WCU group was heterogeneous with respect to the type of impairment and included individuals with spinal cord injury ($n = 11$), spina bifida ($n = 2$), and cerebral palsy ($n = 2$). The remaining five WCU had other neurological-, musculo-skeletal- or joint impairments. The PAL, which was based on the international physical activity questionnaire (IPAQ) score (see Data analysis section for calculations), was lower in WCU (low: $n = 3$, moderate: $n = 8$, high: $n = 9$) compared to ND (moderate: $n = 9$, high: $n = 11$). Sport associations, user organizations, social media, personal contacts and our work environment were the platforms used for participant recruitment. Testing was performed in the Elite Sports Science laboratory at the Norwegian University of Science and Technology (NTNU) in Trondheim, Norway.

Table 1 Participant characteristics with values presented as Mean ± SD

Groups	Sex	Age (years)	Measured BM (kg)	Self-reported BM (kg)	Body Height (cm)	BMI (kg/m^2)
Both	All	35.3 ± 11.8	74.8 ± 15.2	73.9 ± 14.2	174.5 ± 10.9	24.5 ± 4.1
(*n* = 40)	Males	36.3 ± 12.2	81.1 ± 11.9	80.7 ± 10.9	181.9 ± 7.2	24.4 ± 2.8
	Females	34.1 ± 11.5	67.1 ± 15.6	65.5 ± 13.6	165.3 ± 7.1	24.6 ± 5.3
WCU	All	37.4 ± 12.6	74.5 ± 18.6	73.6 ± 16.7	172.5 ± 12.2	24.9 ± 5.3
(*n* = 20)	Males	40.0 ± 12.9	80.4 ± 14.3	79.9 ± 12.3	180.5 ± 8.5	24.5 ± 3.3
	Females	34.1 ± 12.1	67.2 ± 21.3	65.9 ± 18.7	162.7 ± 8.0	25.3 ± 7.2
ND	All	33.3 ± 10.8	75.2 ± 11.4	74.2 ± 11.7	176.2 ± 9.9	24.2 ± 2.4
(*n* = 20)	Males	32.6 ± 10.8	81.9 ± 9.4	81.5 ± 9.9	183.5 ± 5.9	24.3 ± 2.4
	Females	34.0 ± 11.5	67.0 ± 7.8	65.2 ± 6.4	167.3 ± 5.3	23.9 ± 2.6

Abbreviations: *BM* Body mass, *BMI* Body mass index, *WCU* Wheelchair users, *ND* Non-disabled

Study protocol

Three test days were performed within two weeks to avoid significant changes in body composition, and with a minimum of 24 h between each test to minimize the effect of fatigue. The tests were scheduled on approximately the same time of day to account for potential diurnal variations. Participants performed wheelchair propulsion on a treadmill which was set at different speed-incline combinations. Each test day was performed on a specific incline (0.5, 2.5 or 5%) and in a counterbalanced order to avoid sequence bias. Each test day consisted of a 5-min warm-up at 0.5% incline with a self-chosen speed that corresponded to a rating of perceived exertion of 7–9 on the Borg scale [41], followed by three 4-min stages with increasing speeds (Table 2). The stage speeds were pre-set based on pilot testing with the intention of being feasible for all participants and therefore adjusted for sex and level of impairment. Information on age, sex, BM (both self-reported and measured), and body height was collected on the first test day. Participants were asked to refrain from high intensity training and consuming alcohol 24 h prior to testing, avoid caffeine on the day of testing, and not consume any food two hours before testing.

Equipment

BM was measured using a Kistler force plate (Kistler 9286BA; Kistler Instruments AG, Winterthur, Switzerland). For the WCU group this was done while sitting in their own wheelchair and BM was obtained by subtracting the wheelchair mass (range: 6.5–18.2 kg) from the total mass. ND were weighed while standing (without any equipment). Criterion EE was calculated from gas exchange data, which was collected using a facemask (7450 V2 Series, Hans Rudolph KC, KS) attached to a Vyntus CPX ergospirometer with a mixing chamber (Vyaire, Medical GmbH, Germany). The Vyntus CPX was calibrated prior to each test and was found to have excellent accuracy with a measurement error of ~ 1–2% for $\dot{V}O_2$ and $\dot{V}CO_2$ [42]. Criterion HR was measured using a Polar M400HR monitor and Polar H10 chest strap (Polar Electro Oy, Finland). The Polar H10 chest strap was found to have good accuracy with a measurement error below 4% in running [43].

The Apple Watch series 4 (Apple, Inc., CA, USA, software version OS 7.3.3) was used during the data collection and tracks HR using photoplethysmography and movement related features using a built-in accelerometer and gyroscope. Due to the rapid technological development and prolonged data collection (November 2020 – December 2022 including covid-19 related lockdown periods), the AW series 4 may now be considered outdated on certain aspects. Improvements have been reported on the estimation of step counts from the AW series 1 to series 4 [44]. However, no improvement was found for the EE estimation of the AW series 4 compared

Table 2 Overview of the standardized speeds used on the three test days (0.5, 2.5 or 5% incline), for male participants (without tetraplegia) and female participants or male tetraplegic wheelchair users

Stages	0.5% day		2.5% day		5% day	
	Males	Females/Tetraplegia	Males	Females/Tetraplegia	Males	Females/Tetraplegia
1	4 km·h	3 km·h	3 km·h	2 km·h	2 km·h	1 km·h
2	6 km·h	5 km·h	4 km·h	3 km·h	3 km·h	2 km·h
3	8 km·h	7 km·h	5 km·h	4 km·h	4 km·h	3 km·h

to the AW series 1, which still was ~ 30% [14, 15]. Furthermore, during wheelchair propulsion, the error of the AW series 4 HR when compared to the Polar H10 monitor, was ~ 8% [15]. Participants wore the AW on their non-dominant hand. Participant characteristics were registered in an app connected to the watch by the test personnel before testing started, and the activity setting "outdoor push walking pace" was selected for all stages.

WCU used their personal wheelchair and ND used a standardized wheelchair (Küschall K-Series Attract, Invacare, Oslo, Norway, mass: 11.7kg). Propulsion stages were performed on a motorized 5 × 3m treadmill (Forcelink Technology, Culemborg, The Netherlands) with the wheelchairs secured to a mobile traverse safety bar with safety stoppers to prevent participants from rolling off the treadmill (set up visualized in Lyng Danielsson [15]).

Data analysis

EE and HR from AW and the criterion devices were calculated as average values over the entire four-minute period of each stage. Average criterion values were preferred over steady state since the AW only displayed estimated EE and HR average for each entire stage and no continuous data. The Weir formula [45] was used to calculate criterion EE from 10-s average $\dot{V}O_2$ and $\dot{V}CO_2$ value (Eq. 1):

$$EE\left(Kcal \cdot min^{-1}\right) = 3.941 \cdot VO_2\left(L \cdot min^{-1}\right) + 1.106 \cdot VCO_2\,(L \cdot min^{-1}) \tag{1}$$

The PAL of all participants was categorized as low, moderate or high based on the score of the short version international physical activity questionnaire (IPAQ) for ND [46], and an adjusted version for WCUs [47].

Statistical analysis

Statistical analyses were conducted and figures created in RStudio, version 4.2.1 [48]. An α value of 0.05 was used to indicate statistical significance. Of the in total 360 observations, missing EE and HR data resulted in 299 observations for the AW and 339 observations for the criterion data that were included in the analyses. The reasons for the missing data are explained in more detail in our previous study [15].

The R *lme4* package and *lmer* function was used for conducting the linear mixed model analyses. These analyses were chosen due to missing data and the repeated-measures design of our data collection with corresponding dependency in data [49]. The linear mixed model analyses were performed to investigate how much of the variation in the dependent variable (i.e., AW EE) is explained by AW HR, and self-reported BM. Self-reported as opposed to measured BM, which was similar (see Table 1), was used with regards to the ecological validity and allowed for getting the AW user accounts ready prior to the data collection. Identical analyses were conducted with criterion EE and HR values to visualize the actual variation explained by HR or self-reported BM. Furthermore, the analyses on both the AW and criterion data were replicated with only the data of the slowest speed at each incline included, since these speeds correspond closest the activity setting "outdoor push walking pace". Speed and incline were not adjusted for in any of the analyses, as we were interested in the estimation capabilities across all intensities and not within each speed-incline combination. Random-intercept effects were included in all models to allow the HR-EE or BM-EE relationship to vary between participants and to account for dependency in the data, since all participants conducted several stages. Additionally, the following fixed effects were added to the models to further investigate their influence on the estimated AW/criterion EE: body height, sex, age, PAL and group (having a disability or not). Estimated regression coefficients Beta (β) were computed as indicators of the relationship between the fixed effects and dependent variable. Marginal R^2 (R^2m) (fixed effects only) and conditional R^2 (R^2c) (fixed and random effects) values were computed to evaluate the models' ability to explain variation in our data. An R^2m closer to zero indicates that the fixed effects alone do not explain much variation.

The function *anova* of the *afex* package was used to investigate if adding or removing factors from nested mixed models (i.e., models that have an overlap in their fixed effects) significantly improved or reduced model fit. Note that we used the maximum likelihood variance structure (as opposed to the restricted maximum likelihood structure) in the linear mixed model analyses, for the ANOVA model comparisons to be valid. Descriptive comparisons were used for models that were not nested.

Results

The fit of all linear mixed models to the AW and criterion EE data across all speed-incline combinations is visualized in Fig. 1, and the complete results of these analyses are provided in Supplement S1 for the AW data and S2 for the criterion data. For the AW data (Model 1 AW), HR contributed significantly, although to a limited extent, to explaining the variation in AW EE (β: 0.01 kcal·min^{-1}, $p < 0.001$; R^2m:

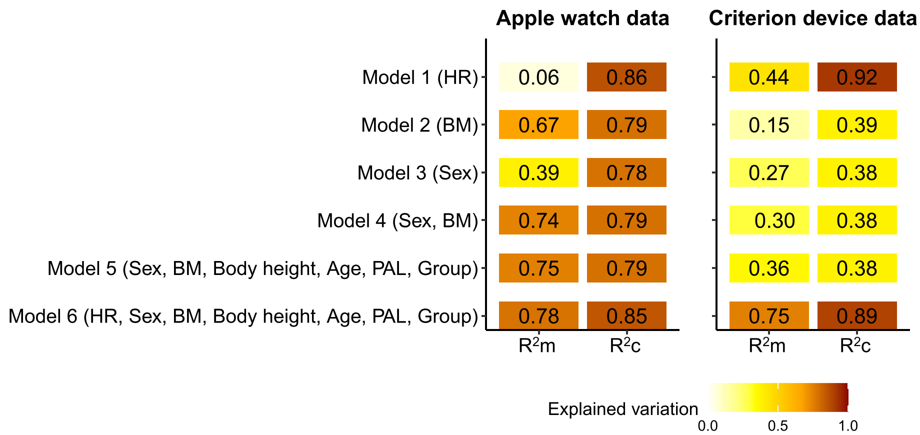

Fig. 1 Fit of the different mixed effect models (fixed effects are provided in brackets) to the Apple Watch and criterion device data as indicated by marginal (R^2m) and conditional R^2 (R^2c). All models include participant ID as a random intercept-effect. Group is divided into wheelchair users and non-disabled. Abbreviations: HR = Heart rate, BM = Body mass, PAL = Physical activity level

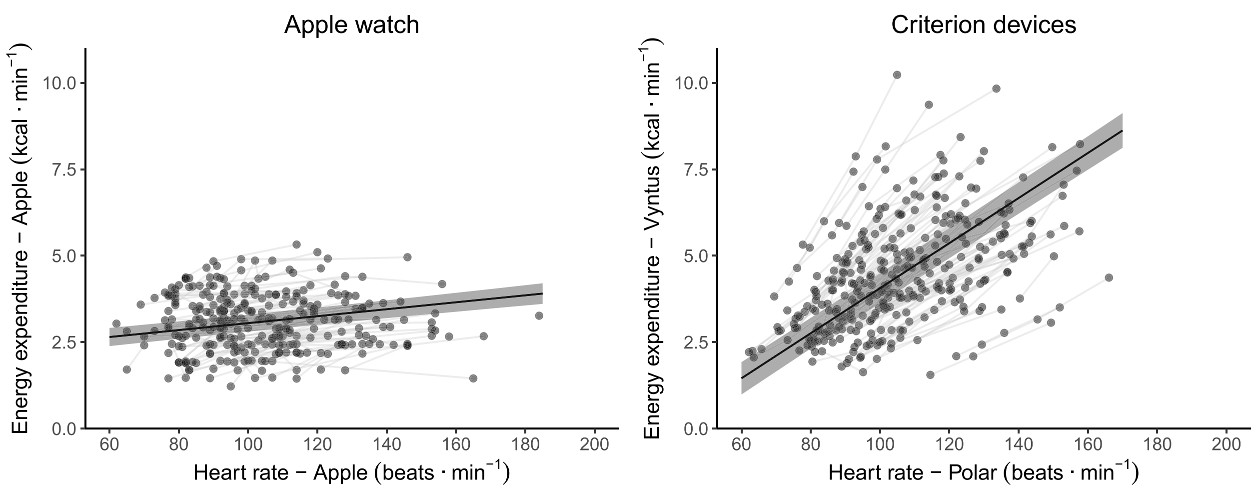

Fig. 2 Linear mixed effect models for energy expenditure with heart rate as the fixed effect and participant ID as random intercept-effect for Apple watch (**A**) and criterion values (**B**). The solid black lines are the regression lines, and the shaded areas are the 95% confidence interval of these regression lines

0.06, R^2c: 0.86) (Figs. 1 & 2). No significant amount of variation in AW EE was explained by AW HR, when running the same analyses only on the first speed of each incline (β: 0.0 kcal·min^{-1}, $p = 0.26$; R^2m: 0.003, R^2c: 0.88; Supplementary Fig. S3). For the criterion data (Model 1 criterion), HR explained significantly, and to a larger extent compared to Model 1 AW, the variation in criterion EE (β: 0.07 kcal·min^{-1}, $p < 0.001$; R^2m: 0.44, R^2c: 0.92) (Figs. 1 & 2).

BM when entered as the sole fixed effect (Model 2 AW) was the one that explained most variation in AW EE (β: 0.04 kcal·min^{-1}, $p < 0.001$; R^2m: 0.67, R^2c: 0.79). In contrast, BM alone (Model 2 criterion) explained less variation in criterion EE (β: 0.04 kcal·min^{-1}, $p < 0.001$; R^2m: 0.15, R^2c: 0.39) (see Figs. 1 & 3).

Additional models were created to investigate whether the combination of fixed factors explains more of the variation in AW and criterion EE (Fig. 1). Sex explained a significant amount of the variation in AW EE (R^2m: 0.39, $p < 0.001$), although less than BM (R^2m: 0.67, $p < 0.001$), but did significantly improve model fit when being added together with BM (R^2m: 0.74, both: $p < 0.001$; comparison AW Model 2 vs 4, $p < 0.001$). There was no significant improvement in explained variation of AW EE when adding body height, age, group and PAL (comparison AW Model 4 vs 5, $p = 0.659$). The amount of variation explained by the models not including HR (Models 2–5)

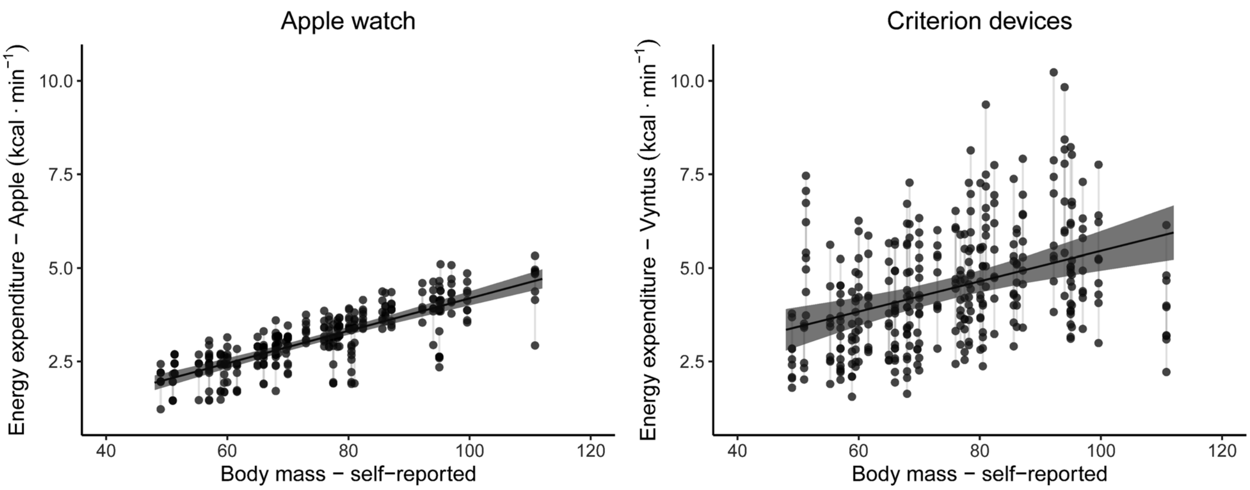

Fig. 3 Linear mixed effect models for energy expenditure, with self-reported body mass as the fixed effect and participant ID as the random-intercept effect for Apple watch (**A**) and criterion values (**B**). The solid black line is the regression line, and the shaded area is the 95% confidence interval of the curve fit

was considerably larger for criterion EE than AW EE (Fig. 1). Lastly, most variation in both AW EE and criterion EE was explained when all fixed effects were entered into the models (Models 6).

Discussion

Results from the present study show that Apple watch (AW) heart rate (HR) explained little of the variation in AW energy expenditure (EE), which is in contrast to criterion HR explaining a considerable amount of criterion EE. Furthermore, body mass (BM) and sex together were the fixed effects that explained most of the variation in AW EE. Adding age, physical activity level (PAL) and having a disability, which were shown to have an influence on EE in previous studies, did not improve model fit. This suggests that the AW EE estimation algorithm mostly adjusts for differences in BM and sex, yet does not sufficiently adjust for increases in exercise intensity.

The finding of little variation in AW EE being explained by AW HR, indicates that exercise intensity is insufficiently considered by the AW EE estimation algorithm. It also clarifies why we found an increased underestimation at higher intensity in a previous study [15]. A possible explanation is that the AW EE estimation algorithm does not further adjust for fluctuations in EE with intensity but uses a fixed EE within the activity settings. The AW has two wheelchair-specific activity settings; the activity setting “outdoor push walking pace” (used in the current study across all stages) likely assumes that WCU exercise at a relatively stable low intensity, while the setting “outdoor push running pace” might be geared towards moderate- to high-intensity exercise. Altogether, our findings suggest that the inclusion of HR or another indicator of exercise intensity (e.g., accelerometry), should lead to improved AW EE estimates within each activity setting.

Interestingly, most of the variation in AW EE was explained by BM alone. In contrast, a lot less of the variation in criterion EE was explained by this factor, although the relationship between the BM and EE (as indicated by the regression slope in Fig. 3) was similar for the AW and criterion values. The apparent explanation is that the AW insufficiently adjusts for exercise intensity and thereby creates artificially low variation in AW EE leading to high R^2m values. On the contrary, for criterion EE, the lack of adjusting for intensity-related variation in EE when entering BM as the sole fixed effect, leads to low R^2m values. Furthermore, in the current study, we found that adding sex in addition to BM improves model fit to the AW EE and criterion EE data. There are two possible explanations: 1) There indeed are physiological differences in the HR-EE relationship between females and males, i.e., at a given HR, EE is higher for males than for females since they have a larger stroke volume [50], more hemoglobin in the blood [51], and more fat free mass especially in the upper body [37]. 2) Alternatively, this finding may be an artefact of the female participants performing wheelchair propulsion stages at a given incline at slower speed compared to their male counterparts. To further improve EE estimates during exercise, investigations are needed into how information on BM and sex (possibly together with other personal information) can be best combined to compensate for the lack of information on fat-free mass.

None of the additional fixed effects (i.e., age, PAL and having a disability or not) explained additional variation in the AW EE (i.e., 75% in AW Model 5 compared to 74% in AW Model 4, Fig. 1). In contrast, some additional variation was explained in the criterion EE by adding these fixed effects (36% in criterion Model 5 instead of 30% in criterion Model 4). This indicates that age, PAL and having a disability are factors that should probably be considered in the AW EE estimation algorithm, given that the findings for variation explained in criterion EE align with previous studies reporting that age, PAL and disability affect body composition and EE [38–40]. Of note, despite the AW EE estimation algorithm being tailor-made for WCU including a specific wheelchair activity setting, it does not seem to account for disability beyond adjusting for the generally lower EE during wheelchair propulsion compared to lower-body or whole-body exercise. This is likely due to the challenges and costs associated with collecting sufficient amounts of data from individuals with various types and levels of impairment, to further individualize the EE estimation algorithms. Creating several smaller datasets like the one collected in the current study and making them publicly available, may in the future contribute to a larger database for better individualization and more accessible technology for individuals of minority groups. This can result in more WCU having access to smartwatches that accurately estimate their energy expenditure. Consequently, this might help them to increase their understanding of a good balance between activity and nutrition, which can lead to better health.

Methodical limitations and future considerations

While the current study includes a sample size large enough to have sufficient power for our analyses, considerably more data will be needed to develop better individualized EE estimation algorithms that are tailored to individual WCU with various impairments.

Furthermore, and as described previously, we used the activity setting "outdoor push walking pace" during all stages. This may have led to an increased underestimation in AW EE as would have been the case when choosing the "outdoor push running pace" setting. We also performed the same analyses on the slowest speed at each incline, which correspond closest the activity setting "outdoor push walking pace" (see Supplementary Figures S3). The results of these additional analyses were similar in that AW HR explained little in AW EE, which further supports that an indicator of exercise intensity is currently lacking. In hindsight, we should have considered the selection of the activity settings more carefully, and either used "outdoor push running pace" for the higher intensity stages or compared the potential difference between the two settings on estimated AW EE.

In addition, participants wore the AW on their non-dominant hand. This choice was mainly made for practical purposes, since participants needed to start and stop the watch for each 4-min stage. The location of the smartwatches is a topic for debate, as the dominant arm is used more in daily life, thereby possibly obscuring activity measurements. However, watch location is unlikely to impact the results of the current study, since the activity during the stages was standardized wheelchair propulsion with relatively similar activity of both arms.

Lastly, all factors included in the present study explained, at best, 78% of the variation in the AW algorithm. We recognize that 100% explained variation for AW EE might be unattainable due to the true variation in measurements and the black-box algorithms relying on additional unknown factors (e.g., accelerometry). However, identifying and including potentially missing factors, most importantly a suitable intensity measure, should further improve EE estimation algorithms.

Conclusion

It remains unclear whether HR is used to adjust for intensity in the AW EE estimation algorithms, as it contributed little to the explained variation in AW EE. This finding is in contrast to criterion HR explaining a considerably higher amount of variation in criterion EE. It also indicates that HR or another indicator of intensity, such as accelerometry, should be used to improve the accuracy of estimated AW EE. In addition, BM as a sole fixed effect explained more than half of the variation in AW EE, suggesting that the AW algorithms adjust the EE estimate based on participants reported BM. Furthermore, the combination of fixed effects BM and sex explained most of the variation in the AW EE. This suggests that the AW algorithm accounts for some of the sex differences in EE — beyond the differences in BM. Adding body height, age, PAL or having a disability did not improve model fit to the AW EE data, but did improve model fit to the criterion EE data. Future investigations are needed to determine which intensity measure best explains the variation in EE in WCU and how to better individualize the EE estimation algorithms based on personal characteristics such as age, PAL, and also disability-related information.

Abbreviations

AW	Apple watch
BM	Body mass
EE	Energy expenditure
HR	Heart rate
ND	Non-disabled individuals
PAL	Physical activity level
R^2c	Conditional R^2
R^2m	Marginal R^2
WCU	Wheelchair users

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s44247-024-00101-z.

Supplementary Material 1.

Supplementary material 2.

Supplementary Material 3.

Acknowledgements
Testing was performed using the laboratory facilities and equipment of NeXt Move, Norwegian University of Science and Technology (NTNU), Faculty of Medicine and Health Sciences at NTNU and the Central Norway Regional Health Authority. We appreciate the continuous contributions from everyone involved in "The digital wheelchair" project, in addition to all the colleagues at the Centre for Elite Sports Research and lab assistants involved in the collection.

Authors' contributions
MLD and JKB analyzed and interpreted the extracted data. MLD wrote the manuscript drafts. RD, BB and MW and JKB critically read and revised the draft versions. The designated authors have participated sufficiently in the work to meet the requirements of co-authorship as specified in the authorship guidelines and have read and approved the final submitted manuscript.

Funding
Open access funding provided by Norwegian University of Science and Technology This study was financed by strategic funding of Faculty of Medicine and Health Science at the Norwegian University of Science and Technology (NTNU), an NTNU innovation scholarship and NTNU Discovery. The funding bodies had no role in the design of this study nor its execution, analyses, interpretation of the data, or decision to submit results.

Availability of data and materials
The datasets used during the current study are available from the corresponding author on reasonable request. The datasets will partially be made publicly available once the project this data collection is a part of is finished.

Declarations

Ethics approval and consent to participate
The data collection, processing and participant consent form were approved by the Norwegian Centre for Research Data (ID: 216680) and the data collection conducted in accordance with the Declaration of Helsinki. Written consent was obtained from all subjects before testing started. The consent form contained information on the study purpose, study design, potential risks and the right to withdraw from the study without providing justification.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Neuromedicine and Movement Science, Centre for Elite Sports Research, Norwegian University of Science and Technology, Trondheim, Norway. [2]Department of Engineering Cybernetics, Norwegian University of Science and Technology, Trondheim, Norway. [3]Department of Neuromedicine and Movement Science, Norwegian University of Science and Technology, Trondheim, Norway. [4]Department of Rehabilitation, Clinic for Spinal Cord Injuries, St. Olavs University Hospital, Trondheim, Norway. [5]Department of Research, Sunnaas Rehabilitation Hospital, Oslo, Norway. [6]Department of Rehabilitation Science and Health Technology, Faculty of Health Sciences, Oslo Metropolitan University, Oslo, Norway.

Received: 23 February 2024 Accepted: 15 May 2024
Published online: 26 July 2024

References

1. Chandrasekaran R, et al. Patterns of use and key predictors for the use of wearable health care devices by US adults: insights from a national survey. J Med Internet Res. 2020;22(10):e22443.
2. Massoomi MR, et al. Increasing and evolving role of smart devices in modern medicine. Eur Cardiol. 2019;14(3):181–6.
3. Heaney J. Energy: Expenditure, Intake, Lack of. In: Gellman MD, Turner JR, editors. Encyclopedia of Behavioral Medicine. Springer, New York: New York; 2013. p. 699–700.
4. Boutari C, et al. A 2022 update on the epidemiology of obesity and a call to action: as its twin COVID-19 pandemic appears to be receding, the obesity and dysmetabolism pandemic continues to rage on. Metabolism. 2022;133:155217.
5. Weil E, et al. Obesity among adults with disabling conditions. JAMA. 2002;288(10):1265–8.
6. Farkas GJ, et al. Energy expenditure following spinal cord injury: a delicate balance. Top Spinal Cord Inj Rehabil. 2021;27(1):92–9.
7. Shea JR, et al. Energy expenditure as a function of activity level after spinal cord injury: the need for tetraplegia-specific energy balance guidelines. Front Physiol. 2018;9:1286.
8. Johnson RK, et al. Total energy expenditure in adults with cerebral palsy as assessed by doubly labeled water. J Am Diet Assoc. 1997;97(9):966–70.
9. Buchholz AC, et al. Greater daily leisure time physical activity is associated with lower chronic disease risk in adults with spinal cord injury. Appl Physiol Nutr Metab. 2009;34(4):640–7.
10. Buchholz AC, et al. Physical activity levels are low in free-living adults with chronic paraplegia. Obes Res. 2003;11(4):563–70.
11. Nightingale TE, et al. Measurement of physical activity and energy expenditure in wheelchair users: methods, considerations and future directions. Sports Med Open. 2017;3(1):10.
12. Fuller D, 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(9):e18694.
13. Germini F, et al. Accuracy and acceptability of wrist-wearable activity-tracking devices: systematic review of the literature. J Med Internet Res. 2022;24(1):e30791.
14. Moreno D, et al. Validity of caloric expenditure measured from a wheelchair user smartwatch. Int J Sports Med. 2020;41(8):505–11.
15. Danielsson ML, et al. Accuracy of the Apple Watch Series 4 and Fitbit Versa for assessing energy expenditure and heart rate of wheelchair users during treadmill wheelchair propulsion: A cross sectional study. JMIR Form Res. 2024;0:e0.
16. Karvonen J, et al. Heart rate and exercise intensity during sports activities. Practical application. Sports Med. 1988;5(5):303–11.
17. Ludwig M, et al. Measurement, prediction, and control of individual heart rate responses to exercise—basics and options for wearable devices. Front Physiol. 2018;9:778.
18. Shcherbina A, et al. Accuracy in wrist-worn, sensor-based measurements of heart rate and energy expenditure in a diverse cohort. J Pers Med. 2017;7(2):3.
19. Falter M, et al. Accuracy of apple watch measurements for heart rate and energy expenditure in patients with cardiovascular disease: cross-sectional study. JMIR Mhealth Uhealth. 2019;7(3):e11889.
20. Düking P, et al. Wrist-worn wearables for monitoring heart rate and energy expenditure while sitting or performing light-to-vigorous physical activity: validation study. JMIR Mhealth Uhealth. 2020;8(5):e16716.
21. Hajj-Boutros G, et al. Wrist-worn devices for the measurement of heart rate and energy expenditure: a validation study for the apple watch 6, polar vantage v and fitbit sense. Eur J Sport Sci. 2023;23(2):165–77.
22. Uphill A, et al. Validity of Apple Watch, Garmin Forerunner® 935 and GENEActiv for estimating energy expenditure during close quarter battle training in Special Forces soldiers. Eur J Sport Sci. 2024;24:614–22.
23. Li R, et al. A critical evaluation of heart rate monitoring to assess energy expenditure in individuals. Am J Clin Nutr. 1993;58(5):602–7.
24. Haskell WL, et al. Simultaneous measurement of heart rate and body motion to quantitate physical activity. Med Sci Sports Exerc. 1993;25(1):109–15.
25. Leonard WR. Measuring human energy expenditure: what have we learned from the flex-heart rate method? Am J Hum Biol. 2003;15(4):479–89.
26. Kortelainen L, et al. A nonlinear mixed model approach to predict energy expenditure from heart rate. Physiol Meas. 2021;42(3):03500.

27. Parak J, et al. Estimating heart rate, energy expenditure, and physical performance with a wrist photoplethysmographic device during running. JMIR Mhealth Uhealth. 2017;5(7):e7437.
28. Gallagher D, et al. Metabolically active component of fat-free body mass: Influences of age, adiposity, and gender. Metabolism. 1996;45(8):992–7.
29. Miller A Jr, et al. Lean body mass as a metabolic reference standard. J Appl Physiol. 1953;5(7):311–6.
30. Weyer C, et al. Determinants of energy expenditure and fuel utilization in man: effects of body composition, age, sex, ethnicity and glucose tolerance in 916 subjects. Int J Obes. 1999;23(7):715–22.
31. Müller MJ, et al. Effect of constitution on mass of individual organs and their association with metabolic rate in humans—a detailed view on allometric scaling. PLoS One. 2011;6(7):e22732.
32. Speakman JR, et al. Associations between energy demands, physical activity, and body composition in adult humans between 18 and 96 y of age. Am J Clin Nutr. 2010;92(4):826–34.
33. Wang Z, et al. Specific metabolic rates of major organs and tissues across adulthood: evaluation by mechanistic model of resting energy expenditure. Am J Clin Nutr. 2010;92(6):1369–77.
34. Ravussin E. Twenty-four-hour energy expenditure and resting metabolic rate in obese, moderately obese, and control subjects. Am J Clin Nutr. 1982;35(3):566–73.
35. Kee AL, et al. Resting energy expenditure of morbidly obese patients using indirect calorimetry: a systematic review. Obes Rev. 2012;13(9):753–65.
36. Hackney AC. Chapter 5 - Energy Expenditure at Rest and During Various Types of Physical Activity. In: Hackney AC, editor. Exercise. Sport, and Bioanalytical Chemistry: Elsevier; 2016. p. 43–52.
37. Bredella MA. Sex differences in body composition. Adv Exp Med Biol. 2017;1043:9–27.
38. Pontzer H, et al. Daily energy expenditure through the human life course. Science. 2021;373(6556):808–12.
39. Poehlman ET. A review: exercise and its influence on resting energy metabolism in man. Med Sci Sports Exerc. 1989;21(5):515–25.
40. da Innocencio Silva Gomes A, et al. Basal and resting metabolic rates of physically disabled adult subjects: a systematic review of controlled cross-sectional studies. Ann Nutr Metab. 2014;65(4):243–52.
41. Borg GA. Psychophysical bases of perceived exertion. Med Sci Sports Exerc. 1982;14(5):377–81.
42. Van Hooren B, et al. Accuracy of respiratory gas variables, substrate, and energy use from 15 CPET systems during simulated and human exercise. Scand J Med Sci Sports. 2024;34(1):e14490.
43. Polar Electro Oy. Polar H10 Heart Rate Sensor System. 2019. Available from: https://www.polar.com/sites/default/files/static/science/white-papers/polar-h10-heart-rate-sensor-white-paper.pdf.
44. Benning N, et al. Measurement performance of activity measurements with newer generation of apple watch in wheelchair users with spinal cord injury. Methods Inf Med. 2021;60:e103–10.
45. Weir JB. New methods for calculating metabolic rate with special reference to protein metabolism. J Physiol. 1949;109(1–2):1–9.
46. IPAQ Research Committee. Guidelines for data processing and analysis of the International Physical Activity Questionnaire (IPAQ). 2005. Available from: https://www.researchgate.net/file.PostFileLoader.html?id=5641f4c36143250eac8b45b7&assetKey=AS%3A294237418606593%401447163075131.
47. Saebu M, et al. Factors associated with physical activity among young adults with a disability. Scand J Med Sci Sports. 2011;21(5):730–8.
48. R Core Team. R: A language and environment for statistical computing Vienna, Austria: R foundation for statistical computing; 2013.
49. Parker RA, et al. Using multiple agreement methods for continuous repeated measures data: a tutorial for practitioners. BMC Med Res Methodol. 2020;20(1):154.
50. St Pierre SR, et al. Sex Matters: a comprehensive comparison of female and male hearts. Front Physiol. 2022;13:831179.
51. Murphy WG. The sex difference in haemoglobin levels in adults - mechanisms, causes, and consequences. Blood Rev. 2014;28(2):41–7.