Workers’ physical activity data contribute to estimating maximal oxygen consumption: a questionnaire study to concurrently assess workers’ sedentary behavior and cardiorespiratory fitness

Tomoaki Matsuo1,2*, Rina So1,2 and Masaya Takahashi2

Abstract

Background: Sedentary behavior (SB) and cardiorespiratory fitness (CRF) are important issues in occupational health. Developing a questionnaire to concurrently assess workers’ SB and CRF could fundamentally improve epidemiological research. The Worker’s Living Activity-time Questionnaire (WLAQ) was developed previously to assess workers’ sitting time. WLAQ can be modified to evaluate workers’ CRF if additional physical activity (PA) data such as PA frequency, duration, and intensity are collected.

Methods: A total of 198 working adults (93 women and 105 men; age, 30–60 years) completed anthropometric measurements, a treadmill exercise test for measuring maximal oxygen consumption (VO2max), and modified WLAQ (m-WLAQ, which included questions about PA data additional to the original questions). Multiple regression analyses were performed to develop prediction equations for VO2max. The generated models were cross-validated using the predicted residual error sum of squares method. Among the participants, the data of 97 participants who completed m-WLAQ twice after a 1-week interval were used to calculate intraclass correlation coefficient (ICC) for the test–retest reliability analyses.

Results: Age ($r = -0.29$), sex ($r = 0.48$), body mass index (BMI, $r = -0.20$), total sitting time ($r = -0.15$), and PA score (total points for PA data, $r = 0.47$) were significantly correlated with VO2max. The models that included age, sex, and BMI accounted for 43% of the variance in measured VO2max (standard error of the estimate (SEE) = 5.04 ml·kg$^{-1}$·min$^{-1}$). These percentages increased to 59% when the PA score was included in the models (SEE = 4.29 ml·kg$^{-1}$·min$^{-1}$). Cross-validation analyses demonstrated good stability of the VO2max prediction models, while systematic underestimation and overestimation of VO2max were observed in individuals with high and low fitness, respectively. The ICC of the PA score was 0.87 (0.82–0.91), indicating excellent reliability.

Conclusions: The PA score obtained using m-WLAQ, rather than sitting time, correlated well with measured VO2max. The equation model that included the PA score as well as age, sex, and BMI had a favorable validity for estimating VO2max. Thus, m-WLAQ can be a useful questionnaire to concurrently assess workers’ SB and CRF, which makes it a reasonable resource for future epidemiological surveys on occupational health.

Keywords: Occupational health, Physical fitness, Reliability, Sitting time, Validity
Background
Many studies have shown that excessive sedentary behavior (SB) increases disease risk [1, 2]. In this mechanismed society, workers are particularly likely to be placed in sedentary situations in the workplace [3, 4]. Therefore, workers’ SB is an important risk factor for occupational health. Similarly, cardiorespiratory fitness (CRF) is a conventional health issue because many studies have shown that low CRF level is strongly associated with increased disease and mortality risks [5]. Recent studies have indicated that midlife CRF plays a role in health-related incidents in later life such as the development of severe diseases [6], increased healthcare costs [7], and decreased longevity [8].

Thus, from the perspective of preventative medicine, both SB and CRF in working adults are key factors in occupational health. However, the relationship between workers’ SB and CRF as well as their interaction effects on disease risk has rarely been investigated in epidemiological surveys. One crucial reason would be inherent in the assessment methodology. The gold standard methods for SB and CRF assessment, such as thigh-worn inclinometer including the activPAL [9] for SB and maximal oxygen consumption (VO\textsubscript{2max}) for CRF, require relatively high cost and considerable time; thus, they have a practical disadvantage for a population-based assessment. In epidemiological surveys, subjective measures such as questionnaires remain useful because they are more cost-effective and present a lower participant burden [10], although the key limitation of the questionnaire method is poor validity. To the best of our knowledge, validated questionnaires to simultaneously assess workers’ SB and CRF have not been proposed.

The Worker’s Living Activity-time Questionnaire (WLAQ) was primarily developed to assess workers’ sitting times in our previous studies [11, 12]. WLAQ can be used to measure a worker’s time spent sitting within four typical domains of a worker’s life: (a) working time, (b) commuting time, (c) nonworking time on a workday, and (d) time on a non-workday. Matsuo et al. [12] evaluated WLAQ and demonstrated favorable test–retest reliability values and criterion (vs. activPAL) validity values for the four sitting times.

Age, sex, and body fat-related values have often been used in VO\textsubscript{2max} estimation models [13–15]. Given that measurement accessibilities differ among different body fat-related values such as body mass index (BMI), waist girth (WG), and %fat, these previous studies [13–15] investigated the predictive power for each body fat item. Furthermore, previous studies [14–18] have shown that physical activity (PA) data, such as frequency, duration, and intensity, contribute to the estimation of VO\textsubscript{2max}. However, the original WLAQ can help assess workers’ sitting times but not other PA data. Thus, WLAQ can be modified for evaluating workers’ CRF if additional PA data are collected, and it can be used along with age, sex, and body fat-related values to develop an equation model for VO\textsubscript{2max} estimation.

Therefore, the purposes of this study were 1) to investigate associations between measured VO\textsubscript{2max} and sitting times and other PA data collected using the modified WLAQ (m-WLAQ); 2) to investigate criterion validity of a developed equation model for estimating VO\textsubscript{2max}; 3) to compare the accuracy of equation models that include BMI, WG, and %fat; and 4) to investigate the test–retest reliabilities of values derived using m-WLAQ.

Methods
Participants
The inclusion criteria were as follows: 1) aged 30–60 years, 2) living in the Tokyo area (Tokyo, Saitama, Chiba, and Kanagawa Prefectures) of Japan, 3) part-time or full-time worker for at least 3 days a week, and 4) no medical conditions that could affect VO\textsubscript{2max} testing. Participants were recruited through a website advertisement. In total, 202 working adults (97 women and 105 men) participated in this study. The participants visited our laboratory to complete anthropometric measurements, a treadmill exercise test, and m-WLAQ. We excluded 4 participants due to insufficient data for the analyses. Consequently, 198 participants (93 women and 105 men) were included in the validity analysis. Furthermore, among the included participants, 97 participants (42 women and 55 men) visited our laboratory for a second time, with an interval of 1 week between the visits. At the second visit, they completed m-WLAQ again for test–retest reliability analyses.

This study was conducted in accordance with the guidelines proposed in the Declaration of Helsinki. The Ethical Committee of the National Institute of Occupational Safety and Health, Japan reviewed and approved the study protocol (ID H2810). Before obtaining written informed consent, the aims and design of this study were explained to each participant.

Measures
Anthropometric measurements
For each participant, height (to the nearest 0.1 cm) was measured once using a wall-mounted stadiometer (YG-200, Yagami, Nagoya, Japan). Body weight (to the nearest 0.1 kg) and %fat with bioelectrical impedance analysis were assessed using a body composition analyzer (InBody-3.2; Biospace, Seoul, Korea). WG (to the nearest 0.1 cm) was measured twice at the level of the umbilicus in the standing position by a skilled member of the research staff. BMI was calculated as weight (in kilograms) divided by the square of height (in meters).
**m-WLAQ**

m-WLAQ was used to evaluate participants’ sitting times during working time, commuting time, and leisure time on a workday and non-workday. A previous study [12] using the original WLAQ demonstrated favorable test–retest reliability values [intraclass correlation coefficient (ICC) = 0.72–0.98] and criterion (activPAL) validity values (Spearman’s ρ = 0.40–0.82) for the four sitting times. The original WLAQ was modified to add several questions for collecting PA data (frequency, duration, and intensity) for developing m-WLAQ. The PA score was calculated as the sum of the points scored for PA data. For its calculation, PA intensity was weighted more heavily than PA duration and frequency (0 and intensity) for developing m-WLAQ. The PA score showed no significant sex differences between groups. Chi-squared tests were used to evaluate the relationship between the measured VO\(_{2\text{max}}\) and other variables. Participants who were older, women, and those with higher body fat-related values had lower VO\(_{2\text{max}}\). Significant negative correlations were observed between VO\(_{2\text{max}}\) and sitting time during leisure time on workday (\(r = -0.27, P < 0.01\)) and total sitting time (\(r = -0.15, P = 0.03\), although sitting times during commuting time, working time, and non-workday were not significantly correlated with VO\(_{2\text{max}}\). Significant positive correlations were observed between PA data (frequency, duration, and intensity) and measured VO\(_{2\text{max}}\) with the largest association (\(r = 0.47, P < 0.01\)) revealed for the PA score.

**Maximal oxygen consumption**

The participants underwent an electrocardiogram-monitoring, exhaustion-limited, graded exercise test on a treadmill (EXCITE RUN, Technogym, Cesena, Italy) using the Bruce protocol to determine VO\(_{2\text{max}}\). During the test, an open-circuit computerized indirect calorimeter (AE-310S, Minato Medical Science, Osaka, Japan) was used to measure ventilation and expired gases. The gas analyzer was calibrated before each trial. Heart rate (HR) was monitored using an electrocardiogram monitor (LifeScope, NIHON KOHDEN, Tokyo, Japan), and a rating of perceived exertion (RPE), using the 6–20 Borg RPE scale, was recorded during the exercise test. The highest 30-s average VO\(_{2}\) value was defined as the VO\(_{2\text{max}}\) value. The exercise test was considered to achieve VO\(_{2\text{max}}\) when three of the following four criteria were satisfied: 1) respiratory exchange ratio >1.10, 2) achievement of maximum HR within 10 bpm of the age-predicted maximal (220−age), 3) RPE >17, and 4) VO\(_{2}\) plateau despite further increases in workload [20, 21].

**Data analysis**

Unpaired Student’s t-tests were performed to evaluate differences between groups. Chi-squared tests were used to analyze categorical variables. Pearson’s correlation coefficients were calculated to evaluate the relationship between the measured VO\(_{2\text{max}}\) and other measurement values. Multiple linear regression analysis was used to develop prediction equations for VO\(_{2\text{max}}\). Changes in the squared multiple correlation coefficient (\(R^2\)) and the standard error of the estimate (SEE) were used to assess the incremental gain in variance explained by the different variables added to the model. In the course of previous studies [13–15], \(R^2\) and SEE were compared among some types of body fat evaluation models, i.e., BMI, WG, and %fat models, to investigate the influence of differences in methodology, because subjects’ body fat is assessed in several ways in epidemiological surveys. The generated models were cross-validated using the predicted residual error sum of squares (PRESS) statistical method [22]. This method calculates the error in prediction for each case when only that case is excluded from generating the estimation model and applying the model to the excluded case. PRESS adjusted \(R^2 (R^2_p)\) is calculated as 1 – (PRESS statistic/SS total), whereas PRESS SEE (SEE\(_p\)) is calculated using the following equation: SEE\(_p\) = \(\sqrt{\text{PRESS statistic}/N}\). The generated models were further validated by comparing the constant errors (CEs) among the subgroups of sex, age, and VO\(_{2\text{max}}\) levels. CE is calculated as the mean difference between measured VO\(_{2\text{max}}\) and predicted VO\(_{2\text{max}}\). The test–retest reliability was examined using ICC and 95% confidence interval (95% CI) with an ICC of <0.40 indicating poor repeatability, 0.40−0.75 indicating fair-to-good repeatability, and >0.75 indicating excellent repeatability [23]. P-value of <0.05 was considered statistically significant. All analyses were conducted using SAS, version 9.4 (SAS Institute Japan, Tokyo, Japan).

**Results**

Table 1 summarizes the demographic characteristics of the participants. We observed higher anthropometric measurements in men than in women. Although sitting times during worktime and on a non-workday were significantly longer in men than in women, sitting time during leisure time on workday was significantly longer in women than in men. The percentage of clerical jobs was high in both sexes, but the percentages of sales and marketing and engineer/researcher were higher in men.

Table 2 shows the questions included in m-WLAQ, answer options (and their assigned points), and the results of the study participants. Although the answer trends were different between women and men for Q8 and Q10, the PA score showed no significant sex difference.

Table 3 shows Pearson’s correlation coefficients between the measured VO\(_{2\text{max}}\) and other variables. Participants who were older, women, and those with higher body fat-related values had lower VO\(_{2\text{max}}\). Significant negative correlations were observed between VO\(_{2\text{max}}\) and sitting time during leisure time on workday (\(r = -0.27, P < 0.01\)) and total sitting time (\(r = -0.15, P = 0.03\), although sitting times during commuting time, working time, and non-workday were not significantly correlated with VO\(_{2\text{max}}\). Significant positive correlations were observed between PA data (frequency, duration, and intensity) and measured VO\(_{2\text{max}}\) with the largest association (\(r = 0.47, P < 0.01\)) revealed for the PA score.
The variables significantly related to the measured VO2max were included as predictors in multiple regression models (Table 4). The model that included age, sex, and BMI accounted for 43% of the variance in measured VO2max (adjusted $R^2 = 0.43$, $SEE = 5.04$ ml·kg$^{-1}$·min$^{-1}$ or 13.1%). When the PA score obtained using m-WLAQ was added to the model that included age, sex, and BMI, the adjusted $R^2$ significantly improved by 16% (adjusted $R^2 = 0.59$, $SEE = 4.29$ ml·kg$^{-1}$·min$^{-1}$ or 11.2%). Similarly, regarding the models using WG or %fat instead of BMI, addition of the PA score improved the variance of the models by 12% or 11%, respectively. The cross-validation results from the PRESS method are also shown in Table 4. The decrease in $R^2$ (approximately 0.01) and increase in $SEE$ value (approximately 0.05 ml·kg$^{-1}$·min$^{-1}$) were small for all three models.

Table 5 presents the results of other cross-validation analyses using CE values. The absolute CE values for the sex and age subgroups were < 1.00 for all three models. Regarding the VO2max subgroups, CE values were negatively high (overestimation) for the low-fitness subgroup and positively high (underestimation) for the high-fitness subgroup, whereas lower CE values were observed for the mid-fitness subgroup in all three models. The test–retest reliabilities were examined using the data of the 97 participants who completed m-WLAQ twice. Table 6 shows the ICC of sitting times and PA-related values obtained using m-WLAQ. The ICC values for sitting times were fair to good (commuting time and non-workday) or excellent (working time and leisure time on workday and total sitting time). All ICC values of PA-related values were excellent except those for exercise duration on non-workday and exercise intensity during leisure time on workday and non-workday, whose values were fair to good. The ICC value of the PA score was 0.87, indicating excellent reliability.

The relationships between measured VO2max and estimated VO2max are shown in Fig. 1. Estimated VO2max correlated well with measured VO2max for all three body fat models; Fig. 1 also shows overestimations in participants with low fitness and underestimations in participants with high fitness for all three models.

**Discussion**

The study showed that 1) the PA score obtained using m-WLAQ, rather than sitting times, was associated with measured VO2max; 2) the equation models that included age, sex, body fat-related values, and PA score obtained using m-WLAQ had favorable validity for estimating VO2max; 3) no appreciable difference was observed in estimated VO2max among the three models with regard to BMI, WG, and %fat; and 4) favorable reliability values were shown for sitting times and the PA score obtained using m-WLAQ.
| Questions                                                                 | Answer options                                                                 | Women  | Men     | P   |
|--------------------------------------------------------------------------|-------------------------------------------------------------------------------|--------|---------|-----|
| Q6 How much breathing-inducing (heart rate increasing) tasks do you perform on an average day during your working hours? | 1. none/almost none (0)                                                       | 81 (87.1) | 76 (72.4) | 0.07 |
|                                                                            | 2. rarely (3)                                                                | 7 (7.5) | 19 (18.1) |     |
|                                                                            | 3. sometimes (5)                                                             | 5 (5.4) | 9 (8.6)   |     |
|                                                                            | 4. often (10)                                                                | 0 (0.0) | 1 (1.0)   |     |
| Q8 In your leisure time on a workday, how much intentional physical activity do you engage in? | 1. none/almost none (0)                                                       | 45 (48.4) | 54 (51.4) | 0.05 |
|                                                                            | 2. 1–3 days a month (1)                                                       | 6 (6.5) | 18 (17.1) |     |
|                                                                            | 3. 1 or 2 days a week (2)                                                     | 27 (29.0) | 18 (17.1) |     |
|                                                                            | 4. ≥3 days a week (3)                                                         | 15 (16.1) | 15 (14.3) |     |
| Q9 If options 2–4 were selected from Q8, please provide the average exercise time per day | 1. < 15 min (1)                                                              | 1 (1.1) | 2 (1.9)   | 0.10 |
|                                                                            | 2. 15–30 min (2)                                                             | 12 (12.9) | 6 (5.7)   |     |
|                                                                            | 3. 31–60 min (3)                                                             | 25 (26.9) | 22 (21.0) |     |
|                                                                            | 4. > 60 min (4)                                                              | 10 (10.8) | 21 (20.0) |     |
|                                                                            | N/A (Q8 = 1)                                                                 | 45 (48.4) | 54 (51.4) |     |
| Q10 If options 2–4 were selected from Q8, please tell us the approximate intensity of the exercise per session | 1. no sweating or panting (0)                                                 | 15 (31.3) | 4 (3.8)   | < 0.01 |
|                                                                            | 2. sweating and panting (3)                                                   | 29 (60.4) | 32 (60.5) |     |
|                                                                            | 3. strained breathing (5)                                                     | 4 (8.3)  | 14 (13.3) |     |
|                                                                            | 4. to the point of exhaustion (10)                                            | 0 (0.0)  | 1 (1.0)   |     |
|                                                                            | N/A (Q8 = 1)                                                                 | 45 (48.4) | 54 (51.4) |     |
| Q13 On holidays, how much intentional physical activity do you engage in?  | 1. none/almost none (0)                                                       | 39 (41.9) | 36 (34.3) | 0.69 |
|                                                                            | 2. 1–2 days a month (1)                                                       | 17 (18.3) | 23 (21.9) |     |
|                                                                            | 3. once a week (2)                                                            | 24 (25.8) | 32 (30.5) |     |
|                                                                            | 4. ≥2 days a week (3)                                                         | 13 (14.0) | 14 (13.3) |     |
| Q14 If options 2–4 were selected from Q13, please provide the average exercise time per day | 1. < 15 min (1)                                                              | 2 (2.2)  | 2 (1.9)   | 0.61 |
|                                                                            | 2. 15–30 min (2)                                                             | 4 (4.3)  | 8 (7.6)   |     |
|                                                                            | 3. 31–60 min (3)                                                             | 27 (29.0) | 27 (25.7) |     |
|                                                                            | 4. > 60 min (4)                                                              | 21 (22.6) | 32 (30.5) |     |
|                                                                            | N/A (Q13 = 1)                                                                 | 39 (41.9) | 36 (34.3) |     |
| Q15 If options 2–4 were selected from Q13, please tell us the approximate intensity of the exercise per session | 1. no sweating or panting (0)                                                 | 10 (10.8) | 10 (9.5) | 0.20 |
|                                                                            | 2. sweating and panting (3)                                                   | 38 (40.9) | 41 (39.0) |     |
Consistent with the findings of a previous study [24], significant negative correlations were observed between sitting times and measured VO$_{2\text{max}}$ (Table 3). However, sitting times were not accepted as an effective explanatory variable for estimating VO$_{2\text{max}}$ in our regression analyses. In contrast, questionnaire-based PA data, such as frequency, duration, and intensity, were significantly correlated with measured VO$_{2\text{max}}$ (Table 3), and the regression model identified the PA score to be the principal explanatory value for the equation models. The PA score was calculated for precise VO$_{2\text{max}}$ prediction in reference to the HUNT study by Nes et al. [17] and a previous exercise intervention study [19]. Nes et al. [17] used some question items regarding PA frequency, duration, and intensity for estimating VO$_{2\text{max}}$ and relative weightings of different responses were set on the basis of their relation to VO$_{2\text{max}}$. In their estimation, PA intensity was weighted more heavily than PA duration and frequency on the PA score. Further, an exercise intervention study [19] emphasized the primacy of PA intensity rather than PA duration and volume in improving VO$_{2\text{max}}$. We followed these studies to develop the PA score, i.e., the questions regarding intensity, such as Q6, Q10, and Q15, were weighted more heavily than other questions (Table 2). The PA score was strongly correlated with VO$_{2\text{max}}$ (Table 3) and functioned well for estimating VO$_{2\text{max}}$ (Table 4). The results of previous studies and the present study suggest that PA intensity can have a potential role in estimating VO$_{2\text{max}}$.

Age, sex, and body fat-related values were significantly correlated with measured VO$_{2\text{max}}$ (Table 3), and these three factors accounted for 43–51% of the variance in measured VO$_{2\text{max}}$ (Table 4). These percentages increased by 11–16% following to addition of the PA score obtained using the questionnaire in the model (Table 4). Jackson et al. [14] suggested questionnaire-based VO$_{2\text{max}}$ prediction models including age, sex, body fat-related values, and the PA score obtained using the questionnaire, demonstrating SEE of 5.35–5.70 ml·kg$^{-1}$·min$^{-1}$. Similarly, Wier et al. [15] suggested questionnaire-based VO$_{2\text{max}}$ prediction models including age, sex, body fat-related values, and the PA score obtained using the questionnaire, demonstrating SEE of 4.72–4.90 ml·kg$^{-1}$·min$^{-1}$. Furthermore, Malek et al. [25] developed a VO$_{2\text{max}}$ prediction equation including age, body weight, height, and questionnaire-based exercise values, which

### Table 2 Questions of m-WLAQ for CRF estimation and the results of this study participants (n = 198) (Continued)

| Questions                                      | Answer options                  | Women (n = 93) | Men (n = 105) | P  |
|------------------------------------------------|---------------------------------|---------------|---------------|----|
| 3. strained breathing (5)                      |                                 | 6 (6.5)       | 17 (16.2)     |    |
| 4. to the point of exhaustion (10)             |                                 | 0 (0.0)       | 1 (1.0)       |    |
| N/A (Q13 = 1)                                  |                                 | 39 (41.9)     | 36 (34.3)     |    |

PA score (total points of Q6, Q8, Q9, Q10, Q13, Q14, and Q15/0–44 points) 8.8 ± 7.7 10.7 ± 8.6 0.11

Values are presented as n (%) or mean ± standard deviation. *If the respondent selected option #1 on Q8, the scores on Q9 and Q10 should be zero. Similarly, if the respondent selected option #1 on Q13, the scores on Q14 and Q15 should be zero. CRF cardiorespiratory fitness, m-WLAQ modified Worker’s Living Activity-time Questionnaire. P values show the results of group difference analyses

### Table 3 Pearson’s correlation coefficients between measured VO$_{2\text{max}}$ and other values (n = 198)

| Age (years) | Sex women:0, men:1 | BMI | WG (cm) | %fat | Sitting time (min·d$^{-1}$) | PA score (points) |
|-------------|---------------------|-----|---------|------|-----------------------------|-------------------|
| r           | 0.48                | –0.20| –0.24   | –0.64| –0.11                       | 0.14              |
| P           | 0.01                | < 0.01| < 0.01  | < 0.01| 0.14                        | 0.01              |
| Working time|                    |      |         |      |                             |                   |
| PA intensity|                    |      |         |      |                             |                   |
| Exercise frequency |                    |      |         |      |                             |                   |
| Exercise duration |                   |      |         |      |                             |                   |
| Exercise intensity |                 |      |         |      |                             |                   |
| r           | 0.29                | 0.35 | 0.42    | 0.31 | 0.36                        | 0.47              |
| P           | 0.01                | 0.01 | < 0.01  | < 0.01| < 0.01                      | < 0.01            |

BMI body mass index, PA physical activity, VO$_{2\text{max}}$ maximal oxygen consumption, WG waist girth. Correlation coefficient (r) and its P value are displayed
had an $SEE$ value of 4.12 ml·kg$^{-1}$·min$^{-1}$. The present study showed results similar to these previous studies, with prediction model $SEE$s of 4.13–4.29 ml·kg$^{-1}$·min$^{-1}$. These $SEE$ values can be replaced with %$SEE$ values of 10.8–11.2%. Other types of $VO_2$max estimation studies reported %$SEE$ values of 11.4% in the 20-m shuttle run test study [26] and 10–15% in wearable device studies [13, 27, 28]. The $SEE$ values in the present study seem to be favorable when compared with those calculated in other $VO_2$max prediction studies.

Regarding the method to validate a regression equation, although the data-splitting method is well known, in which the three models using WG ($SEE$ value of 4.80 ml·kg$^{-1}$·min$^{-1}$), %fat ($SEE$ value of 4.72 ml·kg$^{-1}$·min$^{-1}$), or BMI ($SEE$ value of 4.90 ml·kg$^{-1}$·min$^{-1}$). The present study not only been used in studies with a large sample size such as in that of Matthews et al. (799 participants) [29] and Wier et al. (2801 participants) [15] but also in studies with a small sample size such as those of Malek et al. (115 participants) [25] and Cao et al. (148 participants) [13]. The PRESS method appeared to be appropriate for the present study on 198 participants.

Jackson et al. [14] recommended questionnaire-based $VO_2$max prediction models including age, sex, the PA score obtained using the questionnaire, and body fat-related values such as %fat (skinfold method) and BMI, and they demonstrated $SEE$ values of 5.35 ml·kg$^{-1}$·min$^{-1}$ for the %fat model and 5.70 ml·kg$^{-1}$·min$^{-1}$ for the BMI model. Wier et al. [15] also recommended questionnaire-based $VO_2$max prediction models including age, sex, the PA score obtained using the questionnaire, and body fat-related values such as %fat (skinfold method), WG, and BMI and they showed no considerable differences in accuracy among the three models using WG ($SEE$ value of 4.80 ml·kg$^{-1}$·min$^{-1}$), %fat ($SEE$ value of 4.72 ml·kg$^{-1}$·min$^{-1}$), or BMI ($SEE$ value of 4.90 ml·kg$^{-1}$·min$^{-1}$). The present study

### Table 4: Regression coefficients for prediction of $VO_2$max using selected independent values ($n = 198$)

| Intercept | Independent variables | Adjusted $R^2$ | SEE (ml·kg$^{-1}$·min$^{-1}$) | $R^p$ | $SEEp$ (ml·kg$^{-1}$·min$^{-1}$) |
|-----------|-----------------------|---------------|---------------------------------|-------|--------------------------------|
| BMI model | Age yrs | 63.90 | 0.25 | 8.02 | -0.79 | -0.33 | 0.33 | 0.43 | 5.04 | 1.27 |
| WG model | Sex women:0 men:1 | 59.96 | 0.25 | 7.39 | -0.79 | -0.33 | 0.33 | 0.59 | 4.29 | 1.12 |
| %fat model | %fat | 70.39 | 0.21 | 8.82 | -0.29 | -0.33 | 0.29 | 0.61 | 4.67 | 1.22 |
| | PA score | 64.70 | 0.29 | 2.95 | -0.54 | -0.48 | 0.28 | 0.62 | 4.13 | 1.08 |
| | | 57.50 | -0.24 | 2.83 | -- | -- | -- | -- | 4.03 | -- |

* Each model includes age, sex, and PA score in addition to the titled values (BMI, WG, and %fat) as independent variables

### Table 5: Constant error for subgroups of sex, age, and measured $VO_2$max ($n = 198$)

| Sex | N (%) | Measured $VO_2$max | BMI model* | WG model* | %fat model* |
|-----|-------|---------------------|------------|----------|------------|
|     |       | Predicted $VO_2$max| CE         | Predicted $VO_2$max| CE         | Predicted $VO_2$max| CE         |
| Women | 93 (47.0) | 35.0 ± 5.5 | 35.2 ± 4.0 | -0.17 ± 4.22 | 35.5 ± 4.1 | -0.48 ± 4.17 | 34.9 ± 4.3 | 0.16 ± 3.92 |
| Men | 105 (53.0) | 41.4 ± 6.2 | 41.6 ± 4.2 | -0.18 ± 4.29 | 41.9 ± 4.2 | -0.51 ± 4.11 | 41.3 ± 4.2 | 0.13 ± 4.24 |
| Age |       |           |           |           |           |           |           |           |
| 30–39 | 29 (14.7) | 41.4 ± 6.0 | 42.4 ± 4.1 | -0.91 ± 3.75 | 42.4 ± 4.5 | -0.98 ± 3.59 | 41.8 ± 3.9 | -0.39 ± 3.76 |
| 40–49 | 88 (44.4) | 39.1 ± 7.3 | 38.8 ± 5.4 | 0.32 ± 4.52 | 39.3 ± 5.4 | -0.12 ± 4.49 | 38.7 ± 5.6 | 0.40 ± 4.35 |
| 50–59 | 81 (40.9) | 36.5 ± 5.6 | 36.9 ± 4.5 | -0.45 ± 4.08 | 37.2 ± 4.6 | -0.73 ± 3.92 | 36.4 ± 4.7 | 0.05 ± 3.92 |
| $VO_2$max |       |           |           |           |           |           |           |           |
| Low | 66 (33.3) | 31.4 ± 2.6 | 34.1 ± 3.7 | -2.64 ± 2.88 | 34.2 ± 3.7 | -2.80 ± 2.89 | 33.5 ± 3.6 | -2.07 ± 2.87 |
| Middle | 66 (33.3) | 37.7 ± 1.6 | 38.6 ± 3.5 | -0.91 ± 3.36 | 39.0 ± 3.6 | -1.27 ± 3.45 | 38.3 ± 3.7 | -0.58 ± 3.60 |
| High | 66 (33.3) | 46.0 ± 4.2 | 43.0 ± 3.8 | 3.02 ± 4.23 | 43.4 ± 3.6 | 2.58 ± 3.99 | 42.9 ± 3.7 | 3.08 ± 3.87 |
obtained results similar to those of previous studies, i.e., no considerable difference was observed in accuracy among the three body fat-related variables, i.e., BMI, WG, and %fat (bioelectrical impedance analysis). Although the SEE value of the BMI model (4.29 ml·kg$^{-1}$·min$^{-1}$ or 11.2%) was relatively higher than those of the WG (4.17 ml·kg$^{-1}$·min$^{-1}$ or 10.9%) and %fat (4.13 ml·kg$^{-1}$·min$^{-1}$ or 10.8%) models, which are consistent with the findings reported by Wier et al. [15], the BMI model could be more convenient than the other models because BMI is a basic and less burdensome assessment item in adult health checkups. Therefore, the following equation model is suggested for VO$_{2\text{max}}$ estimation in the present study (using sex = 0 for women and 1 for men): 

$$\text{VO}_2\text{max} = 59.96 + (-0.23 \times \text{age}) + (7.39 \times \text{sex}) + (-0.79 \times \text{BMI}) + (0.33 \times \text{PA score})$$

There are some limitations to the present study. First, response bias may have occurred because the participants had advance knowledge of the experimental procedure, i.e., they could decide to participate in this study after viewing our research advertisement, which may have led to greater inclusion of participants preferring PA or exercise. Second, CE analyses (Table 5) and scatter graphs (Fig. 1) showed that the CRF evaluation model derived in the present study significantly underestimated VO$_{2\text{max}}$ in participants with high fitness and overestimated VO$_{2\text{max}}$ in those with low fitness. This systematic error usually occurs in VO$_{2\text{max}}$ estimation studies [15, 17]. As pointed out by other researchers [15, 17], while underestimation in individuals with high fitness may not be a pressing problem because high fitness relates to low disease and mortality risks, overestimation in individuals with low fitness may be more problematic because low fitness relates to increasing disease risks. A correction method such as a compensation formula or including a convenient stress test should be considered to correct the error, particularly in individuals with low fitness. Third, we could not include approximately 50% of the participants in test–retest reliability analyses because they did not participate in the second round of m-WLAQ. Participant selection bias could have occurred because the selection was not conducted at random but in accordance with participant convenience. Fourth, in recent public health research, moderate-to-vigorous intensity PA (MVPA) has been treated as an important terminology separately from SB [30, 31]. MVPA and SB are defined as accelerometry-measured PA of $\geq 3.0$ metabolic equivalents (METs) and PA of $\leq 1.5$ METs, respectively [32]. m-WLAQ can assess SB but not MVPA.

| Table 6 Test–retest reliability of sitting times and PA-related values measured using data obtained with m-WLAQ at sitting times 1 and 2 (n = 97) |
|----------------|---------|----------------|
| Sitting time   | ICC     | 95% CI         |
| Commuting time | 0.67    | 0.54–0.77      |
| Working time   | 0.94    | 0.91–0.96      |
| Leisure time on workday | 0.79    | 0.70–0.85      |
| Non-workday    | 0.66    | 0.53–0.76      |
| Total          | 0.81    | 0.73–0.87      |
| Exercise frequency |       |               |
| Leisure time on workday | 0.78    | 0.69–0.85      |
| Non-workday    | 0.88    | 0.83–0.92      |
| Exercise duration |       |               |
| Leisure time on workday | 0.82    | 0.68–0.90      |
| Non-workday    | 0.71    | 0.55–0.82      |
| PA or Exercise intensity |       |               |
| Working time   | 0.75    | 0.65–0.83      |
| Leisure time on workday | 0.55    | 0.29–0.73      |
| Non-workday    | 0.48    | 0.25–0.66      |
| PA score       | 0.87    | 0.82–0.91      |

CI: confidence interval, ICC: intraclass correlation coefficient, PA: physical activity, m-WLAQ: modified Worker’s Living Activity-time Questionnaire.

Fig. 1 Relationship between measured and estimated VO$_{2\text{max}}$ in (A) BMI, (B) WG, and (C) %fat models. The regression line and correlation coefficient with P values are displayed. The dashed lines are the identity lines of the measured and estimated VO$_{2\text{max}}$. 

Matsuo et al. BMC Public Health (2020) 20:22
Conclusion

The PA score obtained using m-WLAQ, rather than sitting time, correlated well with measured VO$_{2\text{max}}$ and had a favorable test–retest reliability. The equation model that included PA score with age, sex, and BMI had a favorable validity for estimating VO$_{2\text{max}}$. Thus, the study suggests that m-WLAQ can be a useful questionnaire to assess workers’ CRF and SB.

The findings of this study could help to advance the quality of future epidemiological surveys in occupational health research fields. On the other hand, given the CRF classification based only on age, sex, BMI, and questionnaire results could impose a limitation for detecting differences among individuals, further research is necessary to advance CRF assessment. Use of sophisticated wearable sensors could improve classification [33].

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12889-019-8067-4.

Additional file 1. The modified Worker’s Living Activity-time Questionnaire (m-WLAQ)

Abbreviations

BMI: Body mass index; CE: Constant error; CRF: Cardiorespiratory fitness; ICC: Intraclass correlation coefficients; m-WLAQ: modified Worker’s Living Activity-time Questionnaire; PA: Physical activity; PRESS: Predicted residual sum of squares; SB: Sedentary behavior; SEE: Standard error of estimate; VO$_{2\text{max}}$: Maximal oxygen consumption; WG: Waist girth

Acknowledgments

We thank researchers and staff members of the Research Center for Overwork-Related Disorders in the National Institute of Occupational Safety and Health, Japan for their funding and clerical support to this study.

Authors’ contributions

The contributions of each author were as follows: TM contributed to developing study concept and design, data acquisition, data analysis, and manuscript writing; RS contributed to developing study concept and design, data acquisition, data analysis, and manuscript revisions; and MT contributed to developing study concept and design, data acquisition, data analysis, and manuscript revisions. All authors read and approved the final manuscript.

Funding

The present work was supported by Industrial Disease Clinical Research Grant from the Japanese Ministry of Health, Labour and Welfare (150903–01 and 180902–01). The funding body had no control over the design of the study, data collection, analyses, interpretation, writing, and publication of this work.

Availability of data and materials

Raw data were generated at the Research Center for Overwork-Related Disorders (RECORD) (https://www.jniosh.johas.go.jp/en/groups/overwork.html). On reasonable request, derived data supporting the findings of this study are available from the corresponding author, TM, after approval from RECORD and the Research Ethics Committee.

Ethics approval and consent to participate

This study was conducted in accordance with the guidelines proposed in the Declaration of Helsinki. The Ethical Committee of the National Institute of Occupational Safety and Health, Japan reviewed and approved the study protocol (ID H2810). The aims and design of this study were explained to each participant before they gave their written informed consent.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 15 September 2019 Accepted: 11 December 2019

Published online: 08 January 2020

References

1. Katmazmyr PT, Church TS, Craig CL, Bouchard C. Sitting time and mortality from all causes, cardiovascular disease, and cancer. Med Sci Sports Exerc. 2009;41(5):998–1005.
2. Grontved A, Hu FB. Television viewing and risk of type 2 diabetes, cardiovascular disease, and all-cause mortality: a meta-analysis. JAMA. 2011;305(23):2448–55.
3. Church TS, Thomas DM, Tudor-Locke C, Katzmarzyk PT, Earnest CP, Rodarte RQ, et al. Trends over 5 decades in U.S. occupation-related physical activity and their associations with obesity. PLoS One. 2012;8(5):e19657.
4. van der Ploeg HP, Moller SV, Hanhnez H, van der Beek AJ, Holtermann A. Temporal changes in occupational sitting time in the Danish workforce and associations with all-cause mortality: results from the Danish work environment cohort study. Int J Behav Nutr Phys Act. 2015;12:71.
5. Kodama S, Saito K, Tanaka S, Maki M, Yachi Y, Asumi M, et al. Cardiorespiratory fitness as a quantitative predictor of all-cause mortality and cardiovascular events in healthy men and women: a meta-analysis. JAMA. 2009;301(19):2024–35.
6. Willis BL, Gao A, Leonard D, Defina LF, Berry JD. Midlife fitness and the development of chronic conditions in later life. Arch Intern Med. 2012;172(7):1333–40.
7. Bachmann JM, DeFina LF, Franzini L, Gao A, Leonard DS, Cooper KH, et al. Cardiorespiratory fitness in middle age and health care costs in later life. J Am Coll Cardiol. 2015;66(17):1876–85.
8. Clausen JSR, Marott JL, Holtermann A, Gyntelberg F, Jensen MT. Midlife cardiorespiratory fitness and the long-term risk of mortality: 46 years of follow-up. J Am Coll Cardiol. 2018;72(9):987–95.
9. Kokeye-Keadle S, Libertine A, Lyden K, Staudenmayer J, Freedson PS. Validation of wearable monitors for assessing sedentary behavior. Med Sci Sports Exerc. 2011;43(8):1561–7.
10. Atkin AJ, Gorely T, Clemes SA, Yates T, Edwoodson C, Brage S, et al. Methods of measurement in epidemiology: sedentary behaviour. Int J Epidemiol. 2012;41(5):1460–71.
11. Matsuo T, Sasai H, So R, Ohkawara K. Percentage-method improves properties of Workers’ sitting- and walking-time questionnaire. J Epidemiol. 2016;26(8):405–12.
12. Matsuo T, Sasai H, Ohkawara K. Evaluation of Worker’s living activity-time questionnaire (JNIOSH-WLAQ) primarily to assess workers’ sedentary behavior. Sangyo Eiseigaku Zasshi. 2017;59(6):219–28.
13. Cao ZB, Miyatake N, Higuchi M, Miyachi M, Ishikawa-Takata K, Tabata I. Predicting VO2max with an objectively measured physical activity in Japanese women. Med Sci Sports Exerc. 2010;42(1):179–86.
14. Jackson AS, Blair SN, Mahar MT, Wier LT, Ross RM, Stuteville JE. Prediction of functional aerobic capacity without exercise testing. Med Sci Sports Exerc. 1990;22(6):863–70.
15. Wier LT, Jackson AS, Blair SN. Methods of measurement in epidemiology: sedentary behaviour. Int J Epidemiol. 2011;40(11):2024–35.
16. Wier LT, Jackson AS, Ayers GW, Arenare B. Nonexercise models for estimating VO$_{2\text{max}}$ with waist girth, percent fat, or BMI. Med Sci Sports Exerc. 2006;38(3):355–61.
17. Suraci R, Jackson AS, LaMonte MJ, Morrow JR Jr, Blair SN, Wareham NJ, et al. Assessing cardiorespiratory fitness without performing exercise testing. Am J Prev Med. 2005;29(3):185–93.
18. Nes BM, Janszky I, Vatten LJ, Nilson TJ, Aspernes ST, Wilslef U. Estimating VO$_{2\text{peak}}$ from a nonexercise prediction model: the HUNT study, Norway. Med Sci Sports Exerc. 2011;43(11):2024–30.
19. Peterson MU, Peiper CF, Morey MC. Accuracy of VO$_{2\text{max}}$ prediction equations in older adults. Med Sci Sports Exerc. 2003;35(1):145–9.
20. Matsuo T, Satoome K, Seino S, Eto M, Shimojo N, Matsushita A, et al. Low-volume, high-intensity, aerobic interval exercise for sedentary adults: VO$_{2\text{max}}$, cardiac mass, and heart rate recovery. Eur J Appl Physiol. 2014;114(9):1963–72.
21. Riebe D, Ehrman JK, GMM L. Clinical Exercise Testing and Interpretation. ACSM’s Guidelines for Exercise Testing and Prescription tenth edition: Wolters Kluwer; 2016. p. 111–42.
22. Holiday DB, Ballard JE, McKeeown BC. PRESS-related statistics: regression tools for cross-validation and case diagnostics. Med Sci Sports Exerc. 1995;27(4):612–20.
23. Rosner BA. Fundamentals of biostatistics. 6th ed. Belmont: Thomson Higher Education; 2006. p. 569.
24. Eriksen L, Gronbaek M, Helge JW, Tolstrup JS. Cardiorespiratory fitness in 16025 adults aged 18-91 years and associations with physical activity and sitting time. Scand J Med Sci Sports. 2016;26(12):1435–43.
25. Malek MH, Housh TJ, Berger DE, Coburn JW, Beck TW. A new nonexercise-based VO2(max) equation for aerobically trained females. Med Sci Sports Exerc. 2004;36(10):1804–10.
26. Leger LA, Lambert J. A maximal multistage 20-m shuttle run test to predict VO2max. Eur J Appl Physiol Occup Physiol. 1982;49(1):1–12.
27. Plasqui G, Westerterp KR. Accelerometry and heart rate as a measure of physical fitness: proof of concept. Med Sci Sports Exerc. 2005;37(5):872–6.
28. Kiwon SB, Ahn JW, Lee SM, Lee J, Lee D, Hong J, et al. Estimating maximal oxygen uptake from daily activity data measured by a watch-type fitness tracker: cross-sectional study. JMIR Mhealth Uhealth. 2019;7(6):e13327.
29. Matthews CE, Heil DP, Freedson PS, Pastides H. Classification of cardiorespiratory fitness without exercise testing. Med Sci Sports Exerc. 1999;31(3):486–93.
30. Gebel K, Ding D, Chye T, Stamatakis E, Brown WJ, Bauman AE. Effect of moderate to vigorous physical activity on all-cause mortality in middle-aged and older Australians. JAMA Intern Med. 2015;175(6):970–7.
31. Fukushima N, Kitabayashi M, Kikuchi H, Sasai H, Oka K, Nakata Y, et al. Comparison of accelerometer-measured sedentary behavior, and light- and moderate-to-vigorous-intensity physical activity in white- and blue-collar workers in a Japanese manufacturing plant. J Occup Health. 2018;60(3):246–53.
32. Gibbs BB, Hergenroeder AL, Katzmarzyk PT, Lee IM, Jakicic JM. Definition, measurement, and health risks associated with sedentary behavior. Med Sci Sports Exerc. 2015;47(6):1295–300.
33. Klepin K, Wing D, Higgins M, Nichols J, Godino JG. Validity of cardiorespiratory fitness measured with Fitbit compared to VO2max. Med Sci Sports Exerc. 2019;51(11):2251–6.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.