BMJ Open

Is questionnaire-based sitting time inaccurate and can it be improved? A cross-sectional investigation using accelerometer-based sitting time

Nidhi Gupta,[1] Caroline Stordal Christiansen,[1] Christiana Hanisch,[2] Hans Bay,[1] Hermann Burr,[2] Andreas Holtermann[1,3]

To cite: Gupta N, Christiansen CS, Hanisch C, *et al*. Is questionnaire-based sitting time inaccurate and can it be improved? A cross-sectional investigation using accelerometer-based sitting time. *BMJ Open* 2017;**7**: e013251. doi:10.1136/bmjopen-2016-013251

► Prepublication history and additional material is available. To view please visit the journal (http://dx.doi.org/10.1136/bmjopen-2016-013251).

Received 29 June 2016
Revised 15 October 2016
Accepted 14 December 2016

[1]National Research Centre for the Working Environment, Copenhagen, Denmark
[2]Federal Institute for Occupational Safety and Health (BAuA), Nöldnerstr, Berlin, Germany
[3]Institute of Sports Science and Clinical Biomechanics, University of Southern Denmark, Odense, Denmark

Correspondence to
Dr Nidhi Gupta;
ngu@nrcwe.dk

ABSTRACT

Objectives: To investigate the differences between a questionnaire-based and accelerometer-based sitting time, and develop a model for improving the accuracy of questionnaire-based sitting time for predicting accelerometer-based sitting time.

Methods: 183 workers in a cross-sectional study reported sitting time per day using a single question during the measurement period, and wore 2 Actigraph GT3X+ accelerometers on the thigh and trunk for 1–4 working days to determine their actual sitting time per day using the validated Acti4 software. Least squares regression models were fitted with questionnaire-based siting time and other self-reported predictors to predict accelerometer-based sitting time.

Results: Questionnaire-based and accelerometer-based average sitting times were ≈272 and ≈476 min/day, respectively. A low Pearson correlation (r=0.32), high mean bias (204.1 min) and wide limits of agreement (549.8 to −139.7 min) between questionnaire-based and accelerometer-based sitting time were found. The prediction model based on questionnaire-based sitting explained 10% of the variance in accelerometer-based sitting time. Inclusion of 9 self-reported predictors in the model increased the explained variance to 41%, with 10% *optimism* using a resampling bootstrap validation. Based on a split validation analysis, the developed prediction model on ≈75% of the workers (n=132) reduced the mean and the SD of the difference between questionnaire-based and accelerometer-based sitting time by 64% and 42%, respectively, in the remaining 25% of the workers.

Conclusions: This study indicates that questionnaire-based sitting time has low validity and that a prediction model can be one solution to materially improve the precision of questionnaire-based sitting time.

Strengths and limitations of this study

- Use of a relatively large objective data sample including 9560 valid hours with around 400 days.
- Use of accelerometers and Acti4 software with high sensitivity and specificity for determining sitting.
- Inclusion of the study population of only blue-collar workers limiting the generalisability of the results.
- Findings specific to the single item questionnaire of sitting duration used in this study.

INTRODUCTION

There are several hints in the literature which say that prolonged sitting time is associated with all-cause mortality,[1 2] cardiovascular and metabolic diseases,[1] and chronic diseases such as cancer[1 3] and obesity.[4] Consequently, it is recommended to limit daily the time spent sitting.[1 5] However, most of these findings are based on questionnaire-based sitting time which have been criticised for giving systematically biased and imprecise estimates compared with objective measurements of sitting.[2 6–8]

Previous studies have found low-to-moderate correspondence between questionnaire-based and accelerometer-based sitting time.[6 8 9] However, these studies are either based on few participants or limited accelerometer measures of sitting time.[8–10] Therefore, there is a need to further investigate the correspondence between questionnaire-based and accelerometer-based sitting time.

Despite potentially being inaccurate and biased,[11] the questionnaire-based sitting time will still be needed in large-scale cohort studies and surveillance because it is inexpensive, easy to administer and does not affect the behaviour of the participant.[12] Thus, it is of value to investigate if it is possible to improve the predictive ability of questionnaire-based sitting time by developing a statistical model estimating accelerometer-based sitting time.

Prediction models to estimate accurate measurements have gained attention in the field of body composition[13] and occupational hygiene[14] but are rarely used in the field of sedentary behaviour.

Thus, the main aim of this study was therefore to (1) investigate the agreement between sitting time measured by questionnaire and accelerometers during free living, and (2) build and evaluate a statistical model to predict accelerometer-based sitting time using questionnaire-based sitting time and other available self-reported measures.

MATERIALS AND METHODS

Study design and population

Workers were recruited from seven workplaces in Denmark engaged in blue-collar occupations from the cross-sectional 'New method for Objective Measurements of physical Activity in Daily living (NOMAD)' study.[15] NOMAD was planned to develop and test techniques for collecting valid, objective measurements of physical stress and strain (physical activity and exposure) and compare them with self-reported exposure measures. Information about design, methods, and inclusion and exclusion criteria are described in more detail elsewhere.[15 16]

Blue-collar workers with varying physical exposures at work (ie, construction workers, cleaners, garbage collectors, manufacturing workers, assembly workers, healthcare workers and mobile plant operators) were recruited from workplaces (conveniently) chosen through contact with trade unions and safety representatives. Workplaces were considered eligible if the workers were allowed to participate in the study during working hours. The collection of data was conducted from October 2011 to April 2012.

The study was conducted in accordance with the Helsinki declaration. Informed consent was obtained from all individual participants included in the study.

Procedure

Data collection was conducted over 4 days with research staff visiting the workers at the workplace on days 1 and 4. On the first day, workers interested in participating in the study underwent anthropometric measurements, completed a questionnaire for the information on various candidate predictors and were equipped with accelerometers. On day 4, workers returned the accelerometers and completed a short questionnaire on duration spent on different sedentary and physical activities including an item regarding self-reported sitting time.

Questionnaire-based sitting time

Workers responded to the following question: 'How long time per working day (24 hours) did you spent sitting (including transportation)?' Workers were instructed to take into account the sitting time during the accelerometer wear period and provide a total number of hours and minutes spent sitting. Using similar items, workers were also asked to report their physical activities (slow and fast walking, biking and running time) during a whole day. Subsequently, responses were verified and adjusted for missing values and converted into total minutes. These single items are inspired by the International Physical Activity Questionnaire (IPAQ)[17] and modified Monica Optional Study of Physical Activity Questionnaire.[18]

Accelerometer-based sitting time

Field measurements using two accelerometers (Actigraph GT3X, ActiGraph LLC) for ~4 consecutive days (4×24 hours), a period generally covering at least two working days, were performed.

By using Fixomull (BSN Medical GmbH, Hamburg, Germany), a double side tape (3 M, Hair-Set, St. Paul, Minnesota, USA) and a waterproof film (OpSite flexifix, Smith & Nephew, London, England), two accelerometers were placed at the recommended and standardised position directly on the skin of the thigh and trunk.[19 20] One accelerometer was placed at the medial front of the right thigh, midway between the hip and knee joints.[16 19 21] The other accelerometer was placed at processus spinosus at the level of T1–T2.[16 20–22] The workers were instructed (1) to take off the accelerometers if they caused itching or if any other kind of discomfort such as disturbed sleep occurred, (2) to perform a reference measurement in an upright standing position for 15 s every day, and (3) to fill in a short diary every day with working hours, time in bed (going to bed, to sleep, and getting out of bed), non-wear time and time of reference measurement.[23]

Initialisation of the Actigraph for recording and downloading of data was carried out using the manufacturer's program (Actilife Software V.5.5, ActiGraph LLC, Pensacola, Florida, USA).

The accelerometer data were further analysed using the Acti4 software, estimating the type, duration, intensity and variation of physical activities (walking, running, cycling) and body postures (standing and sitting) across the day(s) with a sensitivity and specificity of more than 98% and 99%, respectively.[19] In short, accelerometer data are first low-pass filtered with a 5 Hz fourth order Butterworth filter and then split up into 2 s intervals with 50% overlap. Afterwards, the individual's reference measurement (ie, standing in an upright position for 15 s on every measured day) and values of the thigh and trunk accelerometer were used to obtain the coordinate transformation between the axis of the accelerometers and the orientation of the thigh and trunk. The occurrence of sitting postures was identified according to the procedure from Gupta *et al*.[16] The sitting posture was defined as the posture in which the inclination of the thigh accelerometer is above 45° and the trunk accelerometer is below 45°.[16 19] Acti4 has determined sitting posture during free-living conditions with a sensitivity and specificity of 98% and 93%.[19]

All non-working days and non-wear periods were excluded from the analysis according to the previously described procedure.[16 23] Only Actigraph recordings including bedtime from days with a minimum of 23 valid hours were included in the analyses. Subsequently, the sitting time per day was calculated as the average of the sitting measurements within 24 hours on all valid days.

Candidate predictors for building the statistical model

As there are not many studies conducted before investigating the potential bias factors of specifically self-reported sitting time, we chose potential predictors which could be theoretically associated with both sitting time and physical activity. Potential predictors for measured sitting time were based on previous studies, which were demographical (age and gender),[24–26] lifestyle related (body mass index (BMI), smoking status and dietary habits),[25–27] physical activity and demands related (standing still at work, sitting at work, walking duration on working days, and leisure-time physical activities),[24 28] work related (rating of perceived exertion (RPE), influence at work, and physical fatigue[16 28 29]) and health related (lower back pain)[16] variables. Age was determined by a unique Danish personal identification number while gender was determined using a question. BMI was calculated from objectively measured height and weight and categorised as normal (≤24.9), overweight (25–29.9) and obese (≥30). Smoking status was determined according to previous studies and categorised into non-smokers, light smokers (≤15 g of tobacco) and heavy smokers (>15 g tobacco).[30] RPE was determined using a single item 'How physically demanding do you normally perceive your working situation?'[22] with a modified scale of 0–9. Self-reported leisure time physical activity (LTPA) was measured according to previous study[31] and dichotomised into low and high LTPA.[32] Influence at work (decision authority) was determined according to the procedure explained in Gupta *et al*.[15] Workers reported their pain intensity in low back on a scale from 0 (no pain) to 9 (worst pain).[16] The workers reported the proportion of working time spent sitting and standing still on a scale of 1 (almost all the time) to 6 (almost never).[33] Physical fatigue was determined using single item 'How physically rested do you feel when you wake up in the morning? Think of the mornings where you have been at work the day before?' with responses ranging from 1 (almost always) to 5 (never).[34] Dietary habits were determined using single item which was inspired by a previous study, 'How often do you usually eat and/or drink candy, ice cream, chocolate, soft drinks?' with responses from 1 (daily) to 4 (rarely).[35]

Statistical analyses

All statistical operations were performed using the R software (R: A language and environment for statistical computing [program]. Vienna, Austria: R Foundation for Statistical Computing, 2014). The level of agreement between questionnaire-based and accelerometer-based sitting times was investigated with Bland-Altman plots[36] and their constant and regression-based limits of agreement were calculated.[37]

Prediction model building

A hierarchical cluster analysis on the candidate predictors was performed using Hoeffding dependence measure since it offers the possibility of including continuous as well as categorical variables in the analysis.[38] The number of clusters was decided among the authors on the basis of the cluster analysis's dendrogram and the available degrees of freedom for statistical analysis. From each resulting cluster, the variable with the largest relative dispersion indicated by largest coefficient of variation was chosen.

For building the prediction model, a 'crude' model was developed fitting the accelerometer-based sitting time as outcome with questionnaire-based sitting time as predictor using a least square regression analysis. Next, a 'full' model including all chosen variables from cluster analysis, together with questionnaire-based sitting time was fitted using multiple least-squares linear regression. In the model, the accelerometer-based sitting time was entered as outcome while all chosen variables from cluster analysis were entered as predictors. The performance of all crude and full calibration models was evaluated by the coefficient of determination (R^2) and R^2 adjusted for the number of terms in the model and the mean square error (MSE) of estimation (ie, the apparent performance). The residuals of the model were scrutinised for being normally distributed and homoscedastic.

Prediction model validation

To evaluate the developed statistical prediction model, a separate resampling validation using ordinary bootstrapping with 500 resamples was used first. For each of the 500 bootstrap resamples, we fitted a model containing the same predictors as those used in the original model. This refitted model was then applied to the original data set, and the model fit parameters were compared with the corresponding parameters obtained in the original model fit. The differences, that is, the 'optimism' of the original model, were averaged across all bootstrap data sets, and this difference was used as an overall measure of optimism, reflecting the extent to which the original calibration model capitalised on chance. The apparent performance of the original model was then adjusted for the estimated optimism to arrive at an expected performance of the model on new data sets.

Second, the model was also evaluated using the split-sample validation method.[39] About two-thirds of the workers were randomly assigned to a 'development' group and the remaining to a 'testing' group. The resulting statistical model on the development group was evaluated in the testing group for prediction of accelerometer-based sitting time.

The validity of the models was evaluated by the goodness of fit, reduction in SD of the difference between accelerometer-based and questionnaire-based sitting time, and minimal systematic bias shown in the Bland and Altman plot graphing difference between measured and predicted accelerometer-based sitting time against their average.

RESULTS

Characteristics of the study population

Out of 358 blue-collar workers who offered participation, 183 (51%) workers who answered to the single item sitting duration and who had at least one valid day with accelerometer-based measurements of ≥23 hours were included in the final analysis (see online supplementary figure A). No significant differences were observed for the age, gender, job seniority, working hours, height and weight between the participants included (n=183) and not included (n=175) in the analysis of this study (results not shown).

In total, workers were measured for 9560 valid hours and on average, workers included in the statistical analyses wore the accelerometer for >2 days with ≈24 hours of measurements per day.

The characteristics of the study group are shown in table 1. Workers were on average 45 years old and a higher proportion of workers were males (60%). Workers were exposed to a wide range of sitting time (157–851 min) on working days based on the accelerometer measures. The questionnaire-based sitting time was materially lower than accelerometer-based sitting time (table 1).

The Bland-Altman plot in figure 1 shows that the questionnaire-based estimates of sitting time were significantly underestimated, that is, bias of ~204 min (≈43%) compared with the accelerometer-based measurements. Both constant and regression-based and limits of agreement were wide, indicating a large interindividual variation between the two measures of sitting time. We found a weak pattern of increasing mean difference between accelerometer-based and questionnaire-based sitting times with increasing average sitting time (R^2=0.02, p=0.04). Additionally, a low positive Pearson correlation was observed between the two measures of sitting time (r=0.32, p<0.001).

Table 1 Characteristics of the blue-collar workers (N=183) involved in the statistical analysis

Variables	N	Per cent	Minimum	Maximum	Mean	SD
Age (years)	183		21	65	44.9	9.8
Gender						
Male	110	60				
Female	73	40				
Leisure time physical activity						
Low	87	48				
High	94	52				
Smoking status						
Non-smoker or ex-smokers	99	59				
≤15 cigarettes (light smokers)	38	23				
>15 cigarettes (heavy smokers)	31	18				
Influence at work in 0–100%	180		0	100	43.7	23.0
BMI (kg/m^2)						
>25	75	41				
25–30	72	39				
>30	36	20				
LBP intensity (0–9)	178		0	9	3.0	2.6
RPE (0–9)	180		0	9	5.5	1.9
Dietary habits (1–4)	179		1	4	2.6	0.9
Physical fatigue (1–5)	180		1	5	2.4	0.9
Slow walking duration (min/day)*	179		0	990	171.6	166.2
Fast walking duration (min/day)*	174		0	600	132.1	127.4
Sitting duration at work (1–6)	182		1	6	4.1	1.2
Standing still duration at work (1–6)	183		1	6	4.6	0.8
Measurements of sitting						
Questionnaire-based sitting time (min/day)*	183		0.0	900.0	271.8	160.5
Accelerometer-based sitting time (min/day)	183		157.2	851.3	475.8	138.3

RPE: 0=not hard work, 9=maximal hard work; influence at work: 0=no influence, 100=highest influence at work; LBP intensity: 0=no pain, 9=worst possible pain; dietary habits: 1=daily, 4=rarely; physical fatigue: 1=almost always, 5=never; sitting duration at work: 1=almost all the time, 6=never; standing still duration at work: 1=almost all the time, 6=never; we did not include self-reported biking and running time as potential predictors due to many missings in the data.
*The data handling strategy was inspired by the IPAQ cleaning guidelines.[40]
BMI, body mass index; IPAQ, International Physical Activity Questionnaire; LBP, low back pain; RPE, rating of perceived exertion.

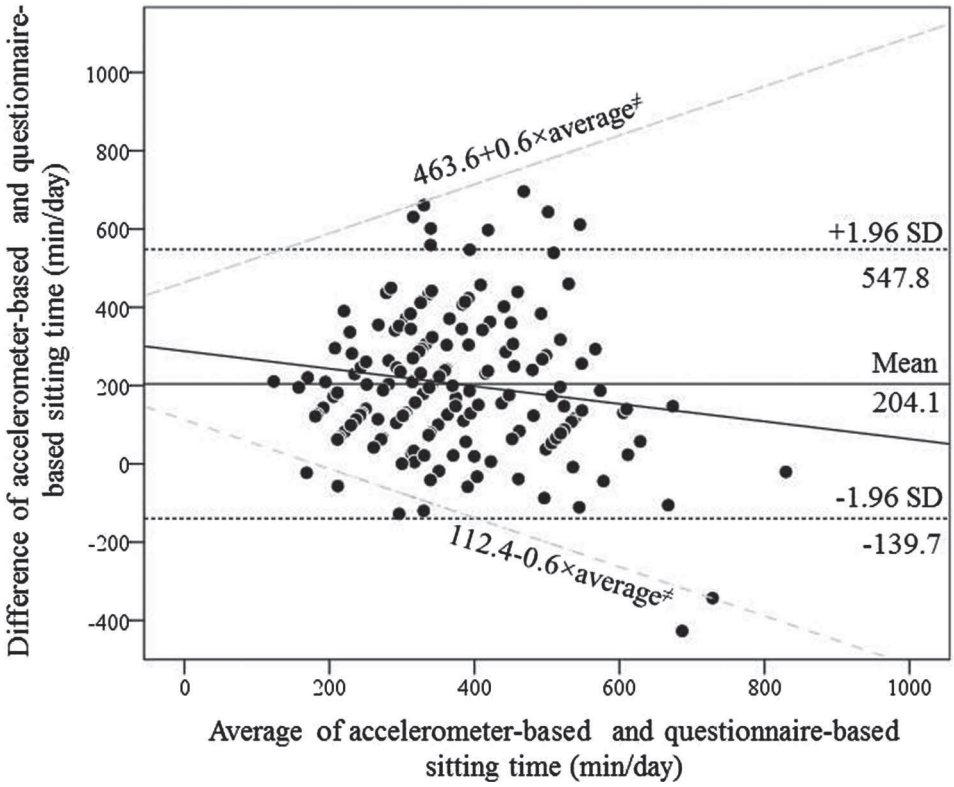

Figure 1 Bland-Altman plot showing the levels of agreement between the difference (vertical axis) and the average (horizontal axis) of accelerometer-based and questionnaire-based sitting time (min/day). Bias (mean difference) is presented as the middle horizontal line, limits of agreement (±1.96 SD) as the dashed horizontal lines, the proportional bias presented with solid regression line and the two diagonal dotted lines were the regression-based limits of agreement; ≠average of accelerometer-based and questionnaire-based sitting time.

Prediction model development

The variables BMI, gender and smoking status were treated as categorical variables and the remaining as linear variables. The dendrogram (see online supplementary figure B) obtained from cluster analysis revealed nine clusters of the variables. Based on the results of relative dispersion of the variables in each cluster, we chose BMI (cluster 1), smoking status (cluster 2), RPE (cluster 3), sitting duration per day (cluster 4), influence at work and slow walking duration per day (cluster 5), low back pain intensity (cluster 6), dietary habits (cluster 7), fast walking duration per day (cluster 8) and gender (cluster 9). The reason for choosing both variables from cluster 5 was that both variables contributed with different type of information.

The final model (Equation 1) was

$$\begin{aligned} \mathrm{SIT}_{\mathrm{accel}} = {} & \alpha_0 + \sum_{(i=1)}^{2} \alpha_1(\mathrm{BMI} = i) \\ & + \sum_{(i=1)}^{2} \alpha_2(\mathrm{Smoke} = i) + \alpha_3(\mathrm{Gender} = i) \\ & + \alpha_4 \mathrm{RPE} + \alpha_5 \mathrm{sitting\ duration} \\ & + \alpha_6 \mathrm{influence\ at\ work} \\ & + \alpha_7 \mathrm{slow\ walk\ duration} + \alpha_8 \mathrm{LBP\ intensity} \\ & + \alpha_9 \mathrm{dietary\ habits} \\ & + \alpha_{10} \mathrm{fast\ walk\ duration} + \in \end{aligned} \quad (1)$$

Where α_0 is the intercept and $\alpha_{1\text{-}10}$ are the regression coefficients for all predictors. In the equation, BMI=I and smoke=i equals 1 if the category is 'normal' and 'non-smokers', respectively, and 0 otherwise; ε is the random error term.

Table 2 shows the result of the statistical prediction model using the linear least square regression analysis, first only fitting questionnaire-based sitting time (crude model) and later all chosen variables from cluster analysis (full model) as potential predictors of the accelerometer-based sitting time. Questionnaire-based sitting time only explained 10% of the variance of the accelerometer-based sitting time. By including chosen potential predictors, the explained variance increased to 41% (R^2 adjusted for terms in the model=37%).

Figure 2 describes the association of accelerometer-based sitting time with questionnaire-based sitting time and predicted accelerometer-based sitting time using the statistical prediction model.

Bootstrapping validation

Validation using resampling bootstrapping resulted in a corrected estimate of R^2 of 32% and MSE of 12167.0. The relative difference between the original and corrected estimates was 9.7% (ie, optimism) when bootstrap models were applied on the original data set.

Split validation

The full model (equation 1) was fitted on the development group of about two-thirds of the workers (n=132). Using the resulting estimates, the accelerometer-based sitting time was predicted in the remaining workers (n=51). Using the developed model from the

Table 2 Results of the fitted models using linear least square analysis predicting objectively measured sitting time

Model	Predictor	Coefficient	p Value
Crude (N=183)	Intercept	400.74	<0.001
	Questionnaire-based sitting time	0.27	<0.001
	R^2=0.10, R^2 adjusted=0.10, MSE=17104.6		
Full model (N=154)	Intercept	367.57	<0.001
	BMI (overweight=1, others=0)	−26.95	0.183
	BMI (obese=1, others=0)	54.58	0.031
	Smoking status (light smokers=1, others=0)	−11.69	0.600
	Smoking status (heavy smokers=1, others=0)	42.97	0.077
	Gender (female=1, male=0)	−46.08	0.017
	RPE (0–9)	−5.55	0.243
	Questionnaire-based sitting time (min/day)	0.31	<0.001
	Influence at work (0–100%)	0.94	0.017
	Slow walk duration (min/day)	−0.16	0.006
	Lower back pain intensity (0–9)	9.88	0.005
	Dietary habits (1–4)	21.40	0.024
	Fast walk duration (min/day)	−0.26	0.001
	R^2=0.41, R^2 adjusted=0.37, MSE=10332.1		

RPE: 0=not hard work, 9=maximal hard work; influence at work: 0=no influence, 100=highest influence at work; lower back pain intensity: 0=no pain, 9=worst possible pain; dietary habits: 1=daily, 4=rarely.
BMI, body mass index; MSE, mean square error; R^2=coefficient of determination; RPE, rating of perceived exertion.

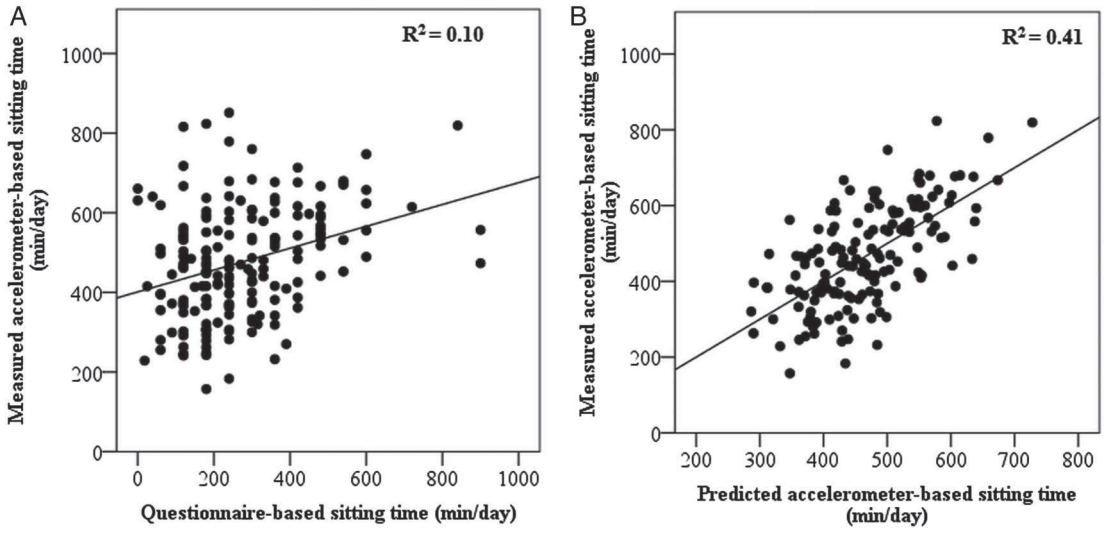

Figure 2 Diagramatic representation of the association between accelerometer-based sitting time per day and questionnaire-based sitting time per day (A) and with improved questionnaire-based sitting time using the developed statistical model including BMI, gender, smoking status, rating of perceived exertion, influence at work, slow and fast walking duration per day, low back pain intensity, and dietary patterns (B) among 183 blue-collar workers.

development group, the mean and SD of the difference between accelerometer-based and questionnaire-based sitting times decreased from ≈184 min to only ≈66 min (decreased by 64%) and from 178.0 min to only 103.9 min (decreased by 42%), respectively, in the testing group. Furthermore, the application of the prediction model improved the accuracy of questionnaire-based sitting time in the testing group from 10% to 36% (figure 3) in predicting the measured accelerometer-based sitting time. The Bland and Altman plot in the testing group indicated no systematic bias, lower mean error and narrower limits of agreement due to the prediction using the developed model (figure 3), compared with questionnaire-based sitting time.

DISCUSSION

This study compared sitting duration (ie, 24 hours) measured via self-administered questionnaire-based and accelerometer-based methods previously shown to have a high sensitivity and specificity in estimating sitting time during free living.[19] Results showed a poor correspondence and level of agreement between these methods of sitting time. Additionally, a statistical prediction model

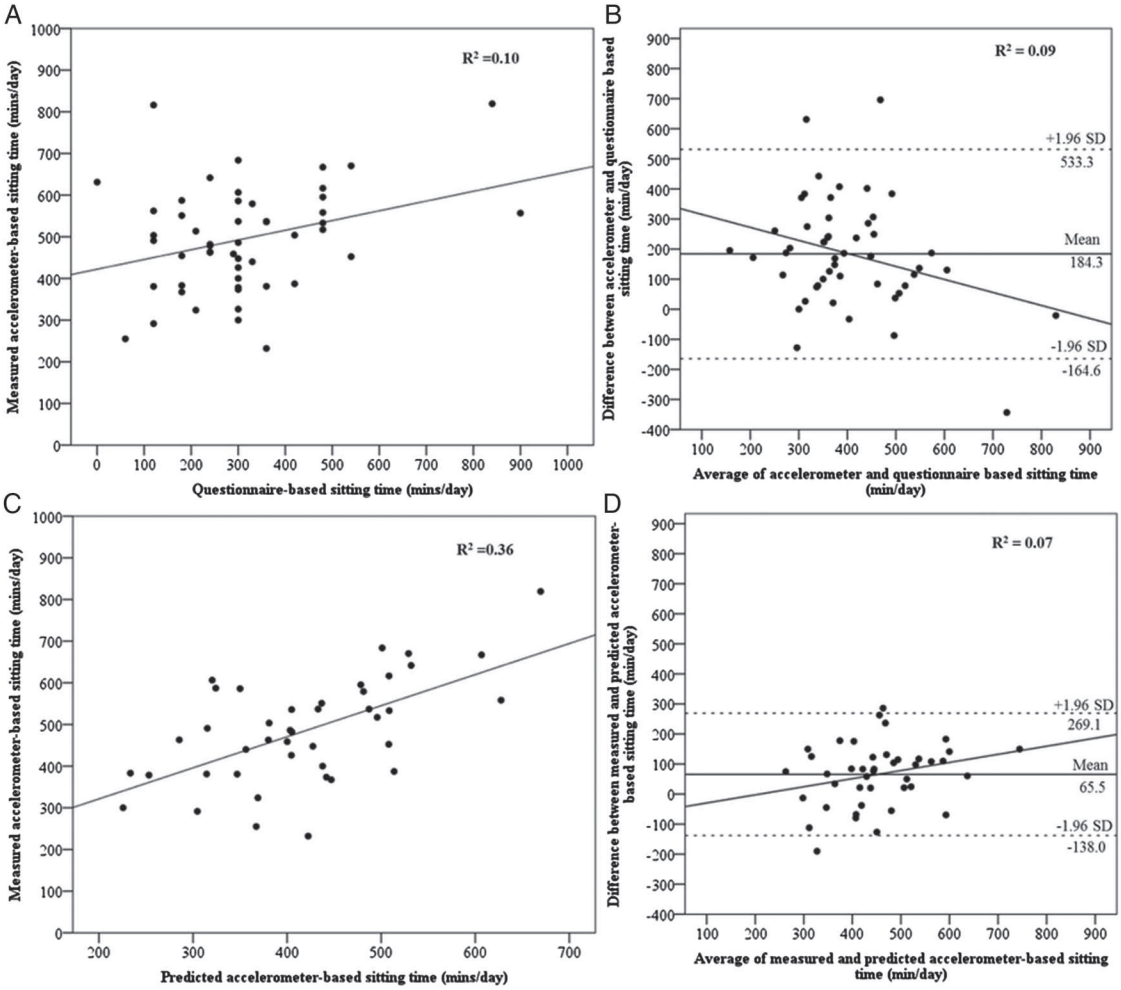

Figure 3 Accuracy of full prediction model developed in the 'development group' for predicting measured accelerometer-based sitting time in the 'testing group'. (A and C) are the regression plots of the prediction of measured accelerometer-based sitting time with questionnaire-based sitting time (A) and with the developed prediction model using questionnaire-based sitting time and other factors (C) among testing sample (n=51). (B and D) are Bland-Altman plots of the prediction of measured accelerometer-based sitting time based on questionnaire-based sitting time (B) and based on prediction values (D). In (B and D), the mean error (mean difference) is presented as the middle horizontal line, limits of agreement (±1.96 SD) as the dashed horizontal lines and the proportional bias presented with regression line.

based on questionnaire-based sitting time explained only 10% of the variance in accelerometer-based sitting time. However, the explained variance was shown to increase to 41% with inclusion of BMI, smoking status, gender, RPE, influence at work, slow and fast walking duration, low back pain, and dietary patterns in the statistical prediction model. The validation using resampling bootstrapping resulted in about only ≈10% optimism. The cross-validation based on the resulting statistical prediction model from the development group (n=132) decreased the mean and SD of the difference between the accelerometer-based and questionnaire-based sitting times by 64% and 42%, respectively, in the testing group (n=51) without introducing any systematic bias and decreasing the actual variation of the sitting time (figure 3).

Our finding shows a low correspondence (r=0.32) between questionnaire-based and accelerometer-based sitting times, indicating a large error associated with questionnaire-based measurements. Beside the low correlation, the questionnaire-based sitting time was underestimated compared with accelerometer-based sitting time by ≈204 min per day (≈57%) along with wide limits of agreement (−139.8 to 547.8 min) between the measurements. Similar findings of low correlation between questionnaire-based and accelerometer-based sitting times and an underestimation of self-reported sitting time and wide limits of agreement were also observed in previous studies using Actigraph[6 7 10 41] or ActivPAL[11 42] as criterion measure. The accuracy of our questionnaire is comparable to most available questionnaires such as IPAQ short (r=0.07 to 0.61) and long (r=0.14 to 0.49), Global Physical Activity Questionnaire (r=−0.02 to −0.40), Australian Women's Activity Survey Questionnaire (r=0.32), Recent Physical Activity Questionnaire (r=0.27), Activity Questionnaire for Adults and Adolescents (r=0.15), and Sedentary

Behavior Questionnaire (r=0.02 to 18).[43] Since the observed large error and under-reporting of sitting time per day using questionnaires can lead to misclassification of exposure to sitting time,[44] questionnaire-based sitting time should be used and interpreted with great care.

Despite the limitations, the questionnaire-based sitting time is still widely used in its 'raw' form in national surveys and large epidemiological studies. Thus, alternative ways to improve the questionnaire-based sitting time are needed. To the best of our knowledge, this is the first study to develop a statistical model to improve questionnaire-based sitting time. When only including questionnaire-based sitting time in the statistical model applied on the whole study group (n=183), only 10% of the variance in the accelerometer-based sitting time was explained. However, inclusion of BMI, smoking status, gender, RPE, influence at work, slow and fast walking duration, low back pain, and dietary patterns in the statistical prediction model substantially increased the explained variance to 41% (R^2 adjusted=0.37). This finding indicates that inclusion of the aforementioned variables in a statistical prediction model improves the precision of questionnaire-based sitting time. The model can be used in predicting objectively measured sitting time based on the combination of the predictors in the model. For example, a 32-year-old male worker who is obese and a non-smoker reported RPE of 8 on a scale of 0–9, influence at work of 37.5% on a scale of 0–100%, eating poor diet 1–2 times/week (3 on a scale of 1–4), and reported sitting in 240 min, slow walking in 540 min and fast walking in 120 min per day, was predicted via our full model to be sitting for 402.8 min per day according to objective measurements.

Although the large improvement in the accuracy seems promising, there is still a large unexplained variance in the accelerometer-based sitting time. This could be explained by the additional predictors which were not available in our study, such as psychosocial variables.[26] We therefore encourage future studies to explore additional predictors which could increase the explained variance of accelerometer-based sitting time. Not all the predictors of our full model are the most commonly used variables in large epidemiological studies and surveys. Thus, we developed a new prediction model based on commonly used variables (ie, age, gender, BMI, smoking status, low back pain and self-reported sitting time) which resulted in an explained variance of 25% (R^2=0.22; see online supplementary table A). Although the improvement could be higher by using the full model, the crude model or the model based on commonly available variables at least gives an opportunity for calibrating the previously collected data on questionnaire-based sitting time.

For a prediction model to be useful, it should perform well in the original data set, and also in new samples of data. Therefore, we used the cross-validation methods to investigate the validity of the statistical prediction model. The results of bootstrapping validation resulted in 10% decrement and 18% increment in R^2 and MSE, respectively, which indicates that the model is not much biased, and gives moderately accurate estimates. Additionally, using the developed model based on a random sample of 75% of the study population (the development group, n=132), the difference between the accelerometer and questionnaire-based sitting time decreased from ≈184 to ≈66 min in the testing group (n=51). Also, the model did not introduce any systematic bias in the improved questionnaire-based sitting time, and the variation of the accelerometer-based sitting time and improved questionnaire-based sitting time was comparable. These results support the internal validity and generalisability of our developed statistical prediction model for improving the precision of questionnaire-based sitting time. However, because the study population only includes blue-collar workers, the results cannot be directly generalised to other populations. This finding also shows the importance of taking the applied predictors into account in analysing and interpreting questionnaire-based sitting time in surveillances and epidemiological studies.

Strengths, limitations, future recommendations and practical implications of the findings

The main strength of this study is the use of a relatively large objective data sample including 9560 valid hours with around 400 days measured in the analysis. Another strength of the study is addressing the inherent limitation of Actigraph software to determine sitting time. Actigraph-based software uses a threshold of 100 counts per minute to identify sitting posture which has been heavily criticised due to its inability to accurately differentiate sitting from standing postures.[45] This leads to incorrect information about temporal patterns of sitting.[45] Therefore, we used Acti4—a posture recognition software which has shown to determine sitting time during free-living conditions with a high sensitivity and specificity.[19] Another strength was the use of cluster analysis to reduce the number of predictors in the statistical model without giving any consideration to their relationship with accelerometer-based sitting time. Further strength of the study was that the recall period of the questionnaire-based sitting time was identical to the wear period of the accelerometers.

A limitation is the study population of blue-collar workers which limits the generalisability of the results. Additionally, our findings are specific to the single-item questionnaire of sitting duration used in this study. The main aim of this study was not to produce a simplified model including fewest variables possible, and we therefore recommend performing simplification of the developed model in the future for convenience in different studies. Before we could recommend using the developed model to improve the questionnaire-based sitting time, the validity of the developed model should be tested in a different study population.

Even if our prediction model could explain only 41% variance of the accelerometer-based sitting time, it can

be used for predicting the accelerometer-based sitting time in future studies where data for all predictors are collected instead of measuring the sitting time using a questionnaire. Our model can also be used to perform retrospective prediction of accelerometer-based sitting time in previous studies where data on all predictors are available. Since the accuracy of the questionnaire used in our study is comparable to other questionnaires, we expect to obtain similar predictive accuracy of our model when using other questionnaires, albeit it needs to be tested in future studies.

CONCLUSION

Our study showed a low correspondence and agreement between questionnaire-based and accelerometer-based sitting times. The developed statistical predictive model on the whole population increased the explained variance in accelerometer-based sitting time from 10% using only questionnaire-based sitting time to 41%. A bootstrapping and a cross-validation supported the validity of the developed prediction model. Thus, the developed statistical prediction model in this study provides a possibility for improving questionnaire-based sitting time among blue-collar workers.

Acknowledgements The authors would like to thank Jørgen Skotte, senior researcher at the NRCWE for his technical assistance in this study.

Contributors AH was the principle investigator for the NOMAD cohort. NG and AH had initiated this particular study. CSC was responsible for collecting the data while HB and NG were responsible for the statistical analysis in the study. NG was responsible for analysing and interpreting data and writing the first draft of the manuscript. NG, CSC, CH, HB, HB and AH reviewed, edited and approved the final draft of this manuscript. All authors have contributed in planning, designing, analysing and finalising the manuscript. All authors take full responsibility for the manuscript and have read and approved the final version of it.

Funding This study was financially supported by the Danish Working Environment Research Fund (journal number 20130069161/9), Denmark. This study was also supported by the internal funding from Federal Institute for Occupational Safety and Health (BAuA), Berlin, Germany and The National Research Centre for the Working Environment (NRCWE), Copenhagen, Denmark.

Competing interests None declared.

Patient consent Obtained.

Ethics approval The study was approved by the Ethics Committee for the Capital Region in Denmark (journal number H-2-2011-047).

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Data used in this study are available by emailing to AH (aho@nrcwe.dk).

REFERENCES

1. Katzmarzyk PT, Church TS, Craig CL, *et al*. Sitting time and mortality from all causes, cardiovascular disease, and cancer. *Med Sci Sports Exerc* 2009;41:998–1005.
2. Chau JY, Grunseit AC, Chey T, *et al*. Daily sitting time and all-cause mortality: a meta-analysis. *PLoS ONE* 2013;8:e80000.
3. Inoue M, Iso H, Yamamoto S, *et al*. Daily total physical activity level and premature death in men and women: results from a large-scale population-based cohort study in Japan (JPHC Study). *Ann Epidemiol* 2008;18:522–30.
4. Parsons TJ, Manor O, Power C. Television viewing and obesity: a prospective study in the 1958 British birth cohort. *Eur J Clin Nutr* 2008;62:1355–63.
5. Stamatakis E, Hamer M, Dunstan DW. Screen-based entertainment time, all-cause mortality, and cardiovascular events: population-based study with ongoing mortality and hospital events follow-up. *J Am Coll Cardiol* 2011;57:292–9.
6. Clark BK, Thorp AA, Winkler EA, *et al*. Validity of self-report measures of workplace sitting time and breaks in sitting time. *Med Sci Sports Exerc* 2011;43:1907–12.
7. Lagersted-Olsen J, Korshoj M, Skotte J, *et al*. Comparison of objectively measured and self-reported time spent sitting. *Int J Sports Med* 2014;35:534–40.
8. Affuso O, Stevens J, Catellier D, *et al*. Validity of self-reported leisure-time sedentary behavior in adolescents. *J Negat Results Biomed* 2011;10:2.
9. Chau JY, Van Der Ploeg HP, Dunn S, *et al*. Validity of the occupational sitting and physical activity questionnaire. *Med Sci Sports Exerc* 2012;44:118–25.
10. Clemes SA, David BM, Zhao Y, *et al*. Validity of two self-report measures of sitting time. *J Phys Act Health* 2012;9:533–9.
11. Chastin SF, Culhane B, Dall PM. Comparison of self-reported measure of sitting time (IPAQ) with objective measurement (activPAL). *Physiol Meas* 2014;35:2319–28.
12. Terwee CB, Mokkink LB, van Poppel MN, *et al*. Qualitative attributes and measurement properties of physical activity questionnaires: a checklist. *Sports Med* 2010;40:525–37.
13. Tian S, Mioche L, Denis JB, *et al*. A multivariate model for predicting segmental body composition. *Br J Nutr* 2013;110:2260–70.
14. Neitzel RL, Daniell WE, Sheppard L, *et al*. Improving exposure estimates by combining exposure information. *Ann Occup Hyg* 2011;55:537–47.
15. Gupta N, Jensen BS, Sogaard K, *et al*. Face validity of the single work ability item: comparison with objectively measured heart rate reserve over several days. *Int J Environ Res Public Health* 2014;11:5333–48.
16. Gupta N, Christiansen CS, Hallman DM, *et al*. Is objectively measured sitting time associated with low back pain? *PLoS ONE* 2015;10:e0121159.
17. Hagstromer M, Oja P, Sjostrom M. The International Physical Activity Questionnaire (IPAQ): a study of concurrent and construct validity. *Public Health Nutr* 2006;9:755–62.
18. Pereira MA, FitzerGerald SJ, Gregg EW, *et al*. A collection of Physical Activity Questionnaires for health-related research. *Med Sci Sports Exerc* 1997;29(6 Suppl):S1–205.
19. Skotte J, Korshoj M, Kristiansen J, *et al*. Detection of physical activity types using triaxial accelerometers. *J Phys Act Health* 2014;11:76–84.
20. Korshoj M, Skotte JH, Christiansen CS, *et al*. Validity of the Acti4 software using ActiGraph GT3X+accelerometer for recording of arm and upper body inclination in simulated work tasks. *Ergonomics* 2014;57:247–53.
21. Hallman DM, Gupta N, Mathiassen SE, *et al*. Association between objectively measured sitting time and neck–shoulder pain among blue-collar workers. *Int Arch Occup Environ Health* 2015;88:1031–42.
22. Lunde LK, Koch M, Knardahl S, *et al*. Musculoskeletal health and work ability in physically demanding occupations: study protocol for a prospective field study on construction and health care workers. *BMC Public Health* 2014;14:1075.
23. Villumsen M, Samani A, Jorgensen MB, *et al*. Are forward bending of the trunk and low back pain associated among Danish blue-collar workers? A cross-sectional field study based on objective measures. *Ergonomics* 2015;58:246–58.
24. Meneguci J, Sasaki JE, da Silva Santos A, *et al*. Socio-demographic, clinical and health behavior correlates of sitting time in older adults. *BMC Public Health* 2015;15:65.
25. Ferrari P, Friedenreich C, Matthews CE. The role of measurement error in estimating levels of physical activity. *Am J Epidemiol* 2007;166:832–40.
26. De Cocker K, Duncan MJ, Short C, *et al*. Understanding occupational sitting: prevalence, correlates and moderating effects in Australian employees. *Prev Med* 2014;67:288–94.
27. Charreire H, Kesse-Guyot E, Bertrais S, *et al*. Associations between dietary patterns, physical activity (leisure-time and occupational) and

television viewing in middle-aged French adults. *Br J Nutr* 2011;105:902–10.
28. Vandelanotte C, Duncan MJ, Short C, *et al*. Associations between occupational indicators and total, work-based and leisure-time sitting: a cross-sectional study. *BMC Public Health* 2013;13:1110.
29. Det Nationale Forskningscenter for Arbejdsmiljø. Arbejdsmiljø og helbred i Danmark 2010 [Work Environment and Health Survey in Denmark 2010]. Secondary Arbejdsmiljø og helbred i Danmark 2010 [Work Environment and Health Survey in Denmark 2010]. 2010. http://www.arbejdsmiljoforskning.dk/~/media/Projekter/AH2012/AH-2012-spoergeskema-endelig-version.pdf
30. Albertsen K, Hannerz H, Borg V, *et al*. Work environment and smoking cessation over a five-year period. *Scand J Public Health* 2004;32:164–71.
31. Holtermann A, Marott JL, Gyntelberg F, *et al*. Occupational and leisure time physical activity: risk of all-cause mortality and myocardial infarction in the Copenhagen City Heart Study. A prospective cohort study. *BMJ Open* 2012;2:e000556.
32. Rasmussen CD, Larsen AK, Holtermann A, *et al*. Adoption of workplaces and reach of employees for a multi-faceted intervention targeting low back pain among nurses' aides. *BMC Med Res Methodol* 2014;14:60.
33. Det Nationale Forskningscenter for Arbejdsmiljø. Arbejdsmiljø og helbred i Danmark 2012. Secondary Arbejdsmiljø og helbred i Danmark 2012. 2012. http://www.arbejdsmiljoforskning.dk/~/media/Projekter/AH2012/AH-2012-spoergeskema-endelig-version.pdf
34. The National Research Centre for the Working Environment. Den Nationale Arbejdsmiljøkohorte—NAK. Secondary Den Nationale Arbejdsmiljøkohorte—NAK. 2005. http://www.arbejdsmiljoforskning.dk/da/arbejdsmiljoedata/nak2005
35. Tilles-Tirkkonen T, Pentikainen S, Lappi J, *et al*. The quality of school lunch consumed reflects overall eating patterns in 11-16-year-old schoolchildren in Finland. *Public Health Nutr* 2011;14:2092–8.
36. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307–10.
37. Brown R, Richmond S. An update on the analysis of agreement for orthodontic indices. *Eur J Orthod* 2005;27:286–91.
38. Harrell FE. *Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis*. New York, USA: Springer, 2001.
39. Dezenberg CV, Nagy TR, Gower BA, *et al*. Predicting body composition from anthropometry in pre-adolescent children. *Int J Obes Relat Metab Disord* 1999;23:253–9.
40. IPAQ. Guidelines for Data Processing and Analysis of the International Physical Activity Questionnaire. Secondary Guidelines for Data Processing and Analysis of the International Physical Activity Questionnaire 2005. https://sites.google.com/site/theipaq/scoring-protocol
41. Van Cauwenberg J, Van Holle V, De Bourdeaudhuij I, *et al*. Older adults' reporting of specific sedentary behaviors: validity and reliability. *BMC Public Health* 2014;14:734.
42. Kozey-Keadle S, Libertine A, Staudenmayer J, *et al*. The feasibility of reducing and measuring sedentary time among overweight, non-exercising office workers. *J Obes* 2012;2012:282303.
43. Healy GN, Clark BK, Winkler EA, *et al*. Measurement of adults' sedentary time in population-based studies. *Am J Prev Med* 2011;41:216–27.
44. Celis-Morales CA, Perez-Bravo F, Ibanez L, *et al*. Objective vs. self-reported physical activity and sedentary time: effects of measurement method on relationships with risk biomarkers. *PLoS ONE* 2012;7:e36345.
45. Atkin AJ, Gorely T, Clemes SA, *et al*. Methods of measurement in epidemiology: sedentary behaviour. *Int J Epidemiol* 2012;41:1460–71.