



**Karolinska
Institutet**

Karolinska Institutet

<http://openarchive.ki.se>

This is a Peer Reviewed Accepted version of the following article, accepted for publication in *Scandinavian Journal of Medicine and Science in Sports*.

2021-05-31

Detecting prolonged sitting bouts with the ActiGraph GT3X

Kuster, Roman P; Grooten, Wilhelmus J A; Baumgartner, Daniel; Blom, Victoria; Hagströmer, Maria; Ekblom, Örjan

Scand J Med Sci Sports. 2020 Mar;30(3):572-582.

Wiley

<http://doi.org/10.1111/sms.13601>

<http://hdl.handle.net/10616/47666>

If not otherwise stated by the Publisher's Terms and conditions, the manuscript is deposited under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

1 Detecting Prolonged Sitting Bouts with the ActiGraph GT3X

2 **Running Head:** Detecting Sitting Bouts with ActiGraph GT3X

3 **Authors:** Roman P Kuster ^(1,2), Wilhelmus J A Grooten ^(1,3), Daniel Baumgartner ⁽²⁾, Victoria
4 Blom ^(4,5), Maria Hagströmer ^(1,3,6), and Örjan Ekblom ⁽⁴⁾

5 1) Division of Physiotherapy, Department of Neurobiology, Care Sciences and Society,
6 Karolinska Institutet, Stockholm (Sweden)

7 2) Institute of Mechanical Systems, School of Engineering, ZHAW Zurich University of
8 Applied Sciences, Winterthur (Switzerland)

9 3) Function Area Occupational Therapy and Physiotherapy, Allied Health Professionals,
10 Karolinska University Hospital, Stockholm (Sweden)

11 4) Åstrand Laboratory of Work Physiology, The Swedish School of Sport and Health
12 Sciences, Stockholm (Sweden)

13 5) Division of Insurance Medicine, Department of Clinical Neuroscience, Karolinska
14 Institutet, Stockholm (Sweden)

15 6) Department of Health Promoting Sciences, Sophiahemmet University, Stockholm
16 (Sweden)

17 **Corresponding Author:** Roman P Kuster, Technikumstrasse 9, 8401 Winterthur, Switzerland.
18 Mail: roman.kuster@alumni.ethz.ch, Tel.: +41-58-934-6522, Fax: +41-58-935-6522

19 **Acknowledgments, Funding, and Conflict of Interest**

20 The authors acknowledge the support of Cahit Atilgan in programming the random forest
21 classifier in Python. The study was not supported by activPAL or ActiGraph. The authors do
22 not declare a conflict of interest. The results of the study are presented clearly, honestly, and
23 without fabrication, falsification, or inappropriate data manipulation.

24 **Abstract**

25 The ActiGraph has a high ability to measure physical activity, however, it lacks an accurate
26 posture classification to measure sedentary behaviour. The aim of the present study was to
27 develop an ActiGraph (waist-worn, 30Hz) posture classification to detect prolonged sitting
28 bouts, and to compare the classification to proprietary ActiGraph data. The activPAL, a highly
29 valid posture classification device, served as reference criterion.¹

30 Both sensors were worn by 38 office workers over a median duration of 9 days. An automated
31 feature selection extracted the relevant signal information for a minute based posture
32 classification. The machine-learning algorithm with optimal feature number to predict the time
33 in prolonged sitting bouts (≥ 5 and ≥ 10 minutes) was searched and compared to the activPAL

34 using Bland-Altman statistics. The comparison included optimised and frequently used cut-
35 points (100 and 150 counts-per-minute (cpm), with and without low-frequency-extension (LFE)
36 filtering).

37 The new algorithm predicted the time in prolonged sitting bouts most accurate (bias ≤ 7
38 minutes/day). Of all proprietary ActiGraph methods, only 150 cpm without LFE predicted the
39 time in prolonged sitting bouts non-significantly different from the activPAL (bias ≤ 18
40 minutes/day). However, the frequently used 100 cpm with LFE accurately predicted total sitting
41 time (bias ≤ 7 minutes/day).

42 To study the health effects of ActiGraph measured prolonged sitting, we recommend using the
43 new algorithm. In case a cut-point is used, we recommend 150 cpm without LFE to measure
44 prolonged sitting, and 100 cpm with LFE to measure total sitting time. However, both cpm cut-
45 points are not recommended for a detailed bout analysis.

46 **Keywords:** activPAL, Automated Feature Selection, Bout Analysis, Machine Learning,
47 Posture Prediction, Sedentary Behaviour

48

49 Introduction

50 Sedentary Behaviour (SB, defined as sitting or reclining with ≤ 1.5 Metabolic Equivalents)² is
51 a substantial part of the modern lifestyle, accounting for the vast majority of waking hours.³
52 Research has linked SB to a plethora of serious chronic diseases and premature deaths.^{4, 5}
53 However, the largest body of evidence is based on imprecise and biased self-reports possibly
54 underestimating the strength of the relationship.^{6, 7} The technological improvements in the past
55 years made it feasible to record SB objectively. Nowadays, studies investigating SB use small
56 and lightweight body worn sensors capable to record free-living behaviour over several days.⁸
57 However, the device-based SB measure is not consistent with its definition,^{9, 10} and research is
58 far away to stipulate evidence based health recommendations.¹¹

59 Probably the most frequently used sensor to measure SB is the ActiGraph (ActiGraph LCC,
60 Pensacola, USA). The ActiGraph with its proprietary counts-per-minute (cpm) was originally
61 developed to measure physical activity.¹² As there is a growing evidence that SB, in particular
62 the time spent in prolonged bouts, is an independent risk factor for human health,¹³⁻¹⁷ ongoing
63 epidemiological studies are interested in measuring both physical activity and SB.⁸ While
64 physical activity only depends on the energy expenditure, the definition of SB includes a
65 posture component: sitting or reclining.² For this reason, it is of high value for the research
66 community to have an algorithm for the ActiGraph to predict prolonged sitting bouts. In
67 particular, those ≥ 5 and ≥ 10 minutes assumed to be most relevant for human health.¹⁷

68 To measure sitting, a pragmatic cut-point of < 100 cpm for the sensor vertical axis is most
69 frequently used,¹⁸ although there are inconsistent findings whether other cut-points, between
70 22 to 150 cpm, or machine-learning approaches like the soj3x detect sitting more accurately.^{1, 3, 19-21}
71 As the cpm measure does not consider body posture, sophisticated machine-learning
72 algorithms use the ActiGraph raw data to detect sitting.^{22, 23} However, these algorithms were

73 developed without considering feature relevance. We therefore do not know whether they
74 extract all relevant signal information to classify posture. It is very common to use extensive
75 feature lists informed by author experience or published algorithms.^{21, 24-27} Only a few studies
76 so far investigated feature relevance,²⁸ but rarely as tool for feature selection,^{27, 29} and never in
77 combination with a posture classification algorithm. Furthermore, machine-learning algorithms
78 are typically optimized to have a high sensitivity and specificity to predict posture in a certain
79 predefined window length (typically 1 minute), but not with respect to predict health-relevant
80 bout lengths.^{13, 17} Most algorithms were developed in more or less controlled laboratory
81 settings, not covering the true variability of real life.^{26, 27, 30} Moreover, many algorithm
82 developments were tailored to special population groups like breast cancer survivor or
83 overweight females.^{24, 28}

84 The aim of the present study was therefore to develop a new ActiGraph posture classification
85 algorithm to detect prolonged sitting bouts in a healthy population with sedentary occupations,
86 and to compare the new algorithm to classifications based on proprietary ActiGraph data.

87 Materials and Methods

88 Study Overview

89 The ActiGraph was calibrated against the activPAL (PAL Technologies, Glasgow, SCO) in a
90 healthy office worker population using machine-learning applied on sensor raw data collected
91 in free-living. To build the algorithm, an automated feature selection based on feature relevance
92 was used. Since poor health outcome is assumed to be related to the time spent in prolonged
93 sitting,^{13, 14, 16} a subsequent bout analysis identified the optimal feature number to predict the
94 time in bouts ≥ 5 and ≥ 10 minutes.¹⁷ Moreover, optimized cut-points for proprietary ActiGraph
95 data were developed and, together with frequently used existing cut-points and the inclinometer
96 function, included in the bout analysis.

97 Participants

98 A convenient sample of 38 participants from the GIH Brain-Health study was used.³¹ The
99 Brain-Health Study investigated the association between physical activity pattern and
100 cognition, mental health and sleep in office workers. Participants were recruited from two
101 worksites in the area of Stockholm. Office workers able to perform one week of accelerometer
102 assessment were included. Each participant signed an informed consent prior to study inclusion.
103 Ethical approval to re-use the Brain-Health data was granted by the regional ethics board (DNR
104 2018/2315-32).

105 Data Collection

106 Participants were instructed to wear an ActiGraph wGT3X-BT at the right waist (firmware
107 versions 1.9.1/1.9.2/2.5.0/3.2.1 used, 30 Hz, elastic belt) and an activPAL3 (considered as
108 reference criterion) on the right thigh (firmware 4.2.4, 20 Hz, taped), both attached as
109 recommended by the manufacturers. Participants kept a diary and noted when the ActiGraph
110 was not worn at the waist (e.g. during water based activities, sleep).

111 Data Preparation

112 Proprietary software of the sensor manufacturers were used to download sensor data and
113 generate comma separated raw data and event files for the activPAL (activPAL3, v7.2.38), as
114 well as raw data and 1-second episode files with and without low-frequency-extension (LFE)
115 filtering for the ActiGraph (ActiLife, v6.13.3). All files were load into MATLAB 2018a (v9.4,
116 Mathworks Inc., Nattick, USA). Adjacent events in the activPAL event file with the same
117 activity code were summarized and treated as single activities.³² Subsequently, the following
118 data preparation steps were carried out (for a detailed description see data processing plan in
119 Supporting Information 1): Valid recording time included all days with <95% of the time spent
120 in mode activPAL code, ≥ 500 steps and ≥ 12 hours recording. On the first/last day, valid
121 recording time was limited to the time after/before the first/last 45-second non-sedentary
122 activPAL activity. Sleep time was then removed using the Winkler algorithm (Version A)³².
123 Since the algorithm is known to underestimate sleep time,³² step tolerance was increased from
124 20 to 50 and two additional criteria using the thigh rotation angle around the longitudinal axis
125 applied.³³ Before matching the sensor data, the signals were synchronized as the sensor clocks
126 were out of sync. The offset was neither constant for all sensors nor for a single recording over
127 time. The time course of the offset between the two sensors over each recording was determined
128 by 1) finding the largest cross-correlation between the two normalized sensor x-axes of non-
129 overlapping 3 hour episodes to get the average offset of each 3 hour episode; 2) linear
130 approximation of the offset over all 3 hour episodes; 3) applying the linear approximated offset
131 to the ActiGraph time. Next, ActiGraph non-wear episodes were excluded based on the diary
132 information, sensor contradiction, and prolonged non-wear. Sensor contradiction was defined
133 as the time when the 3 dimensional ActiGraph raw signal remained constant while the activPAL
134 detected a posture change or classified the time as active (ActiGraph likely not worn).
135 Prolonged non-wear was defined as the time when the 3 dimensional ActiGraph raw signal
136 remained constant for ≥ 90 minutes. Last, to prevent excessive fragmentation of the data with
137 respect to the bout analysis, short episodes between excluded episodes were removed.

138 **Minute Extraction** – Valid minutes were extracted in two different ways, one for the algorithm
139 and cut-point development (training minutes) and one for the bout analysis (testing minutes).
140 The training minutes included only minutes with constant activPAL classifications (sitting,
141 standing, and active). All activPAL events ≥ 1 minute were identified, and as many minutes as
142 possible extracted. An event of e.g. 4.5 minutes of sitting was split in 4 single minutes, the
143 first/last minute starting/ending 15 seconds after/before the event started/ended. The testing
144 minutes were extracted according to daytime (starting at midnight) and included all available
145 minutes on days with ≥ 10 recording hours, similar as in typical epidemiological studies.^{4,5}

146 Machine Learning Algorithm Development

147 **Feature Calculation and Selection** – A total of 563 ActiGraph signal features were calculated
148 for each training minute, of which 409 in the time and 154 in the frequency domain (see feature
149 table in Supporting Information 2). Features were calculated for each sensor axis and the vector
150 magnitude, the low pass filtered sensor axes and vector magnitude (Butterworth 2nd order,
151 0.5Hz cut-off), and the 3d angle of the low pass filtered data. To identify the relevance of each
152 signal feature, a random forest classifier programmed in Python was used. The classifier run

153 100 times, and the 100 most relevant signal features were subsequently inputted into a
154 sequential forward feature selection to get the final feature ranking. A MATLAB bagged
155 classification tree ensemble (using standard properties with five bags) iteratively selected the
156 feature with highest cross-validity on the holdout subjects in each round, similar as in our
157 previous study,³⁴ until a maximum cross-validity was found. The feature selected in each round
158 was assigned to the corresponding rank.

159 **Algorithm Training** – Based on the ranking, the training properties for each feature number
160 were optimized using MATLAB’s built-in hyper-parameter optimisation function for learner
161 ensembles (fitcensemble), again using the holdout subject approach. The optimisation searched
162 for the best ensemble learner method (Bag, AdaBoost M2, RUS Boost), split criterion (gdi,
163 twoing, deviance), number of trees (10 to 500), minimum leave size (1 to $n/2$, n = number of
164 minutes), maximum number of splits (1 to $n-1$), and learning rate (0 to 1). Further details about
165 the optimisation properties can be accessed online (www.mathworks.com/help/stats/fitcensemble.html). Subsequently, 38 holdout algorithms were trained for each feature number
166 (one for each subject) and used in the bout analysis to identify the optimal feature number. A
167 detailed description on how classification trees are trained can be found elsewhere.³⁵
168

169 Optimized Cut-point Development

170 Beside the machine-learning algorithms, posture classifications based on cpm data for the
171 vertical axis and vector magnitude as well as steps-per-minute were developed, all with and
172 without LFE. The 1-second episode counts and steps were summarized for the extracted training
173 minutes, and cut-points from 0 to 5’000 to identify sitting and standing inspected. Similar as
174 for the machine-learning, the cut-points with highest cross-validity on the holdout subjects were
175 selected and used in the bout analysis to identify the most accurate one.

176 Bout Analysis

177 For each testing minute, the selected features as well as the cpm and steps-per-minute were
178 calculated. The trained holdout algorithms (machine-learning) and cut-points (proprietary
179 ActiGraph data) were then used to predict body posture of each minute. All ActiGraph
180 predictions as well as the activPAL reference criterion (the proprietary event file) were
181 subsequently aggregated in sitting and standing bouts of certain lengths for each day and
182 subject. A sitting/standing bout was defined as the time the prediction model/activPAL event
183 file classified a person continuously in sitting/standing, without the allowance of any other body
184 posture or walking. Additionally, the two most frequently used cpm cut-points, 100 and 150 for
185 the vertical axis,¹⁸ and the inclinometer function were included in the bout analysis (all with
186 and without LFE). For the inclinometer function, each testing minute was assigned to the most
187 dominant posture. Note that the proprietary activPAL event file uses another resolution (0.1
188 seconds) for the behaviour classification than the developed ActiGraph prediction models (60
189 seconds).

190 Evaluation and Statistics

191 **Data Preparation** – After rejecting the normal distribution with Lilliefors test, descriptive
192 results for data preparation are presented with median (interquartile-range).

193 **Algorithm and Cut-point Development** - To analyse cross-validity, the balanced holdout
194 sensitivity and specificity, which is the average of all sensitivities and specificities over all
195 holdout subjects, was used. For the machine-learning, the balanced sensitivity and specificity
196 was weighted according to the fraction of each behaviour in the training data. For the
197 proprietary ActiGraph data, the cut-points to detect sitting and standing were searched
198 independently. Accordingly, the balanced sensitivity and specificity was calculated for each
199 posture separately. The holdout approach (also called leave-one-subject-out) trained the
200 algorithm/cut-point on all but one subject (the holdout), and used the trained algorithm/cut-
201 point to predict the posture on the holdout subject. This procedure was repeated until every
202 subject served once as holdout, and the cross-validity was calculated among all holdout
203 predictions.

204 **Bout Analysis** - With respect to detrimental health effects of prolonged sitting,¹³⁻¹⁷ the daily
205 time spent in sitting bouts ≥ 5 min and ≥ 10 min was considered most important.¹⁷ Accordingly,
206 the algorithm and cut-point with lowest absolute bias to predict the time spent in these bouts
207 was selected. Additional bout lengths and number of bouts per day are presented to inspect the
208 prediction performance in detail. For standing, there is no evidence that certain bout lengths are
209 more relevant for health than others are. Accordingly, only total time spent standing was
210 analysed. Bias was calculated according to Bland-Altman statistics by subtracting the activPAL
211 reference criterion from the ActiGraph holdout prediction.³⁶ In case the bias depended on the
212 mean, the regression approach was used. To simplify comparison, data is in either case
213 (standard or regression approach) presented at the mean of both methods with bias and standard
214 error. Significant differences of the ActiGraph methods to the activPAL were detected using
215 the 95% confidence interval of the bias.

216 Results

217 Subjects of the present analysis were 25 men and 13 women. Mean \pm SD was 71.2 \pm 10.2 kg for
218 body mass and 42.3 \pm 8.4 years for age. Subjects wore the sensors for 9 (0) days (median with
219 inter-quartile range in brackets). Sensor offset at first valid data entry was 5.9 (8.7) seconds and
220 increased with 1.0 (1.3) seconds a day. Data preparation and minute extraction resulted in
221 200'704 training minutes (3'345 hours) and 255'569 testing minutes (4'260 hours). The posture
222 in which the time was spent is shown in Table 1.

223 **Machine Learning Algorithm** - The automated feature selection identified 26 relevant signal
224 features (maximum cross-validity), for each of which an algorithm was trained (see feature
225 ranking information in Supporting Information 2). However, the lowest absolute bias to predict
226 the sitting time in bouts ≥ 5 and ≥ 10 minutes was found for the algorithm with 14 features. This
227 algorithm combined 16 decision trees and predicted the time non-significantly different from
228 the activPAL (Table 2, absolute bias ≤ 7 minutes). The detailed bout analysis (from < 5 to ≥ 30
229 minutes, Table 2) shows that the time and number of bouts < 15 minutes was overestimated by
230 the algorithm, while longer bouts were accurately predicted. For standing, the bias was non-
231 significantly different from the activPAL (Table 2).

232 **Optimised Cut-points** - All optimised cut-points for proprietary ActiGraph data (cut-points
233 shown in Table 2) significantly underestimated the time in sitting bouts ≥ 5 and ≥ 10 minutes,
234 except steps-per-minute without LFE (accurate for bouts ≥ 10 minutes, overestimation for bouts
235 ≥ 5 minutes, Table 2). The detailed bout analysis uncovers that the time and number of short
236 bouts was generally overestimated and long bouts generally underestimated. For standing, the
237 optimised cpm cut-points for data without LFE predicted the time non-significantly different
238 from the activPAL, but the bias depended on total standing time (marked with † in Table 2).

239 **Existing Cut-points and Inclinometer Function** – The existing cut-points for proprietary
240 ActiGraph data significantly underestimated the time in the two bout lengths, except 150 cpm
241 without LFE (absolute bias ≤ 18 minutes, Table 3). However, the 100 cpm with LFE accurately
242 predicted total sitting time without consideration of a minimum bout length. The detailed bout
243 analysis shows again that short bouts were generally overestimated and long bouts generally
244 underestimated, both mostly significant (Table 3). The inclinometer function significantly
245 underestimated the time in the two bout lengths as well as total sitting and standing time.

246 Discussion

247 This study developed a new posture classification algorithm for ActiGraph raw data to predict
248 the time spent in prolonged sitting bouts as well as total standing time. The posture prediction
249 of the new algorithm does not differ from the activPAL. For sitting, the bias was $< 0.0\%$ for
250 bouts ≥ 5 minutes and -1.8% for bouts ≥ 10 minutes. For standing, the bias was -4.9% for total
251 time without consideration of a minimum bout duration. The algorithm to predict the posture
252 directly from the ActiGraph raw data file as exported by ActiLife is provided on MATLAB
253 Central File Exchange (URL is inserted provided that your journal approves the publication).

254 The study also optimised cut-points for proprietary ActiGraph data. Of these, there was only
255 one accurately predicting the time spent in sitting bouts ≥ 10 minutes: the step count with a cut-
256 point of 3 steps-per-minute (without LFE). All others substantially underestimated prolonged
257 sitting. For standing, the developed cpm cut-points without LFE accurately predicted total time
258 (vertical axis and vector magnitude). However, the longer the time spent standing the larger the
259 bias.

260 Moreover, two frequently used existing cpm cut-points were included in the bout analysis: 100
261 and 150 cpm on the vertical axis.¹⁸ While the 150 cpm without LFE accurately predicted the
262 time in prolonged sitting bouts (bias of ≤ 18 minutes or $\leq 4.6\%$), all others underestimated
263 prolonged sitting. However, 100 cpm on the vertical axis with LFE very accurately predicted
264 the total time spent sitting (bias of ≤ 7 minutes or $\leq 1.4\%$). The result for the 100 cpm with LFE
265 is in line with Matthews et al. 2018 and the overestimation of short bouts (< 20 minutes) and
266 underestimation of long bouts (≥ 30 minutes) in line with Kerr et al. 2018.^{3, 24} The results for
267 the 150 cpm to detect prolonged sitting is in line with the recommendation in Kim et al. 2015.¹
268 However, due to the significant overestimation of bouts < 25 minutes and underestimation of
269 bouts ≥ 30 minutes, a detailed bout analysis is not recommended with the 150 cpm.

270 For all cpm cut-points, there was a substantial difference between the data with and without
271 LFE, highlighting that the decision whether LFE is used or not has a great bearing, and should

272 future studies sensitize to report the use of LFE.¹⁸ Although the results of the existing cut-points
273 (Table 3) were not directly compared to the optimised cut-points for methodological reasons
274 (Table 2), it is evident that the optimised cut-points performed worse in the bout analysis despite
275 the slightly higher balanced sensitivity and specificity (see cross-validity table in Supporting
276 Information 3). The existing cut-points had far higher sensitivities (+18%) and far lower
277 specificities (-20%) to detect sitting. From this, we conclude that sensitivity and specificity is
278 not a universal measure to infer to the accuracy in the bout analysis. Future studies developing
279 new algorithms to measure prolonged sitting might therefore consider the use of other
280 optimisation criteria than balanced sensitivity and specificity, combine it as in this study with a
281 subsequent bout analysis, or weight the sensitivity more than the specificity. In our data set, a
282 weighting factor between 1.16 and 1.85 for sensitivity would have turned the best method for
283 proprietary ActiGraph data to predict total sitting time (100 cpm on the vertical axis with LFE)
284 also into the one with highest balanced sensitivity and specificity.

285 The ActiGraph inclinometer function performed worst and underestimated prolonged sitting as
286 well as total standing time by more than 2 hours a day or -32 to -54%. For total sitting time, our
287 data (bias of -21% and -22%) is in line with Kim et al 2015 who compared the inclinometer
288 function to an automated wearable camera.¹

289 Methodological Consideration

290 The machine-learning algorithm development started with an extensive feature number (563)
291 calculated for an immense amount of training data (200'704 minutes) collected in entirely free-
292 living over several days. The data was labelled with the activPAL, a well-known and highly
293 valid sensor to measure body posture that is seen as the method of choice to measure sitting in
294 free-living.^{1,20,37} Before building the algorithm, a random forest classifier in combination with
295 a sequential forward feature selection identified the most relevant signal features. Although
296 machine-learning is able to minimize the impact of non-relevant signal features on the predicted
297 output, this approach was key to end up with an algorithm having only a few features with
298 limited complexity despite the good bout performance. Our algorithm uses 14 features in
299 combination with 16 trees, while other algorithms use more than 40 features with 500 trees.^{24,}
300 ³⁵ In a general sense, an algorithm with only a few features and simple architecture is less prone
301 to overfitting and thus more likely to have a better generalizability than an algorithm with many
302 features and complex architecture, although the algorithm with many features typically
303 performs better on the training data.³⁸ In this regard, we recommend to develop algorithms with
304 as few features as necessary, and to treat each feature for each signal dimension independently
305 to ensure the algorithm performance is not reduced with non-relevant and/or redundant
306 features.³⁹ Our final algorithm e.g. uses the signal power of the sensor y-axis, but not the signal
307 power of the other two sensor axes. For this reason, we recommend to forego predefined feature
308 lists and to use an automated selection procedure. Interestingly, the algorithm uses only 2
309 features from the low-pass filtered data, but 12 features from the non-filtered raw data (see
310 Supporting Information 2). While the low-pass filtered acceleration signal reflects the waist
311 orientation versus gravity (which is often referred to as inclinometer function) that is sensitive
312 to body shape and sensor placement, the non-filtered data reflects waist movements and is less
313 sensitive to body shape and sensor placement. Accordingly, the presented algorithm primarily

314 detects the different motion pattern of the waist while sitting and standing, and not a different
315 waist orientation.

316 After optimising the training properties for each feature number, the algorithms were developed
317 with the training minutes and evaluated on the testing minutes. The clear distinction between
318 training and testing minutes further helped to limit the algorithm complexity and prevent
319 overfitting. For this reason, the algorithm with 14 features was selected, although the one with
320 26 had the highest cross-validity in the training data, again supporting our observation that a
321 high cross-validity does not imply a good bout performance. Although the two data sets
322 (training: minutes with constant activPAL classification, testing: all minutes on days with ≥ 10
323 hours) are not independent of each other as they use the same data recording and subjects, the
324 start times of the minutes were always different and thus the features not congruent. Even more
325 importantly, 28% of the testing minutes contained more than one activPAL posture
326 classification, similar as the data of a typical field study. The combination with the holdout
327 subject approach makes the algorithm to a large degree independent of the training minutes and
328 increases its generalizability. Nevertheless, a future study using the presented algorithm should
329 use exactly the same sensor settings: mounting the ActiGraph wGT3X-BT at the right waist
330 with an elastic belt and record with 30Hz. However, the study recorded data over several days,
331 and the raw data looks like the sensors were not always worn in the same way (e.g. upside
332 down). For this reason, the results of this study should not be compared to studies collecting
333 data on a daily basis in the presence of a researcher.¹ Since the ActiGraph raw data is already
334 pre-processed, we do not know whether our algorithm depends on the ActiLife software version
335 used in this study.

336 The Bland-Altman comparison used the data of each device similar as a typical field study does:
337 The proprietary activPAL event file with a resolution of 0.1 seconds, and the ActiGraph
338 predictions on a minute-by-minute level. The bout comparison is therefore questionable for
339 very short bouts (< 1 minute) as the ActiGraph might fail to detect them. However, there is some
340 evidence that prolonged bouts are health-relevant, and $> 90\%$ of the daily sitting time was spent
341 in bouts ≥ 5 minutes (activPAL data, Table 2). We therefore accepted this limitation for very
342 short bouts but were able to use the sensors exactly the way as they are used in field studies.
343 From a health perspective, we do not feel that sitting bouts < 1 minute are of critical importance.

344 Furthermore, the ActiGraph step count (without LFE) allows for 2 steps per minute although
345 the minute is still classified as sitting. This might imply that different sitting bout definitions
346 were used in this study, which was not the case. The fact that the ActiGraph records 2 steps in
347 a minute does not mean that a subject actually took 2 steps. We quite often noticed single steps
348 in a minute, even though the activPAL classified the entire minute as sitting. Accordingly, the
349 ActiGraph step count should be interpreted with caution when only a handful of steps are
350 recorded.

351 Unless the algorithm is tested in another study population than office workers, its application
352 in other populations should take place with caution. Our office worker spent 8.0 hours a day
353 sitting in 47 bouts, of which almost 50% in bouts ≥ 30 minutes (activPAL data). The female
354 breast cancer survivors in Kerr et al. 2018 spent 8.1 hours a day sitting in 49 bouts, of which

355 approximately 56% in bouts ≥ 30 minutes (activPAL data).²⁴ The NHANES 2003-2006 study
356 population in Kim et al. 2015 spent 8.0 hours a day sitting in 93 bouts, of which only 20% in
357 bouts ≥ 30 minutes (ActiGraph data with 100 cpm cut-point on the vertical axis).¹⁷ However, if
358 comparing the NHANES data to the 100 cpm in this study (without LFE), our subjects spent
359 8.4 hours a day sitting in 76 bouts, of which only 26% in bouts ≥ 30 minutes. Thus, it seems that
360 our office workers are not fundamentally different from other study populations, but we do not
361 know whether they are representative. The office workers in Keown et al. 2018 spent 9.8 hours
362 a day sitting in 49 bouts, of which 67% in bouts ≥ 30 minutes (activPAL data).⁴⁰ However, the
363 comparison to NHANES data highlights that using the 100 cpm on the vertical axis without
364 LFE is not the preferred choice to analyse the time in and number of sitting bouts. Even the 150
365 cpm does not allow such a detailed analysis.

366 This study developed an algorithm to detect prolonged sitting bouts since there is some evidence
367 that long-lasting, uninterrupted sitting might have detrimental health effects.^{13, 14, 16} To date, we
368 do not know which bout length separates detrimental sitting from non-detrimental sitting. One
369 frequently cited study reports that either 5 or 10 minutes could be a reasonable choice,
370 especially as compared to no minimum bout length.¹⁷ Unfortunately, the study used the 100
371 cpm on the vertical axis to detect sitting bouts. As can be seen in our data (Table 3), the 100
372 cpm is not appropriate to detect sitting bouts, and further research is warranted to identify what
373 separates detrimental from non-detrimental sitting. For this reason, we decided to treat bouts
374 ≥ 5 and ≥ 10 minutes equal, even though bouts longer than 10 minutes are included twice. In
375 regard of the bout length, we decided to develop a minute-based posture classification. Other
376 authors used shorter durations of e.g. 5 seconds on a very similar dataset to better handle posture
377 changes within a minute.²⁴ In our data set, 28% of all testing minutes contained at least one
378 posture change, while 72% were spent in the same posture. There is some evidence that
379 reducing the window size reduces cross-validity,²³ with unknown effects on the bout analysis.
380 However, we felt that reducing the window size is superfluous for an algorithm aiming to detect
381 prolonged sitting bouts. A shorter window size increases the computational demands that could
382 be a severe limitation for large data sets. However, the minute based approach might partially
383 explain the new algorithm's overestimation of short bouts.

384 The combined analysis of two sensors requires that they record synchronously. However, we
385 noticed a substantial offset between the two sensor clocks. The start offset could be a
386 consequence of using different clocks (i.e. computers) to initialise the sensors, and the
387 increasing offset must be a consequence of inexact sensor frequency. We could not find
388 evidence that other studies observed the same issue, but recommend future studies to inspect
389 the raw data in detail and ensure their synchronicity.

390 The feature calculation of the final algorithm does not allow to straightforwardly convert the
391 MATLAB code into a universal computer language like C++ since MATLAB specific
392 functions are used. Thus, the use of the algorithm requires a MATLAB license including two
393 toolboxes (signal processing, statistics and machine learning), all together resulting in an annual
394 or perpetual license fee of 600 or 1200 USD. Compared to available freeware, however, the
395 advantages of MATLAB clearly outweighed the disadvantages for this project. Given the costs

396 of the ActiGraph sensors for large field studies, we do not consider the license fee to be a serious
397 limitation. The final algorithm published on MATLAB Central File Exchange (URL is inserted
398 provided that your journal approves the publication) directly predicts the posture from the
399 ActiGraph raw data csv-file and creates a new csv-file in the same data format. The algorithm
400 can be used even without previous MATLAB experience.

401 Practical Implications

402 The results of this study show that there is no single ActiGraph method accurately predicting
403 the sitting time in certain bout length as well as total sitting time. We therefore recommend
404 future studies to choose a method depending on the study aim. To analyse the total sitting time
405 regardless of a minimum bout duration, our recommendation is to use the 100 cpm cut-point
406 with LFE. To analyse prolonged sitting, our recommendation is to use the developed machine-
407 learning algorithm or the 150 cpm without LFE. The machine-learning algorithm is the most
408 accurate choice, and allows for a very detailed analysis of bouts ≥ 15 minutes (time in and
409 number of bouts) that should be avoided with the 150 cpm. Moreover, the algorithm includes
410 an accurate total standing time prediction. For the cut-point methods, the study highlights that
411 the decision whether LFE is used or not is of utmost importance and should be explicitly
412 reported. Regarding the future algorithm development to detect prolonged sitting, we
413 recommend considering also other optimisation criteria than sensitivity and specificity with
414 respect to an accurate bout prediction. The present study analysed the classification capability
415 of the ActiGraph GT3X to detect prolonged sitting, which should not be equated with SB. For
416 SB, the sitting classification must be combined with an activity classification and other cut-
417 points than those investigated in this study might solve the SB classification better.

418 Perspective

419 To study the health effects of ActiGraph measured prolonged sitting, we recommend using the
420 new algorithm available on MATLAB Central File Exchange. In case a cpm cut-point should
421 be use, the 150 cpm without LFE is the best choice. To analyse total sitting time without
422 consideration of a minimum bout length, the 100 cpm cut-point is the most appropriate choice
423 only in combination with LFE data. However, we do not recommend using the cpm cut-points
424 for a detailed sitting bout analysis. Further research is warranted to validate the new algorithm
425 in an independent sample and different population.

426 References

- 427 1. Kim Y, Barry VW, Kang M. Validation of the actigraph gt3x and activpal accelerometers for the assessment of sedentary
428 behavior. *Meas Phys Educ Exerc Sci* 2015; 19:125-137
- 429 2. SBRN. Letter to the editor: Standardized use of the terms "sedentary" and "sedentary behaviours". *Applied Physiology,*
430 *Nutrition, and Metabolism* 2012; 37:540-542
- 431 3. Matthews CE, Kozey-Keadle S, Moore SC, Schoeller DS, Carroll RJ, Troiano RP, Sampson JN. Measurement of active and
432 sedentary behavior in context of large epidemiologic studies. *Med Sci Sports Exerc* 2018; 50:266-276
- 433 4. Amirfaiz S, Shahril MR. Objectively measured physical activity, sedentary behavior, and metabolic syndrome in adults:
434 Systematic review of observational evidence. *Metab Syndr Relat Disord* 2019; 17:1-21

- 435 5. Dohrn IM, Sjoström M, Kwak L, Oja P, Hagström M. Accelerometer-measured sedentary time and physical activity-a
436 15 year follow-up of mortality in a Swedish population-based cohort. *J Sci Med Sport* 2018; 21:702-707
- 437 6. Chastin SFM, Dontje ML, Skelton DA, Cukic I, Shaw RJ, Gill JMR, Greig CA, Gale CR, Deary IJ, Der G, Dall PM, Seniors
438 USPT. Systematic comparative validation of self-report measures of sedentary time against an objective measure of
439 postural sitting (activPAL). *Int J Behav Nutr Phys Act* 2018; 15:21
- 440 7. de Rezende LF, Rodrigues Lopes M, Rey-Lopez JP, Matsudo VK, Luiz Odo C. Sedentary behavior and health outcomes:
441 An overview of systematic reviews. *PLoS One* 2014; 9:e105620
- 442 8. Lee IM, Shiroma EJ. Using accelerometers to measure physical activity in large-scale epidemiological studies: Issues and
443 challenges. *Br J Sports Med* 2014; 48:197-201
- 444 9. Holtermann A, Schellewald V, Mathiassen SE, Gupta N, Pinder A, Punakallio A, Veiersted KB, Weber B, Takala EP,
445 Draicchio F, Enquist H, Desbrosses K, Garcia Sanz MP, Malinska M, Villar M, Wichtl M, Strelb M, Forsman M, Lusa S,
446 Tokarski T, Hendriksen P, Ellegast R. A practical guidance for assessments of sedentary behavior at work: A perosh
447 initiative. *Appl Ergon* 2017; 63:41-52
- 448 10. Kang M, Rowe DA. Issues and challenges in sedentary behavior measurement. *Meas Phys Educ Exerc Sci* 2015; 19:105-
449 115
- 450 11. Stamatakis E, Ekelund U, Ding D, Hamer M, Bauman AE, Lee IM. Is the time right for quantitative public health guidelines
451 on sitting? A narrative review of sedentary behaviour research paradigms and findings. *Br J Sports Med* 2019; 53:377-
452 382
- 453 12. Freedson PS, Melanson E, Sirard J. Calibration of the computer science and applications, inc. Accelerometer. *Med Sci*
454 *Sports Exerc* 1998; 30:777-781
- 455 13. Bellettiere J, Winkler EAH, Chastin SFM, Kerr J, Owen N, Dunstan DW, Healy GN. Associations of sitting accumulation
456 patterns with cardio-metabolic risk biomarkers in Australian adults. *PLoS One* 2017; 12:e0180119
- 457 14. Benatti FB, Ried-Larsen M. The effects of breaking up prolonged sitting time: A review of experimental studies. *Med Sci*
458 *Sports Exerc* 2015; 47:2053-2061
- 459 15. Buckley JP, Hedge A, Yates T, Copeland RJ, Loosemore M, Hamer M, Bradley G, Dunstan DW. The sedentary office: An
460 expert statement on the growing case for change towards better health and productivity. *Br J Sports Med* 2015;
461 49:1357-1362
- 462 16. Healy GN, Dunstan DW, Salmon J, Cerin E, Shaw JE, Zimmet PZ, Owen N. Breaks in sedentary time: Beneficial
463 associations with metabolic risk. *Diabetes Care* 2008; 31:661-666
- 464 17. Kim Y, Welk GJ, Braun SI, Kang M. Extracting objective estimates of sedentary behavior from accelerometer data:
465 Measurement considerations for surveillance and research applications. *PLoS One* 2015; 10:e0118078
- 466 18. Migueles JH, Cadenas-Sanchez C, Ekelund U, Delisle Nystrom C, Mora-Gonzalez J, Lof M, Labayen I, Ruiz JR, Ortega FB.
467 Accelerometer data collection and processing criteria to assess physical activity and other outcomes: A systematic
468 review and practical considerations. *Sports Med* 2017; 47:1821-1845
- 469 19. Clarke-Cornwell AM, Farragher TM, Cook PA, Granat MH. Empirically derived cut-points for sedentary behaviour: Are
470 we sitting differently? *Physiol Meas* 2016; 37:1669-1685
- 471 20. Kozey-Keadle S, Libertine A, Lyden K, Staudenmayer J, Freedson PS. Validation of wearable monitors for assessing
472 sedentary behavior. *Med Sci Sports Exerc* 2011; 43:1561-1567
- 473 21. Lyden K, Keadle SK, Staudenmayer J, Freedson PS. A method to estimate free-living active and sedentary behavior from
474 an accelerometer. *Med Sci Sports Exerc* 2014; 46:386-397
- 475 22. de Almeida Mendes M, da Silva ICM, Ramires VV, Reichert FF, Martins RC, Tomasi E. Calibration of raw accelerometer
476 data to measure physical activity: A systematic review. *Gait Posture* 2018; 61:98-110
- 477 23. Farrahi V, Niemela M, Kangas M, Korpelainen R, Jamsa T. Calibration and validation of accelerometer-based activity
478 monitors: A systematic review of machine-learning approaches. *Gait Posture* 2019; 68:285-299
- 479 24. Kerr J, Carlson J, Godbole S, Cadmus-Bertram L, Bellettiere J, Hartman S. Improving hip-worn accelerometer estimates
480 of sitting using machine learning methods. *Med Sci Sports Exerc* 2018; 50:1518-1524
- 481 25. Liu S, Gao RX, Freedson PS. Computational methods for estimating energy expenditure in human physical activities.
482 *Med Sci Sports Exerc* 2012; 44:2138-2146

483 26. Payey TG, Gilson ND, Gomersall SR, Clark B, Trost SG. Field evaluation of a random forest activity classifier for wrist-
484 worn accelerometer data. *J Sci Med Sport* 2017; 20:75-80

485 27. Zhang S, Rowlands AV, Murray P, Hurst TL. Physical activity classification using the genea wrist-worn accelerometer.
486 *Med Sci Sports Exerc* 2012; 44:742-748

487 28. Ellis K, Kerr J, Godbole S, Staudenmayer J, Lanckriet G. Hip and wrist accelerometer algorithms for free-living behavior
488 classification. *Med Sci Sports Exerc* 2016; 48:933-940

489 29. Chowdhury AK, Tjondronegoro D, Chandran V, Trost SG. Ensemble methods for classification of physical activities from
490 wrist accelerometry. *Med Sci Sports Exerc* 2017; 49:1965-1973

491 30. Bastian T, Maire A, Dugas J, Ataya A, Villars C, Gris F, Perrin E, Caritu Y, Doron M, Blanc S, Jallon P, Simon C. Automatic
492 identification of physical activity types and sedentary behaviors from triaxial accelerometer: Laboratory-based
493 calibrations are not enough. *J Appl Physiol* (1985) 2015; 118:716-722

494 31. Pantzar A, Jonasson LS, Ekblom O, Boraxbekk CJ, Ekblom MM. Relationships between aerobic fitness levels and cognitive
495 performance in swedish office workers. *Front Psychol* 2018; 9:2612

496 32. Winkler EA, Bodicoat DH, Healy GN, Bakrania K, Yates T, Owen N, Dunstan DW, Edwardson CL. Identifying adults' valid
497 waking wear time by automated estimation in activpal data collected with a 24 h wear protocol. *Physiol Meas* 2016;
498 37:1653-1668

499 33. Lyden K, John D, Dall P, Granat MH. Differentiating sitting and lying using a thigh-worn accelerometer. *Med Sci Sports
500 Exerc* 2016; 48:742-747

501 34. Kuster R, Huber M, Hirschi S, Siegl W, Baumgartner D, Hagströmer M, Grooten W. Measuring sedentary behavior by
502 means of muscular activity and accelerometry. *Sensors (Basel)* 2018; 18

503 35. Ellis K, Kerr J, Godbole S, Lanckriet G, Wing D, Marshall S. A random forest classifier for the prediction of energy
504 expenditure and type of physical activity from wrist and hip accelerometers. *Physiol Meas* 2014; 35:2191-2203

505 36. Bland J, Altman D. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999; 8:135-160

506 37. Lyden K, Kozey Keadle SL, Staudenmayer JW, Freedson PS. Validity of two wearable monitors to estimate breaks from
507 sedentary time. *Med Sci Sports Exerc* 2012; 44:2243-2252

508 38. Kate RJ, Swartz AM, Welch WA, Strath SJ. Comparative evaluation of features and techniques for identifying activity
509 type and estimating energy cost from accelerometer data. *Physiol Meas* 2016; 37:360-379

510 39. Liu S, Gao RX, John D, Staudenmayer JW, Freedson PS. Multisensor data fusion for physical activity assessment. *IEEE
511 Trans Biomed Eng* 2012; 59:687-696

512 40. Keown MK, Skeaff CM, Perry TL, Haszard JJ, Peddie MC. Device-measured sedentary behavior patterns in office-based
513 university employees. *J Occup Environ Med* 2018; 60:1150-1157

514

515 Tables

516 Table 1: Overview of the recorded time, data preparation, and time used for the algorithm/cut-point development (training) and the bout analysis (testing data). Absolute time in hours per subject except sleep (hours per night), relative time in percentage of total time per subject. Indicated is the median with interquartile-range (iqr).

	Absolute Time [hours/subject]		Relative Time [%]	
	Median (iqr)	Range	Median (iqr)	Range
Valid Recording Time	200.7 (7.3)	[81.3 - 224.2]		
- Sleep	8.5 (1.3)	[6.9 - 10.6]		
- ActiGraph Non-Wear	8.2 (11.5)	[2.1 - 39.5]		
- Short Episode	0.9 (1.1)	[0.0 - 3.6]		
Remaining Time	121.3 (15.6)	[37.1 - 149.6]		
Training Data	90.2 (15.1)	[30.3 - 111.7]		
- Sitting	62.2 (14.9)	[22.8 - 94.3]	72.6 (15.3)	[49.8 - 91.3]
- Standing	17.3 (11.8)	[5.8 - 37.7]	21.0 (10.9)	[6.0 - 38.8]
- Active	6.5 (5.5)	[1.5 - 13.1]	7.2 (5.6)	[2.7 - 15.5]
Testing Data	114.3 (24.6)	[31.5 - 149.4]		
- Sitting	60.7 (19.5)	[20.9 - 98.7]	55.4 (13.4)	[38.4 - 78.2]
- Standing	36.3 (14.1)	[8.0 - 58.8]	31.5 (10.9)	[13.8 - 44.7]
- Active	15.2 (8.3)	[2.6 - 26.4]	13.6 (4.9)	[8.0 - 21.7]

The training data contains only minutes with constant activPAL classification, the testing data contains all minutes on days with ≥ 10 hours.
Abbreviations: interquartile-range (iqr)

Table 2: Bias of the machine-learning algorithm and the optimised cut-points for proprietary ActiGraph data to the activPAL (reference criterion). Indicated is the mean \pm standard error for the reference criterion, and bias \pm standard error for the ActiGraph methods. Time in minutes per day.

	Reference Criterion	Machine Learning Algorithm	Y_{cpm}	$Y_{cpm(LFE)}$	VM_{cpm}	$VM_{cpm(LFE)}$	Step	$Step_{LFE}$
Sitting			< 16 cpm	< 23 cpm	< 69 cpm	< 170 cpm	< 3 spm	< 5 spm
Time in Bout								
- ≥ 5	441.2 \pm 12.7	0.2 \pm 6.6	-185.6 \pm 13.6 *	-190.2 \pm 13.9 *	-168.9 \pm 14.9 *	-126.0 \pm 15.4 *	38.5 \pm 11.6 *	-101.5 \pm 13.8 *
- ≥ 10	397.4 \pm 12.6	-7.1 \pm 7.4	-234.9 \pm 13.1 *	-237.7 \pm 13.1 *	-223.2 \pm 14.5 *	-176.9 \pm 15.6 *	6.5 \pm 11.9	-132.2 \pm 13.2 *
-total	481.7 \pm 12.5	17.6 \pm 6.7 *	-98.7 \pm 12.5 *	-109.9 \pm 13.5 *	-86.4 \pm 13.1 *	-57.2 \pm 13.9 *	80.4 \pm 11.5 *	-55.4 \pm 14.1 *
-<5	40.4 \pm 2.1	17.5 \pm 2.7 *	86.9 \pm 3.6 * (†)	80.4 \pm 3.5 * (†)	82.5 \pm 3.4 * (†)	68.8 \pm 3.2 * (†)	42.0 \pm 2.3 * (†)	46.1 \pm 2.8 *
-5-9	43.8 \pm 1.8	7.3 \pm 1.9 *	49.3 \pm 3.5 * (†)	47.5 \pm 3.6 * (†)	54.3 \pm 3.3 * (†)	50.9 \pm 3.0 * (†)	32.0 \pm 2.2 * (†)	30.7 \pm 3.0 * (†)
-10-14	42.2 \pm 1.6	5.2 \pm 1.5 *	11.8 \pm 2.9 * (†)	10.6 \pm 3.0 * (†)	14.9 \pm 3.1 * (†)	23.3 \pm 2.8 * (†)	21.0 \pm 2.1 * (†)	13.9 \pm 2.2 * (†)
-15-19	43.9 \pm 2.0	-1.6 \pm 2.3	-10.1 \pm 2.7 *	-11.3 \pm 2.8 *	-6.7 \pm 2.9 * (†)	-1.0 \pm 3.1 (†)	9.3 \pm 2.5 *	3.3 \pm 2.7
-20-24	37.8 \pm 2.3	0.2 \pm 2.0	-18.9 \pm 3.3 *	-18.3 \pm 3.3 *	-17.7 \pm 3.5 *	-9.8 \pm 3.6 *	5.7 \pm 2.4 *	-3.2 \pm 2.8
-25-29	37.7 \pm 2.2	-3.1 \pm 1.7	-23.7 \pm 2.3 * (†)	-23.6 \pm 2.3 * (†)	-21.3 \pm 2.5 *	-18.4 \pm 2.9 *	0.7 \pm 2.0	-12.2 \pm 2.5 *
- ≥ 30	235.9 \pm 11	-8.0 \pm 7.7	-194.0 \pm 7.9 * (†)	-195.1 \pm 7.9 * (†)	-192.4 \pm 8.0 * (†)	-171.0 \pm 10.2 *	-30.3 \pm 10.3 *	-134.0 \pm 10.6 *
Number of Bouts								
- ≥ 5	19.4 \pm 0.5	2.0 \pm 0.3 *	4.0 \pm 0.9 * (†)	3.5 \pm 1.0 * (†)	5.3 \pm 0.9 * (†)	6.7 \pm 0.8 * (†)	8.1 \pm 0.6 *	4.2 \pm 0.8 * (†)
- ≥ 10	13.4 \pm 0.4	0.4 \pm 0.2	-4.4 \pm 0.6 * (†)	-4.6 \pm 0.6 * (†)	-3.8 \pm 0.6 * (†)	-1.7 \pm 0.6 * (†)	2.7 \pm 0.4 *	-1.1 \pm 0.5 *
-total	46.7 \pm 1.7	8.4 \pm 2.2 *	46.4 \pm 2.9 * (†)	42.9 \pm 2.8 * (†)	45.0 \pm 2.7 * (†)	38.3 \pm 2.4 * (†)	26.3 \pm 1.7 *	25.1 \pm 2.1 *
-<5	27.3 \pm 1.5	6.4 \pm 2.0 *	42.4 \pm 2.5 * (†)	39.5 \pm 2.4 * (†)	39.7 \pm 2.4 * (†)	31.6 \pm 2.2 * (†)	18.2 \pm 1.4 *	20.9 \pm 1.7 *
-5-9	6.0 \pm 0.3	1.6 \pm 0.3 *	8.4 \pm 0.5 * (†)	8.0 \pm 0.5 * (†)	9.0 \pm 0.5 * (†)	8.4 \pm 0.4 * (†)	5.4 \pm 0.3 * (†)	5.3 \pm 0.4 * (†)
-10-14	3.4 \pm 0.1	0.6 \pm 0.1 *	1.2 \pm 0.2 * (†)	1.1 \pm 0.2 * (†)	1.5 \pm 0.2 * (†)	2.2 \pm 0.2 * (†)	1.9 \pm 0.2 * (†)	1.4 \pm 0.2 * (†)
-15-19	2.5 \pm 0.1	0.0 \pm 0.1	-0.5 \pm 0.2 *	-0.6 \pm 0.2 *	-0.3 \pm 0.2 (†)	0.0 \pm 0.2 (†)	0.6 \pm 0.1 *	0.3 \pm 0.2
-20-24	1.7 \pm 0.1	0.1 \pm 0.1	-0.8 \pm 0.2 *	-0.8 \pm 0.1 *	-0.8 \pm 0.2 *	-0.4 \pm 0.2 *	0.3 \pm 0.1 *	-0.1 \pm 0.1
-25-29	1.4 \pm 0.1	-0.1 \pm 0.1	-0.9 \pm 0.1 * (†)	-0.9 \pm 0.1 * (†)	-0.8 \pm 0.1 *	-0.7 \pm 0.1 *	0.0 \pm 0.1	-0.4 \pm 0.1 *
- ≥ 30	4.4 \pm 0.2	-0.1 \pm 0.1	-3.4 \pm 0.2 * (†)	-3.4 \pm 0.2 * (†)	-3.4 \pm 0.2 * (†)	-2.9 \pm 0.2 *	-0.3 \pm 0.2	-2.2 \pm 0.2 *
Standing			< 403 cpm	< 398 cpm	< 1379 cpm	< 1484 cpm	< 11 spm	< 42 spm
Time in Bout								
-total	261.5 \pm 10.4	-12.7 \pm 6.4	-13.5 \pm 9.8 (†)	-23.8 \pm 11.5 * (†)	-6.5 \pm 10.7 (†)	-44.9 \pm 12.7 *	-141.8 \pm 7.5 * (†)	60.8 \pm 13.4 *

Positive bias indicates an overestimation, negative an underestimation. Significant differences to the reference criterion marked with *, biases depending on the time/number in bout indicated with † (regression approach used to calculate bias).

Abbreviations: vertical sensor axis (y), counts-per-minute (cpm), steps-per-minute (spm), low-frequency-extension filtering (LFE), vector magnitude (VM).

Table 3: Bias of the existing ActiGraph methods (counts-per-minute (cpm) and inclinometer function) to the activPAL (reference criterion). Indicated is the mean \pm standard error for the reference criterion, and bias \pm standard error for the ActiGraph methods. Time in minutes per day.

	Reference Criterion	Y_{cpm}	$Y_{cpm(LFE)}$	Y_{cpm}	$Y_{cpm(LFE)}$	Inclinometer	Inclinometer _{LFE}
Sitting		< 100 cpm	< 100 cpm	< 150 cpm	< 150 cpm		
Time in Bout							
- \geq 5	441.2 \pm 12.7	-24.4 \pm 10.6 *	-59.4 \pm 11.5 *	14.5 \pm 9.8	-12.9 \pm 10.4	-148.3 \pm 20.2 *	-143.3 \pm 20 *
- \geq 10	397.4 \pm 12.6	-67 \pm 11.6 *	-105.3 \pm 12.4 *	-18.4 \pm 10.2	-49.6 \pm 11.0 *	-180.3 \pm 19.5 *	-175.5 \pm 19.7 *
- total	481.7 \pm 12.5	23.3 \pm 10.7 *	-6.8 \pm 11.1	54.4 \pm 10.5 *	28.4 \pm 10.6 *	-105.8 \pm 19.6 *	-100.4 \pm 19.2 *
-<5	40.4 \pm 2.1	47.7 \pm 2.7 * (†)	52.6 \pm 2.7 * (†)	39.9 \pm 2.7 *	41.2 \pm 2.7 *	42.6 \pm 3.4 * (†)	43 \pm 3.4 * (†)
-5-9	43.8 \pm 1.8	42.6 \pm 2.5 * (†)	45.9 \pm 2.5 * (†)	32.9 \pm 2.3 * (†)	36.8 \pm 2.3 * (†)	32 \pm 2.6 * (†)	32.2 \pm 2.6 *
-10-14	42.2 \pm 1.6	28.6 \pm 1.8 * (†)	28.6 \pm 2 * (†)	23.3 \pm 2.0 * (†)	25.0 \pm 1.7 * (†)	10.7 \pm 2.5 * (†)	10.8 \pm 2.6 * (†)
-15-19	43.9 \pm 2.0	12 \pm 2.7 * (†)	6.9 \pm 2.8 *	13.3 \pm 3.0 * (†)	11.8 \pm 3.1 *	-7.1 \pm 3.2 *	-5.8 \pm 3.1
-20-24	37.8 \pm 2.3	2.8 \pm 2.6	-1.1 \pm 3.2	5.6 \pm 2.1 *	4.5 \pm 2.2 *	-10.6 \pm 2.9 *	-10.7 \pm 2.7 *
-25-29	37.7 \pm 2.2	-4.4 \pm 2.9	-10.4 \pm 2.8 *	0.3 \pm 2.4	-2.9 \pm 2.7	-18.8 \pm 2.9 *	-18.6 \pm 2.9 *
- \geq 30	235.9 \pm 11	-106 \pm 10.1 *	-129.3 \pm 9.8 *	-60.8 \pm 10.0 *	-88.0 \pm 9.6 *	-154.6 \pm 14.1 *	-151.3 \pm 14.4 *
Number of Bouts							
- \geq 5	19.4 \pm 0.5	8.8 \pm 0.6 * (†)	8.3 \pm 0.6 * (†)	8.2 \pm 0.6 *	8.2 \pm 0.6 *	2.6 \pm 0.8 * (†)	2.7 \pm 0.8 * (†)
- \geq 10	13.4 \pm 0.4	1.8 \pm 0.4 *	0.7 \pm 0.5	2.6 \pm 0.4 *	2.0 \pm 0.4 *	-3 \pm 0.6 * (†)	-2.8 \pm 0.6 * (†)
- total	46.7 \pm 1.7	29.2 \pm 2.1 *	30.8 \pm 2.2 *	24.8 \pm 2.0 *	25.1 \pm 2.0 *	18.7 \pm 2.4 *	18.9 \pm 2.4 *
-<5	27.3 \pm 1.5	20.4 \pm 1.8 *	22.5 \pm 1.9 *	16.6 \pm 1.7 *	17.0 \pm 1.6 *	16.2 \pm 2.3 *	16.2 \pm 2.4 *
-5-9	6.0 \pm 0.3	7 \pm 0.3 * (†)	7.6 \pm 0.4 * (†)	5.5 \pm 0.3 * (†)	6.2 \pm 0.3 * (†)	5.5 \pm 0.4 * (†)	5.6 \pm 0.4 *
-10-14	3.4 \pm 0.1	2.6 \pm 0.2 * (†)	2.6 \pm 0.2 * (†)	2.1 \pm 0.2 * (†)	2.3 \pm 0.1 * (†)	1.1 \pm 0.2 * (†)	1.1 \pm 0.2 * (†)
-15-19	2.5 \pm 0.1	0.8 \pm 0.2 * (†)	0.5 \pm 0.2 *	0.9 \pm 0.2 * (†)	0.8 \pm 0.2 *	-0.3 \pm 0.2	-0.3 \pm 0.2
-20-24	1.7 \pm 0.1	0.2 \pm 0.1	0 \pm 0.1	0.3 \pm 0.1 *	0.2 \pm 0.1 *	-0.4 \pm 0.1 *	-0.4 \pm 0.1 *
-25-29	1.4 \pm 0.1	-0.1 \pm 0.1	-0.4 \pm 0.1 *	0.0 \pm 0.1	-0.1 \pm 0.1	-0.7 \pm 0.1 *	-0.7 \pm 0.1 *
- \geq 30	4.4 \pm 0.2	-1.6 \pm 0.2 *	-2 \pm 0.2 *	-0.7 \pm 0.2 *	-1.2 \pm 0.2 *	-2.6 \pm 0.2 *	-2.5 \pm 0.3 *
Standing							
Time in Bout							
- total	261.5 \pm 10.4	-	-	-	-	-140.8 \pm 19.1 * (†)	-138.6 \pm 19.2 * (†)

Positive bias indicates an overestimation, negative an underestimation. Significant differences to the reference criterion marked with *, biases depending on the time/number in bout indicated with † (regression approach used to calculate bias).

Abbreviations: vertical sensor axis (y), counts-per-minute (cpm), low-frequency-extension filtering (LFE).

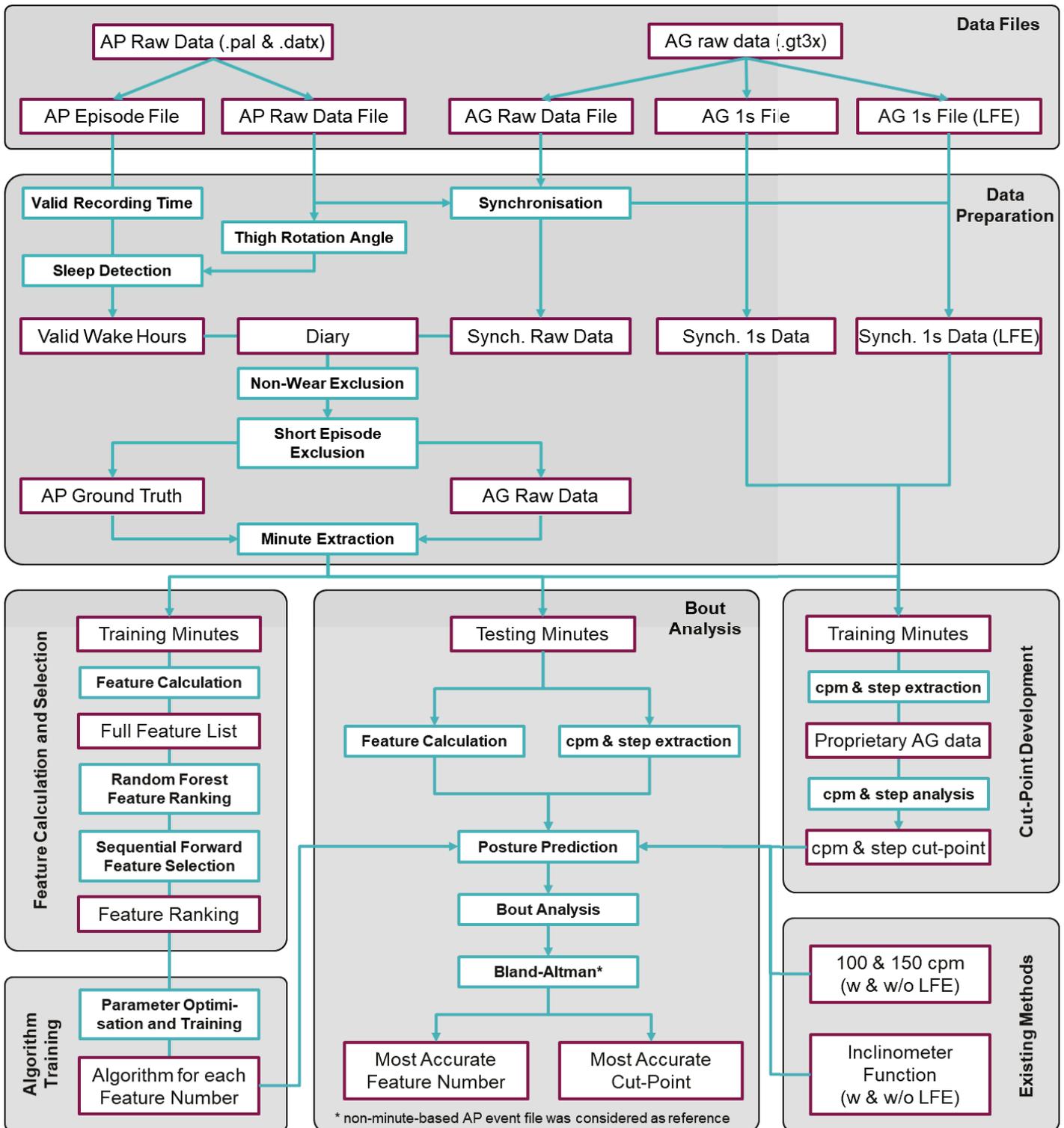
519 **List of Supporting Information Files**

520 Supporting Information 1. Data processing plan including detailed data preparation
521 description (.pdf)

522 Supporting Information 2. Feature table with ranking information and MATLAB code on how
523 to calculate the features (.pdf)

524 Supporting Information 3. Cross-validity table of all presented methods (.pdf)

Supporting Information 1: Data processing plan for acivPAL (AP) and ActiGraph (AG) data, including detailed data preparation description (bottom).



Valid Recording Time
A valid day consists of (criteria applied to activPAL (AP) data):

- <95% of day spent in mode AP Code
- ≥500 steps
- ≥12 hours recorded

On first/last valid day:

- recording start/stop defined by first/last 45-second non-Sedentary AP activity

Thigh Rotation Angle
Orientation of the thigh along its longitudinal axis versus vertical room axis (Lyden et al. 2016, 33; used to detect sleep)

- Sitting/Standing ≈ 0°
- Lying on the side ≈ ± 90°
- Lying on the stomach ≈ 180°

Sleep Detection
Longest AP Sedentary Bout from noon to noon, expanded if surrounding 15 minute window contains:

- AP Sedentary ≥2 hours
- AP Sedentary ≥0.5 hours & ≤50 steps
- AP Standing between Sleeping and Sedentary with 0 steps
- AP Sedentary with thigh rotation >65° & ≤50 steps
- AP Sedentary ≥15 min with thigh rotation >65° & ≤100 steps (applied to surrounding 60 min)

1-3) Winkler et al. 2016 (32); 4-5) Lyden et al. 2016 (33)

Synchronisation
Find largest cross-correlation between normalized sensor x-axes of non-overlapping 3 hour bouts, maximum 2 minute lag. Delay linear approximated and applied to ActiGraph (AG) time.

Non-Wear Exclusion
Each AP episode overlapping an AG episodes ≥1 second with constant AG signal on all sensor axes excluded if:

- AP recorded posture change
- AP classified the episode as Active
- AG episode ≥90 min

Self-Reported: Time excluded if participants reported the AG was not worn (diary).

Short Episode Exclusion
Time between two excluded episodes if:

- <5 min
- <10 min & both excluded episodes ≥2 min
- <60 min & shorter than both excluded episodes, each ≥10 min

Supporting Information 2 – Table 1: Table of all features including ranks for the top 100. From all 563 features, the 100 most relevant ones (identified by the random forest classifier) are indicated with the rank of the sequential forward feature selection. The final algorithm uses the 14 top ranked features (rank marked in bolt). Of these, 4 were selected from the vector magnitude and z-axis, respectively, 3 from the x-axis, 2 from the y-axis, and 1 from the dynamic time warping between x- and y-axis. Most features are based on the raw data (12) and 2 on the filtered data. No feature based on the 3d-angle was included in the final algorithm.

Features	raw data							filtered data				filtered angles			time	usage count
	x	y	z	VM	xy	xz	yz	x	y	z	VM	x	y	z		
Time Domain																
1 st Percentile				88						66				17		3
5 th Percentile				39						76				91		3
10 th Percentile			52	93						62				70		4
25 th Percentile			43	98						18						3
50 th Percentile (Median)			29											51		2
75 th Percentile				83												1
90 th Percentile				97												1
95 th Percentile				25												1
99 th Percentile				59												1
Inter-quartile range				36												1
Minimum				92						2						1 2
Maximum				19												1
Range		32		100												2
Mean			5							67				99		1 3
Standard Deviation (SD)				82												1
Coefficient of Variation (CV)				46												1
Skewness				49												1
Kurtosis				4												1 1
Summed absolute Signal Change from Frame to Frame	27	64	22	35				26	47	38	33					8
Lag 1 Frame Autocorrelation											61					1
Lag 1 Second Autocorrelation																0
3 rd Moment				1												1 1
4 th Moment				40												1
Number of Peaks											42					1
Number of Prominent Peaks	10	60	54	50				65		23						1 6
entropy				95												1
Number of Zero-Crossings																0
Mean Time between adjacent Zero-Crossings																0
Median Time between adjacent Zero-Crossings																0
SD of the Time between adjacent Zero-Crossings																0
Number of Median-Crossings										31						1
Mean Time between adjacent Median-Crossings																0
Median Time between adjacent Median-Crossings																0
SD of Time between adjacent MedianCrossings																0
Dynamic Time Warping (DTW) between Axes				3												1 1
DTW between 1 st Derivative of the axes				20	86	37										3
Covariance between axes				79												1
Correlation between axes				24												1
Daytime															21	1
SD of all non-overlapping 5 Seconds Mean																0
SD of all non-overlapping 5 Seconds CV				85												1
Frequency Domain																
Mean Frequency		78	15	96						16						4
Power at Mean Frequency ±0.1Hz		63	73	57				68	11			58	45			1 7
Median Frequency				44												1
Power at Median Frequency ±0.1Hz		80	55					90	53			30	48			6
Mean Frequency between 0.3 to 3Hz		56														1
Power at Mean Frequency ±0.1 Hz between 0.3 to 3Hz																0
Median Frequency between 0.3 to 3Hz		28								41						2
Power at Median Frequency ±0.1Hz between 0.3 to 3 Hz																0
Total Signal Power		9	71					77	89			74	84			1 6
Power below 0.3 Hz		94	75					69	87			72				5
Power between 0.3 and 3 Hz		13	12	6												3 3
Power above 3 Hz		8	34	7	14											3 4
Harmonic Power				81												1
Harmonic Frequency																
Usage Count																
top 14 (final algorithm)	3	2	2	4	1				2							14
top 100	4	12	12	28	4	1	1	1	7	10	6	1	4	8	1	100

Supporting Information 2 – Table 2: Instructions and MATLAB code to calculate the signal features. * marks features for which NaN and ±Inf were replaced with zero.

Dimensions	Instructions / MATLAB Code
rawdata: RAWDATA(:,1:3)	x, y, and z, as recorded
vector magnitude: RAWDATA(:,4)	= sqrt(RAWDATA(:,1).^2+RAWDATA(:,2).^2+RAWDATA(:,3).^2)
filtered data: RAWDATA(:,5:8)	= filter(b,a, RAWDATA(:,1:4)); with CutoffFreq = 0.5; sampfreq = 30; [b,a] = butter(2,CutoffFreq / (sampfreq/2));
filtered angle x: [~,RAWDATA(:,9),~]	= cart2sph(RAWDATA(:,6),RAWDATA(:,7),RAWDATA(:,5));
filtered angle y: [~,RAWDATA(:,10),~]	= cart2sph(RAWDATA(:,7),RAWDATA(:,5),RAWDATA(:,6));
filtered angle z: [~,RAWDATA(:,11),~]	= cart2sph(RAWDATA(:,5),RAWDATA(:,6),RAWDATA(:,7));
Minute Data	
Start frame of each minute (frameID)	= 1:1800:(NumberOfMinutes-1)*1800;
Data of each Minute (MinData)	= RAWDATA(minuteID:minuteID+1799,dimension) % for dimension = 1:11;
# Features	
Time Domain	
11 1 st Percentile	prctile(MinData,1);
11 5 th Percentile	prctile(MinData,5);
11 10 th Percentile	prctile(MinData,10);
11 25 th Percentile	prctile(MinData,25);
11 50 th Percentile (Median)	prctile(MinData,50);
11 75 th Percentile	prctile(MinData,75);
11 90 th Percentile	prctile(MinData,90);
11 95 th Percentile	prctile(MinData,95);
11 99 th Percentile	prctile(MinData,99);
11 Inter-quartile range	iqr(MinData)
11 Minimum	min(MinData);
11 Maximum	max(MinData);
11 Range	max(MinData) - min(MinData);
11 Mean	nanmean(MinData);
11 Standard Deviation (SD)	nanstd(MinData);
11 Coefficient of Variation (CV) *	nanstd(MinData)/nanmean(MinData);
11 Skewness *	skewness(MinData);
11 Kurtosis *	kurtosis(MinData);
11 Summed absolute Signal Change from Frame to Frame	sum(abs(diff(MinData)));
11 Lag 1 Frame Autocorrelation *	lag = autocorr(MinData,sampfreq); lag(2);
11 Lag 1 Second Autocorrelation *	lag = autocorr(MinData,sampfreq); lag(sampfreq+1);
11 3 rd Central Moment	moment(MinData(isnan(MinData)~=1),3);
11 4 th Central Moment	moment(MinData(isnan(MinData)~=1),4);
11 Number of Peaks	length(findpeaks(MinData, 'Threshold',1e-4,'MinPeakHeight', mean(MinData) + (max(MinData)-min(MinData))/4));
11 Number of Prominent Peaks	length(findpeaks(MinData, 'Threshold',1e-6,'MinPeakProminence', (max(MinData)-min(MinData))/4));
11 entropy	entropy(MinData);
11 Number of Zero-Crossings	C = midcross(MinData(isnan(MinData)~=1),sampfreq); length(C);
11 Mean Time between adjacent Zero-Crossings	if size(C,1) < 2; 60; else; mean(diff(C)); end
11 Median Time between adjacent Zero-Crossings	if size(C,1) < 2; 60; else; median(diff(C)); end
11 SD of the Time between adjacent Zero-Crossings	if size(C,1) < 2; 0; else; std(diff(C)); end
11 Number of Median-Crossings	zci = @(MinData) find(MinData(-.*circshift(MinData(-), [-1 0]) <= 0); C = zci(MinData); length(C);
11 Mean Time between adjacent Median-Crossings	if size(C,1) < 2; 60; else; mean(diff(C)); end
11 Median Time between adjacent Median-Crossings	if size(C,1) < 2; 60; else; median(diff(C)); end
11 SD of Time between adjacent MedianCrossings	if size(C,1) < 2; 0; else; std(diff(C)); end
3 Dynamic Time Warping (DTW) between Axes	dtw(MinData(:,1), MinData(:,2)); % for x-y, (:,1) and (:,3) for x-z, (:,2) and (:,3) for y-z
3 DTW between Signal Changes from Frame to Frame	dtw(diff(MinData(:,1)), diff(MinData(:,2))); % for x-y, (:,1) and (:,3) for x-z, (:,2) and (:,3) for y-z
3 Covariance between axes	CovTemp = nancov(MinData(:,1:3)); CovTemp(1,2) % for x-y; CovTemp(1,3) % for x-z; CovTemp(2,3) % for y-z;
3 Correlation between axes	corr(MinData(:,1),MinData(:,2)); % for x-y, (:,1) and (:,3) for x-z, (:,2) and (:,3) for y-z
1 Daytime	TIMESINCEFIRSTDAY(frameID,1) - floor(TIMESINCEFIRSTDAY(frameID,1));
11 SD of all non-overlapping 5 Seconds Mean	for i = 1:12; TempMean(i) = nanmean(MinData((i-1)*150+1:(i-1)*150+150,:)); end; std(TempMean)
11 SD of all non-overlapping 5 Seconds CV	for i = 1:12; TempStd(i) = nanstd(MinData((i-1)*150+1:(i-1)*150+150,:)); TempCV(i) = TempStd(i) ./ TempMean(i); end; std(TempCV)
Frequency Domain	
11 Mean Frequency *	MeanFreq = meanfreq(MinData,sampfreq);
11 Power at Mean Frequency ±0.1Hz	L = [MeanFreq-0.1 MeanFreq+0.1]; if L(1) < 0; L(2) = L(2)+abs(L(1)); L(1) = 0; end; if L(2) > 15; L(1) = L(1) - (L(2)-15); L(2) = 15;end; bandpower(MinData,sampfreq,L);
11 Median Frequency *	MedFreq = medfreq(MinData,sampfreq);
11 Power at Median Frequency ±0.1Hz	L = [MedFreq-0.1 MedFreq+0.1]; if L(1) < 0; L(2) = L(2)+abs(L(1)); L(1) = 0; end; if L(2) > 15; L(1) = L(1) - (L(2)-15); L(2) = 15;end; bandpower(MinData,sampfreq,L);
11 Mean Frequency between 0.3 to 3Hz *	MeanFreqLow = meanfreq(MinData,sampfreq,[0.3 3]);
11 Power at Mean Frequency ±0.1 Hz between 0.3 to 3Hz	L = [MeanFreqLow-0.1 MeanFreqLow+0.1]; if L(1) < 0; L(2) = L(2)+abs(L(1)); L(1) = 0; end; if L(2) > 15; L(1) = L(1) - (L(2)-15); L(2) = 15;end; bandpower(MinData,sampfreq,L);
11 Median Frequency between 0.3 to 3Hz *	MedFreqLow = medfreq(MinData,sampfreq,[0.3 3]);
11 Power at Median Frequency ±0.1Hz between 0.3 to 3 Hz	L = [MedFreqLow-0.1 MedFreqLow+0.1]; if L(1) < 0; L(2) = L(2)+abs(L(1)); L(1) = 0; end; if L(2) > 15; L(1) = L(1) - (L(2)-15); L(2) = 15;end; bandpower(MinData,sampfreq,L);
11 Total Signal Power	bandpower(MinData,sampfreq,[0 15]);
11 Power below 0.3 Hz	bandpower(MinData,sampfreq,[0 0.3]);
11 Power between 0.3 and 3 Hz	bandpower(MinData,sampfreq,[0.3 3]);
11 Power above 3 Hz	bandpower(MinData,sampfreq,[3 15]);
11 Harmonic Power *	[~,harmpow,~] = thd(MinData,sampfreq); harmpow(1);
11 Harmonic Frequency *	[~,~,harmfreq] = thd(MinData,sampfreq); harmfreq(1);

Supporting Information 3: Cross-validity table for all optimized and existing methods to detect sitting, standing, and being active, including cut-off for the cut-off based methods (in counts-per-minute (cpm) and steps per minute (spm)). The balanced sensitivity and specificity (Balanced) is the mean of sensitivity and specificity over the indicated/all posture. Data analysed on a subject-by-subject level and averaged over all subjects with median and non-parametric 95% confidence interval in brackets (after rejecting normal distribution with Lilliefors test). The activPAL served as reference criterion.

	Cut-Off		Overall		Sitting		Standing			Active		
	Sitting	Standing	Balanced	Balanced	Sensitivity	Specificity	Balanced	Sensitivity	Specificity	Balanced	Sensitivity	Specificity
ML Algorithm	-	-	90.4 [87.9 - 92.4]	87.8 [84.0 - 90.7]	95.6 [94.7 - 97.2]	79.6 [74.0 - 85.2]	85.2 [79.8 - 87.6]	74.8 [65.5 - 78.8]	96.1 [95.0 - 97.4]	99.2 [98.9 - 99.5]	98.4 [97.9 - 99.1]	99.9 [99.9 - 100.0]
Y_{cpm}	< 16 cpm	< 403 cpm	76.9 [74.5 - 78.0]	71.1 [66.9 - 73.2]	72.0 [67.3 - 77.7]	68.6 [63.1 - 77.7]	63.9 [61.6 - 66.5]	53.6 [47.4 - 58.9]	75.9 [72.1 - 80.2]	96.9 [95.8 - 97.5]	96.3 [93.2 - 97.8]	97.6 [96.9 - 98.5]
Y_{cpm(LFE)}	< 23 cpm	< 398 cpm	76.7 [74.5 - 78.4]	71.4 [67.1 - 73.7]	71.8 [66.8 - 76.1]	71.8 [66.3 - 80.2]	63.8 [60.5 - 66.2]	54.9 [48.6 - 60.3]	74.8 [71.8 - 78.8]	96.6 [96.0 - 97.3]	97.2 [95.4 - 98.9]	97.0 [95.9 - 97.9]
VM_{cpm}	< 69 cpm	< 1379 cpm	76.6 [74.1 - 77.7]	69.8 [65.8 - 71.8]	71.7 [69.3 - 77.4]	66.4 [55.5 - 74.1]	62.8 [60.6 - 65.2]	51.2 [41.3 - 58.7]	76.0 [72.6 - 80.8]	97.8 [97.2 - 98.3]	97.3 [96.0 - 98.7]	98.5 [98.1 - 98.7]
VM_{cpm(LFE)}	< 170 cpm	< 1484 cpm	75.9 [73.2 - 76.9]	69.0 [64.4 - 72.1]	76.9 [74.5 - 82.9]	59.6 [49.4 - 67.9]	61.1 [57.9 - 62.4]	39.9 [33.1 - 51.0]	81.2 [77.7 - 84.6]	97.8 [97.5 - 98.3]	97.8 [96.8 - 99.0]	98.2 [97.6 - 98.4]
Step	< 3 spm	< 11 spm	70.7 [69.7 - 72.3]	61.6 [59.9 - 66.5]	95.2 [94.6 - 96.1]	29.8 [25.2 - 40.9]	51.9 [51.2 - 52.7]	8.0 [7.0 - 11.3]	96.0 [95.5 - 96.7]	98.6 [98.1 - 99.1]	97.7 [96.3 - 98.8]	99.7 [99.6 - 99.8]
Step_{LFE}	< 5 spm	< 42 spm	75.5 [72.6 - 78.1]	66.5 [62.2 - 71.6]	76.8 [73.7 - 80.8]	57.7 [49.5 - 66.5]	61.2 [56.9 - 63.3]	43.4 [35.2 - 49.6]	79.6 [77.5 - 83.1]	99.4 [99.2 - 99.7]	99.4 [98.9 - 99.8]	99.8 [99.4 - 99.8]
Y_{cpm}	< 100 cpm	-	-	67.8 [64.3 - 72.3]	90.7 [88.7 - 92.7]	45.3 [38.0 - 55.6]	-	-	-	-	-	-
Y_{cpm(LFE)}	< 100 cpm	-	-	70.1 [65.7 - 73.7]	87.1 [84.5 - 89.7]	54.1 [44.7 - 63.2]	-	-	-	-	-	-
Y_{cpm}	< 150 cpm	-	-	66.6 [63.2 - 71.4]	94.2 [93.0 - 95.3]	39.3 [32.0 - 49.5]	-	-	-	-	-	-
Y_{cpm(LFE)}	< 150 cpm	-	-	68.5 [65.0 - 72.8]	91.9 [90.7 - 93.7]	45.2 [37.9 - 57.1]	-	-	-	-	-	-
Inclinometer	-	-	-	33.8 [29.6 - 43.9]	27.4 [23.4 - 32.4]	44.5 [33.6 - 58.2]	47.3 [45.6 - 48.1]	0.9 [0.3 - 2.9]	90.5 [86.4 - 93.9]	-	-	-
Inclinometer_{LFE}	-	-	-	33.5 [29.4 - 43.7]	27.5 [23.5 - 32.5]	43.8 [33.5 - 57.7]	47.3 [45.6 - 48.1]	0.9 [0.3 - 2.9]	90.5 [86.3 - 93.9]	-	-	-

Abbreviations: machine learning (ML), vertical axis (y), counts-per-minute (cpm), low-frequency-extension (LFE), vector magnitude (VM), steps-per-minute (spm)